Published in final edited form as:

Biometrics. 2017 June ; 73(2): 401–409. doi:10.1111/biom.12575.

A Framework for Bayesian Nonparametric Inference for Causal effects of Mediation

Chanmin Kim^{1,*}, Michael J. Daniels², Bess H. Marcus³, and Jason A. Roy⁴

¹Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115

²Department of Statistics and Data Sciences & Department of Integrative Biology, The University of Texas, Austin, TX 78712

³Department of Family Medicine and Public Health, University of California, San Diego, CA 92093

⁴Department of Biostatistics, University of Pennsylvania, Philadelphia, PA 19104

Summary

We propose a Bayesian non-parametric (BNP) framework for estimating causal effects of mediation, the natural direct and indirect effects. The strategy is to do this in two parts. Part 1 is a flexible model (using BNP) for the observed data distribution. Part 2 is a set of uncheckable assumptions with sensitivity parameters that in conjunction with Part 1 allows identification and estimation of the causal parameters and allows for uncertainty about these assumptions via priors on the sensitivity parameters. For Part 1, we specify a Dirichlet process mixture of multivariate normals as a prior on the joint distribution of the outcome, mediator and covariates. This approach allows us to obtain a (simple) closed form of each marginal distribution. For Part 2, we consider two sets of assumptions: (a) the standard sequential ignorability (Imai et al., 2010) and (b) weakened set of the conditional independence type assumptions introduced in Daniels et al. (2012) and propose sensitivity analyses for both. We use this approach to assess mediation in a physical activity promotion trial.

Keywords

Causal inference; Dirichlet process mixture; Sensitivity Analysis

1. Introduction

Inference for causal effects are of interest in many fields. Social and behavioral researchers, in particular, are interested in situations where potential mediators are present on the causal pathway, and want to assess the effect of an intervention on outcome of interest both through and around the mediator; the former is an indirect effect and the latter is a direct effect in

* ckim@hsph.harvard.edu.

Supplementary Materials

Web appendices, Tables, and Figures (which further describe model specification, posterior computation, simulation set-up, sensitivity analyses, and model checking) referenced in Sections 2.2, 3.2, 3.3, 4, 5, 6, 7, 8 along with the code are available with this paper at the *Biometrics* website on Wiley Online Library.

mediation analysis. To define the causal effects, we use a potential outcome framework (Rubin, 1974). Assuming SUTVA (Stable Unit Treatment Value Assumption; Rubin, 1980), for a vector of randomized interventions $\mathbf{Z} \in \{0, 1\}^{\otimes n}$, the random variable $M_{i,z}$, which denotes the value of the mediator M_i for the i th subject that would have been observed had a vector of interventions $\mathbf{Z}$ been set to $\mathbf{z}$, is equal to M_{i,z_i} (i.e., potential values of the mediator do not vary with interventions assigned to others) and the random variable $Y_{i,z\mathbf{m}}$, which denotes the value of the outcome for the i th subject that would have been observed had a vector of interventions $\mathbf{Z}$ been set to $\mathbf{z}$ and values of the mediator $\mathbf{M}$ been set to $\mathbf{m}$, is equal to Y_{z_i,m_i} (i.e., potential values of the outcome do not vary with interventions and values of the mediator for others). Throughout the paper, we suppress the index i for notational simplicity and focus on assessing the natural effects (Robins and Greenland, 1992; Pearl, 2001). The natural indirect effect on the population level is defined as $E[Y_{1,M_1} - Y_{1,M_0}]$ and the natural direct effect on the population level as $E[Y_{1,M_0} - Y_{0,M_0}]$, where we use the notation, $Y_{z,M_z'}$, the value of the outcome that would have been observed if an individual had been assigned to intervention z with (possibly hypothetically) mediator M set to its value under z' . These are distinguished from controlled effects (Pearl, 2001) which define the causal effects while intervening to fix the mediator M to a value m . For example, a controlled direct effect is defined by $E[Y_{1,m} - Y_{0,m}]$.

There are a large number of methods for mediation analysis in the literature proposed from a frequentist perspective (Pearl, 2001; MacKinnon et al., 2002; Robins, 2003; Preacher and Hayes, 2004; Petersen et al., 2006; VanderWeele, 2009; Imai et al. 2010, Albert and Nelson, 2011; Tchetgen Tchetgen et al. 2012; Valeri and VanderWeele, 2013), with fewer approaches using Bayesian inference. The latter include Yuan and MacKinnon (2009), Elliott et al. (2010), Schwartz et al. (2011), Daniels et al. (2012) and Mattei et al. (2013). Yuan and MacKinnon (2009), who focus on the standard structural equation model framework (Baron and Kenny, 1986), compute the posterior distribution of products of coefficients, and infer causal effects of interest. Elliott et al. (2010) develop an approach for estimating nonparametric bounds of principal strata causal effects of a dichotomous mediator and a dichotomous outcome by using prior distributions over a possible range of values. Schwartz et al. (2011) use a Bayesian nonparametric model, a Dirichlet process mixture model, to construct the distribution of principal strata of continuous mediators. Their model identifies the strata of continuous mediators and explores the latent structure of the data automatically. Mattei et al. (2013) develop a Bayesian principal stratification inference method for multiple outcomes not based on structural assumptions but based on flexible distributional assumptions. However, in the last three papers, the causal effects are defined within principal strata, so-called principal causal effects, and are different from what we estimate, the natural direct and indirect effects. In addition, Bayesian nonparametric models have not been used in context of estimating natural direct and indirect effects.

For estimation and identification, we approach the problem in two parts. The first part is a flexible model for the observed data. For this, we specify a Bayesian nonparametric model (BNP), in particular, a multivariate Dirichlet process mixture of normals; this will provide flexibility of modeling as well as computational ease. However, other choices are possible.

We are careful to specify models so that the total effect is invariant to the uncheckable assumptions, as the total effect is estimable for randomized interventions.

The second part is a set of uncheckable assumptions that provide identification of the causal effects of interest. For this, we propose two different sets of assumptions. The first is (the standard) sequential ignorability (Imai et al., 2010). The second is a modification of the assumptions in Daniels et al. (2012) which are weakened here by conditioning on baseline covariates. Both of these assumptions have sensitivity parameters for which uncertainty about the (uncheckable from data) assumptions can be characterized via prior distributions. Also, we can infer covariate-specific causal effects, which may help researchers to understand the effect of the intervention among populations having different characteristics. Any set of identifying assumptions can be used in this framework.

Our approach differs from existing approaches by the way in which we completely separate the observed data model from uncheckable identifying assumptions. For example, VanderWeele (2010) reviews assumptions for the identification of natural direct and indirect effects and states that if those assumptions hold and (semi-) parametric regression models are correctly specified (which are part of uncheckable assumptions), then the natural direct and indirect effects are estimable. In our approach, we distinguish the observed data model from the uncheckable identifying assumptions. For the former we minimize parametric modeling assumptions; for the latter we allow informative priors to account for uncertainty about (necessary) uncheckable assumptions. Our approach is also different from Daniels et al. (2012) in that we accommodate all types of outcome, mediator and covariate variables via flexible, general BNP models while their approach is confined to the binary outcomes without covariates.

We apply our method to the STRIDE data (Marcus et al. 2007) which is a randomized clinical trial to evaluate the effect of (print- and telephone-based) interventions on physical activity. This study contains several measures of processes of change obtained from questionnaires. Our focus is on behavioral process as a potential mediator that can affect physical activity. The primary outcome is the amount of weekly moderate to vigorous physical activity minutes at 12 months. We assess the causal effects for different values of age and body mass index (BMI).

The remainder of the article is organized as follows. In Section 2, we define the causal effects of interest conditional on baseline covariates and propose a Bayesian nonparametric model for the observed data. Section 3 outlines two sets of assumptions, each sufficient for identification. Section 4 summarizes posterior computation. In Section 5, we outline strategies for sensitivity analysis. In Section 6, we conduct simulation studies to assess performance of the BNP model. In Section 7, we use our approach to estimate the causal effects in the STRIDE data.

2. Notation and definition of causal effects and specification of the observed data model

We define a $(q - 2)$ vector of baseline covariates, $\mathbf{X}$ which are completely observed. The binary treatment is denoted as Z . We define the potential mediator M_z as the value of the mediator if the subject receives intervention $Z = z$. Then, observed mediators are defined as $M_{\text{obs}} = ZM_1 + (1 - Z)M_0$. The potential outcome $Y_{z,M_z'}$ denotes the value of the outcome that Y under intervention $Z = z$ with mediator M set to the value that would be observed under $Z = z'$. Only one potential outcome can be observed, $Y_{\text{obs}} = ZY_{1,M_1} + (1 - Z)Y_{0,M_0}$.

We define natural direct and indirect effects conditional on baseline covariates $\mathbf{X} = \mathbf{x}$ as $NIE(\mathbf{x}) = E(Y_{1,M_1} - Y_{1,M_0} | \mathbf{x})$ and $NDE(\mathbf{x}) = E(Y_{1,M_0} - Y_{0,M_0} | \mathbf{x})$. The natural indirect effect, $NIE(\mathbf{x})$, quantifies the effect of the intervention through the mediator for a fixed value of covariates $\mathbf{X} = \mathbf{x}$. $NIE(\mathbf{x})$ corresponds to the arrows that flow from Z to M to Y in Figure 1.

The natural direct effect, $NDE(\mathbf{x})$, quantifies the effect of the intervention on the outcome by setting the mediator M to its natural value M_0 (the value of the mediator in the absence of the intervention) given a fixed value of the baseline covariates $\mathbf{X}$; this corresponds to the horizontal arrow from Z to Y in Figure 1. The total effect is the sum of two effects $TE(\mathbf{x}) = NIE(\mathbf{x}) + NDE(\mathbf{x}) = E(Y_{1,M_1} - Y_{0,M_0} | \mathbf{x})$. After integrating out the baseline covariates, we obtain the marginal causal effects NIE , NDE and TE .

In what follows, we first provide details on BNP models for the observed data (Section 2.1). Then, in Section 3, we provide two different sets of assumptions to identify the causal effects of interest given the observed data distribution.

2.1 A Bayesian Nonparametric Model for the Observed Data

The observed data models can be specified nonparametrically. Here, we specify Dirichlet process mixtures (DPM) of multivariate normals (Escobar and West, 1995; Muller et al., 1996; Jara et al., 2011) for the q -dimensional joint distribution of observed data Y_{obs} , M_{obs} and $\mathbf{X}$. For each intervention $Z = 0, 1$,

$$(Y_{\text{obs},i}^z, M_{\text{obs},i}^z, \mathbf{X}_i^z) \sim N_q(\boldsymbol{\mu}_{z,i}, \boldsymbol{\Sigma}_{z,i}), i=1, \dots, n_z,$$

$$(\boldsymbol{\mu}_{z,i}, \boldsymbol{\Sigma}_{z,i}) \sim G_z, i=1, \dots, n_z, \quad (1)$$

$$G_z \sim DP(\alpha_z \mathcal{G}_z),$$

where $(Y_{\text{obs},i}^z, M_{\text{obs},i}^z, \mathbf{X}_i^z)$ denotes the i -th observed data $(Y_{\text{obs}}, M_{\text{obs}}, \mathbf{X})$ under intervention z and n_z is the number of observations under intervention z . The base distribution $\mathcal{G}_z$ is taken to be the conjugate normal-inverse-Wishart distribution (NIW),

$N(\boldsymbol{\mu}_z; \mathbf{m}_z, \kappa_0^{-1} \boldsymbol{\Sigma}_z) \mathcal{W}^{-1}(\boldsymbol{\Sigma}_z; \nu_z, \boldsymbol{\Psi}_z)$. The inverse-Wishart is parameterized such that

$E[\Sigma_z] = \Psi_z^{-1} / (v_z - q - 1)$. For the hyper-priors, we follow the similar specification suggested by Taddy (2008) and Jara et al. (2011). See the supplementary materials for more details. We specify a Gamma prior $G(1, 1)$ on the mass parameter α_z . Though the results from the BNP model are generally not sensitive to reasonable changes in the (hyper-) prior specifications (Taddy, 2008), we still conduct sensitivity analyses for different specifications in the supplementary materials. When the outcome, mediator and/or covariates are binary or ordinal, we can introduce normal latent variables (Johnson and Albert, 1999) and proceed similarly (see the supplementary materials for details).

3. Identifying assumptions and inference on causal effects

Given the observed data distribution, we present two sets of assumptions both of which are sufficient to estimate the causal effects of interest. Versions of both sets of assumptions (some stronger) have been proposed in the literature. The key point is that the same model for the observed data, specified in Section 2, can be combined with either of these sets of assumptions. In that way, we separate the modeling of the observed data from the causal identifying assumptions.

3.1 First Set of Assumptions (Sequential Ignorability)

We first infer causal effects of mediation based on the sequential ignorability assumption. Let $\mathbf{W}$ be a set of baseline covariates, assumed to be all confounders among variables. Following the definition from Imai et al. (2010):

$$\{Y_{z',m}, M_z\} \perp\!\!\!\perp Z | \mathbf{W} = \mathbf{w} \quad (2)$$

$$Y_{z',m} \perp\!\!\!\perp M_z | Z = z, \mathbf{W} = \mathbf{w}, \quad (3)$$

for $z, z' \in \{0, 1\}$. It is also assumed that $0 < p(Z = z | \mathbf{W} = \mathbf{w})$ and $0 < p(M_z = m | Z = z, \mathbf{W} = \mathbf{w})$ for $z, z' \in \{0, 1\}$, all $\mathbf{w} \in \mathcal{W}$, and $m \in \mathcal{M}$. The set of confounders, $\mathbf{W}$, may or may not be the same as the baseline covariates $\mathbf{V}$ introduced in the next section. For the observed data model in the previous section, we need to fit the model for $(Y_{z,M_z}, M_z, \mathbf{X})$ where $\mathbf{X} = \mathbf{W} \cup \mathbf{V}$.

Note that this assumption must hold without exposure induced outcome and mediator confounders (i.e., none of the mediator-outcome confounders are affected by exposure). See the supplementary materials for details on modifying our approach to accommodate this setting. It is also worth noting that under a randomized intervention the first step of the sequential ignorability holds unconditional on covariates $\mathbf{W} = \mathbf{w}$.

In the STRIDE example, the first step of the sequential ignorability holds by the experimental design since participants were randomized to the intervention group. The second step of the sequential ignorability states that given pretreatment covariates, behavioral processes under the treatment status z are independent of the average physical

activity minutes under the treatment status z' , possibly $z = z'$, with behavioral processes set to m which has been called cross-world independence (Richardson and Robins, 2013).

3.2 Second Set of Assumptions (Mediator Induction Equivalence)

We also consider an alternative set of assumptions to identify the natural direct effect and indirect effect. We call this set the ‘mediator induction equivalence’ assumptions based on the key assumption (Assumption 1 in what follows). The proposed assumptions are weaker than previous assumptions proposed in Daniels et al. (2012) since they now only hold conditional on baseline covariates $\mathbf{V} = \mathbf{v}$. Again, we point out that the baseline covariates $\mathbf{W}$ in the first set of assumptions and the baseline covariates $\mathbf{V}$ in the second set of assumptions are different in their use. For sequential ignorability in Section 3.1, covariates $\mathbf{W}$ are incorporated to adjust for confounding relationships among variables, especially between the mediator and the outcome. However, for mediator induction equivalence in Section 3.2, covariates $\mathbf{V}$ are incorporated to weaken the identifying assumptions proposed in Daniels et al. (2012), which are based on the conditional distributions of the potential outcomes.

The following set of assumptions will identify the parameters of the the potential outcomes and mediators distributions, $f(y_{1,M_1}|m_0, m_1, \mathbf{v})$, $f(y_{1,M_0}|m_0, m_1, \mathbf{v})$, $f(y_{0,M_0}|m_0, m_1, \mathbf{v})$ and $f(m_0, m_1|\mathbf{v})$, which are sufficient to estimate the NIE and NDE.

Assumption 1—

$$f(y_{1,M_0}|M_0=m, M_1, \mathbf{V}=\mathbf{v}) = f(y_{1,M_1}|M_0, M_1=m, \mathbf{V}=\mathbf{v}),$$

the conditional distribution of the outcome is the same whether the corresponding mediator was induced by $Z = 1$ or $Z = 0$. This assumption equates the conditional distribution of unobservable potential outcome Y_{1,M_0} to the conditional distribution of observable outcome Y_{1,M_1} if they are conditioning on the same value m for mediators M_0 and M_1 , respectively. It is worth noting that this assumption implies $f(y_{1,M_0}|M_0 = m, \mathbf{V} = \mathbf{v}) = f(y_{1,M_1}|M_1 = m, \mathbf{V} = \mathbf{v})$.

Under Assumption 1, the potential value of the outcome under the intervention with realized mediator value m , $Y_{1,m}$, is the same regardless of from which treatment the value for the mediator, m , was induced. Assumption 1 is plausible if Y_{1,M_Z} is only affected by the mediator under the same treatment M_Z not by the mediator under the opposite treatment M_{1-Z} after conditioning on baseline covariates $\mathbf{V}$. Thus, Y_{1,M_1} and Y_{1,M_0} would result in the same expected response $Y_{1,m}$ if the mediator value m is the same. With the mediator M_Z and the covariates $\mathbf{V}$ capturing all the information of the mediator value under the opposite treatment, M_{1-Z} , this assumption can be more plausible. This assumption is somewhat related to sequential ignorability assumption if we assume the same baseline covariates (which is hard to assume in practice given their different roles); details can be found in the supplementary materials.

In STRIDE, we assume the conditional distributions of the average activity minutes under the intervention $Z = 1$ and the control $Z = 0$ are the same conditional on the same value m

for the mediator, behavioral processes, and the same values for age, BMI and baseline behavioral processes.

Since the potential mediators, M_0 and M_1 , cannot be observed simultaneously, we need an assumption to identify their joint distribution.

Assumption 2—The joint distribution of mediators conditional on baseline covariates is assumed to follow a Gaussian copula model (Nelsen, 1999) with rank correlation ρ between mediators. Note that this construction of the joint distribution of the mediators puts no restrictions on the models for the marginals.

3.3 Estimation of NIE and NDE

Recall that the set of the baseline covariates in the observed data model X denotes the union of two sets of covariates used in Section 3.1 and Section 3.2, $W \cup V$. By the randomization of the treatment, the distributions $f(x)$, $f(m_Z|x)$, $f(y_{Z,M_Z}|m_Z, x)$, and $f(y_{Z,M_Z})$ are estimable from $(Y_{\text{obs}}, M_{\text{obs}}, X)$.

With the first set of assumptions (sequential ignorability), we can nonparametrically estimate the $NIE(w)$ and $NDE(w)$ for some covariates $W = w$ as well as the average natural indirect and direct effect of the population. The proof follows trivially from Imai et al (2010), so we omit the details.

With the mediator induction equivalence assumptions, we can estimate the distribution of *a priori counterfactual* Y_{1,M_0} based on the identified parameters of the observed data model. Then, we estimate the $NIE(v)$ and $NDE(v)$ for covariates $V = v$ as well as the average natural indirect and direct effects. See the next section and the supplementary materials for details.

In general, inference on the NIE and NDE is driven by the specific uncheckable assumptions made and not by the observed data model (due to its flexibility).

4. Posterior Computation

To draw posterior samples from the parameters in observed data model, we use the algorithms summarized in Neal (2000) (Algorithm 8; see the supplementary materials). Since the NIE, NDE and TE are functions of the parameters of the observed data models, for each posterior sample of those parameters, we do several post-processing steps by plugging samples into the functions to obtain the posterior distributions of the NIE, NDE and TE and the corresponding point estimates and credible intervals. See the supplementary materials for the complete algorithm.

5. Sensitivity Analysis

In this section, we discuss a new strategy to assess sensitivity to violations of (3) in the sequential ignorability assumption and a default strategy for the sensitivity parameter ρ introduced in Assumption 2 in the mediator induction equivalence assumptions.

Under the sequential ignorability set of assumptions, (2) is guaranteed in our setting of a randomized experiment. However, (3) is based on an untestable, no unobserved confounding relationships between the mediator and the outcome. Although a few approaches to sensitivity analysis have been proposed (Imai et al, 2010; Albert and Wang, 2014; VanderWeele and Chiba, 2014), they are restricted to settings, such as linear models (or generalized linear models), that do not apply to our setting. We propose a simple strategy next.

To introduce a method for sensitivity to (3), we first need to recap identification results from Imai et al. (2010) which are re-stated in the supplementary materials. If we let $g_0(m, \mathbf{w}) = E(Y_{1,m} | M_0 = m, Z = 0, \mathbf{W} = \mathbf{w}) - E(Y_{1,m} | Z = 0, \mathbf{W} = \mathbf{w})$ and $g_1(m, \mathbf{w}) = E(Y_{1,m} | M_1 = m, Z = 1, \mathbf{W} = \mathbf{w}) - E(Y_{1,m} | Z = 1, \mathbf{W} = \mathbf{w})$, then $E(Y_{1,M_0} | \mathbf{W} = \mathbf{w})$ can be re-expressed without using (3) as follows,

$$E(Y_{1,M_0} | \mathbf{W} = \mathbf{w}) = \int \{g_0(m, \mathbf{w}) - g_1(m, \mathbf{w}) + E(Y | M_1 = m, Z = 1, \mathbf{W} = \mathbf{w})\} dF_{M|Z=0, \mathbf{W}=\mathbf{w}}(m),$$

where (3) implies $g_0(m, \mathbf{w}) = 0$ and $g_1(m, \mathbf{w}) = 0$ for all m and $\mathbf{w}$. Thus, $g_0(m, \mathbf{w})$ and $g_1(m, \mathbf{w})$ are sensitivity parameters.

To calibrate them, we note that, from (2), $g_0(m, \mathbf{w})$ and $g_1(m, \mathbf{w})$ can be re-expressed as

$$g_0(m, \mathbf{w}) = E(Y_{1,m} | M_0 = m, \mathbf{W} = \mathbf{w}) - E(Y_{1,m} | \mathbf{W} = \mathbf{w}), \quad (4)$$

$$g_1(m, \mathbf{w}) = E(Y_{1,m} | M_1 = m, \mathbf{W} = \mathbf{w}) - E(Y_{1,m} | \mathbf{W} = \mathbf{w}). \quad (5)$$

We calibrate $g_0(m, \mathbf{w})$ and $g_1(m, \mathbf{w})$ by first computing the total amount of variability of the outcomes under $Z = 0$ and $Z = 1$ explained by $\mathbf{W}$. For $g_1(m, \mathbf{w})$ in (5), we estimate the coefficient of determination among the treated ($Z = 1$) with $\mathbf{W}$ as covariates (but not M_1) and denote it as $R_1(\mathbf{w})$. Then, we set the absolute value of the difference between conditional

expectations in (5), $|g_1(m, \mathbf{w})|$, to be less than $\sqrt{\text{var}(Y | Z=1) \times (1 - R_1(\mathbf{w}))} \times k_1$ where k_1 is a certain percent of the remaining total variance that is not explained by $\mathbf{W}$. With this specification, we restrict the additional contribution of M_1 to the outcome model under $Z = 1$ to be less than $k_1\%$ (sensitivity parameter) of the total variance that is not explained by $\mathbf{W}$.

The coefficient of determination $R_1(\mathbf{w}) = 1 - \frac{\sum z_i (y_i - f_i)^2}{\sum z_i (y_i - \bar{y}_i)^2}$ is calculated using the posterior samples of f_i , the expected value of y_i conditional on the mediator and covariates.

Similarly for (4), we might expect $|g_0(m, \mathbf{w})|$ to be less than

$\sqrt{\text{var}(Y | Z=1) \times (1 - R_1(\mathbf{w}))} \times k_0$. As such, we set k_0 since we expect that M_0 does not explain as much of the variance of $Y_{1,m}$ as M_1 . Thus, we have two sensitivity parameters, k_0 and k_1 , bounded in the unit square.

For the mediator induction equivalence set of assumptions, we assume the (rank) correlation between M_0 and M_1 , ρ , in Assumption 2 is same across all values of $\mathbf{V}$. Under the typical assumption of positive correlation, a default approach would be to specify $\rho \in [0, 1)$. Note that, from the post-processing steps for the mediator induction equivalence (in the supplementary materials), it is observed that the posterior means of the indirect and direct effects do not depend on ρ . We can also examine sensitivity towards violation of Assumption 1; we relegate the details of this sensitivity analysis to the supplementary materials and note that the NIE and NDE are effected by ρ in that case.

6. Simulation

To assess the performance of the proposed BNP model, we conduct simulation studies under sequential ignorability. Here, we compare the BNP model with a standard Bayesian parametric model and a frequentist semiparametric model (Imai et al., 2010, Tingley et al., 2014); for the latter we use the R package `mediation` (CRAN). A simple Bayesian parametric model for the joint distribution (of the data under $Z = 0, 1$) is specified as follows,

$$(Y_{\text{obs},i}^z, M_{\text{obs},i}^z, X_{1,i}^z, X_{2,i}^z) \sim N_4(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z), i=1, \dots, n_z,$$

$$(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z) \sim N(\boldsymbol{\mu}_z; \mathbf{m}_z, \kappa_0^{-1} \boldsymbol{\Sigma}_z) \mathcal{W}^{-1}(\boldsymbol{\Sigma}_z; \nu_z, \boldsymbol{\Psi}_z), \quad (6)$$

where all hyper-priors are specified as in Section 2.1. For the frequentist semiparametric model, we specify the mediator and outcome models as $E(M_{\text{obs}}) = \alpha + f_1(Z) + f_2(X_1) + f_3(X_2)$, $E(Y_{\text{obs}}) = \beta + g_1(Z) + g_2(M) + g_3(X_1) + g_4(X_2)$, respectively, where Generalized Additive Models (GAM) are used to fit the Y_{obs} and M_{obs} models with cubic splines $f(\cdot)$ and $g(\cdot)$. To fit those, we use the `mediate()` function with setting `boot = TRUE` for the nonparametric bootstrap to calculate uncertainty estimates.

We now compare these models under three cases all of which have the sample size $n = 240$ (the STRIDE data size) where half of the sample receives the treatment ($Z = 1$) and two covariates are distributed as $W_1 \sim N(45.25, 9.61)$ and $W_2 \sim N(28.55, 5.56)$. The distributions of covariates are based on the STRIDE data. Three cases are generated as follows: (1) Case 1. Mediator (skew normal) and Outcome (Mixture of Normals); (2) Case 2. Mediator (skew normal) and Outcome (Mixture of Normals with Interaction); (3) Case 3. Mediator (skew normal) and Outcome (Normal with nonlinear terms). See the supplementary materials for the detailed specification.

Table 1 shows results of the simulation study based on 500 replications. The results show that all of the estimators from frequentist and BNP models are approximately unbiased for Case 1 (all absolute biases are < 0.05). Also, MSEs are approximately the same for both models.

For Case 2, the Bayesian nonparametric model outperforms both the Bayesian parametric and frequentist semiparametric models in terms of bias and MSE. The frequentist semiparametric model performs especially poorly for Case 2 since its semiparametric regression structure does not include an interaction between Z and M ; this approach highly depends on the assumed model structure. MSE for the NIE is the smallest while bias for the NIE is the largest because the frequentist model estimates the NIE incorrectly but with less variability. If we add a spline term for interaction $Z \times M$, we get more biased results (biases are -1.26 , -0.58 , -1.84 for the NIE, NDE and TE, respectively). On the contrary, the BNP model shows the stable performance regardless of the presence of an interaction term.

In the setting of the nonlinear terms (Case 3), the estimates from the BNP and the frequentist models are approximately unbiased while the parametric model produces relatively biased estimates for the NIE and NDE. Overall, the BNP model performs well for data having non-standard distributions and/or non-linear terms.

Table 1 contains the average computation time of each model for each replication. For the Bayesian parametric model, we use our own Stan (Stan Development Team, 2014) code with 40,000 iterations (20,000 burn-in). For the BNP model, we use our own R package, `BNPMediation` (available in GitHub), implemented using `DPpackage` (Jara et al., 2011) with 40,000 iterations. The BNP model requires slightly more computation time (based on Mac OS $\times$ with 2.4 GHz Intel Core i5 processor and 8 GB RAM) in comparison to the frequentist model; however, the time is not prohibitively larger.

We also examined sensitivity of results to specification of the hyper-priors and the effectiveness of the sensitivity analysis approach proposed for the sequential ignorability assumption. Results were relatively insensitive to the specification of the hyper-priors, even under relatively diffuse priors. In addition, we demonstrate how the sensitivity analysis approach can result in credible intervals that cover the true causal parameters in the presence of unmeasured confounding. Further details can be found in the supplementary materials.

7. Application

7.1 STRIDE project

STRIDE (Marcus et al., 2007) was a randomized clinical trial to evaluate the effectiveness of interventions delivering individually tailored messages in increasing physical activity among sedentary adults. Participants ($N = 239$) were randomized to one of 3 treatment arms: (1) telephone-based intervention; (2) print-based intervention; or (3) contact control. Since the telephone- and print-based interventions were each delivering the same theory-based, individually tailored physical activity intervention (but just delivered via different channels), we combine telephone-based and print-based interventions into a single intervention group ($n_1 = 161$) versus the contact control group ($n_0 = 78$). Physical activity questionnaires were used to obtain secondary outcomes such as behavioral processes (e.g., enlisting social support), cognitive processes (e.g., seeking out information on physical activity), decisional balance (i.e., weighing the pros and cons of adopting physical activity) and self-efficacy; these were collected at baseline, 6 and 12 months. In the original study (Marcus et al., 2007), these secondary outcomes measured at 6 months were the proposed mediators. The primary

outcome was the minutes per week of moderate to vigorous intensity physical activity at 12 months. We examine ‘behavioral processes’ at 6 months as a potential mediator in our analysis. See the supplementary materials for how behavioral processes were measured.

Several characteristics of subjects were collected at baseline including age, gender, race, marital status, educational attainment, cigarette use and body mass index. We consider age, body mass index (BMI) and baseline behavioral processes as covariates of interest, $\mathbf{V}$, in the analysis since, among available demographic variables, baseline BMI is highly correlated with individuals’ physical activity minutes (Sevick et al., 2007) and age and BMI are the top two demographic variables to predict stage of adoption (including behavioral processes) in the questionnaire design (Marcus et al, 1992). Baseline behavioral processes explains the mediator (behavioral processes measured at 6 months) the most. Note that covariates, $\mathbf{V}$, are used to weaken the mediator induction equivalence assumptions and to define conditional causal effects.

In the sequential ignorability assumptions, since it is randomized study, we only need to adjust for confounding between the behavioral processes (mediator) and physical activity minutes per week (outcome) under an assumption of no exposure induced confounding relationship. We include baseline cognitive processes and baseline self efficacy which are all moderately correlated with behavioral processes and the outcome (Marcus et al., 2007b) as well as age, BMI and baseline behavioral processes to adjust for confounding in (3).

Fifty-one participants had incomplete information on the primary outcome and/or mediator. For these, we (implicitly) assume ignorable missingness. Missing outcomes and mediators are iteratively sampled via a Monte Carlo Markov chain (MCMC) data augmentation algorithm conditional on observed data and parameters. Finally, we log transformed the outcome due to skewness and to avoid placing mass at negative values.

7.2 Priors and Sensitivity Parameter

For the Dirichlet process mixture model in (1), we set $v_z = v_\Psi = 16 (= 2(q + 1))$ and $\tau_1 = 10$, $\tau_2 = 5$ in the base distribution as stated in Section 2.1. Other (hyper)priors were specified in Section 2.1 and the supplementary materials. For the sensitivity parameter ρ in Assumption 3 of the mediator induction equivalence set of assumptions, we specify a $\text{Unif}(0, 1)$ prior. For the sensitivity analysis in the sequential ignorability assumptions, we set $k_0 \sim \text{Unif}(0, 10)$ and $k_1 \sim \text{Unif}(k_0, 20)$. See Section 7.3 for details.

To illustrate the need for nonparametric models for the observed data in STRIDE, we show improved fit of the BNP model vs. the parametric model in the supplementary materials.

7.3 Results

For the observed data posterior, we ran 50000 MCMC iterations and used the last 10000 iterations, with thinning of 5 to obtain $N = 2000$ posterior samples of $\{(\boldsymbol{\mu}_{z, c_i}, \boldsymbol{\Sigma}_{z, c_i}); z = 0, 1, i = 1, \dots, n\}$. Details on monitoring convergence of the posterior samples from the base distributions is provided in the supplementary materials.

Table 2 contains the total effect and natural direct and indirect effects for typical settings of baseline covariate values (age and BMI) under both (1) sequential ignorability and (2) mediator induction equivalence. The total effects have significant effects (i.e., 95% credible intervals exclude 0) for all covariates values considered under both sets of assumptions except for age=55 and BMI=35. The natural indirect effects have 95% C.I. that cover zero for almost all scenarios of older subjects (age=55) under both sets of assumptions. It is worth noting that the natural indirect effect (the natural direct effect) increases (decreases) as BMI increases.

The NDE for a given set of covariates values is the average difference in moderate intensity physical activity minutes per week (on a log scale) at the end of the study between the intervention ($z = 1$) and the control ($z = 0$), if the subject's behavioral processes were set to what it would have been in the absence of the intervention. Hence, statistically significant positive values of NDE suggest that the effect of the intervention on moderate intensity physical activity that is not through its effect on behavioral processes is positive under either set of assumptions.

The NIE for a given set of covariates values is the average minutes of moderate intensity physical activity per week (on the log scale) if subjects were in the intervention arm, compared with the average minutes of moderate intensity physical activity per week if subjects were also in the intervention arm but their behavioral processes were set to what it would have been in the absence of the intervention. The 95% C.I.s including zero for the NIEs in Table 2 indicate that the effects of the intervention that are due to changes in behavioral processes are not statistically significant. Overall, when subjects have high BMIs, the telephone and print based intervention are more effective in terms of the average minutes change than those from subjects who have lower BMIs. With regard to age, older subjects with high BMI or younger subjects with low BMI are more likely to change their average minutes of moderate intensity physical activity under the intervention. Note that the TEs are invariant to the two sets of assumptions and the differences between estimates of the natural direct/indirect effects under mediator induction equivalence and under sequential ignorability are small here. In the supplementary materials, we also examine (marginal over covariates) causal effects under both sets of assumptions.

We consider deviations from sequential ignorability assumption in terms of several pairs of values for k_0 and k_1 (in Section 5), $\{(5, 5), (5, 10), (5, 15), (10, 10), (10, 15), (15, 20)\}$, under the constraint that $k_0 \leq k_1$. The above pairs are combinations of M_1 and M_0 explaining additional 5%, 10%, 15% or 20% of the total variance of the outcome that is not explained

by covariates. In Section 5, since we set $|g_z(m, \mathbf{w})| < \sqrt{\text{var}(Y|Z=1) \times (1 - R_1(\mathbf{w}))} \times k_z$ for each intervention $Z = z$, we separately consider positive and negative cases of $g_z(m, \mathbf{w})$. Table 3 contains the results under the different scenarios. The results indicate as k_0 and k_1 get closer (i.e., M_0 additionally explains as much of the total variance of the outcome as M_1) the NIE and the NDE converge to the values under the sequential ignorability assumptions. On the contrary, the more M_1 explains the total variance versus M_0 , the more the effects deviate from the effects under the sequential ignorability assumptions. Recall that the sequential ignorability implies $k_0 = k_1 = 0$. The estimates of the TEs are invariant for all

values of k_0 and k_1 . Table 3 also contains the results under uniform distributions, $\text{Unif}(0, 10)$ and $\text{Unif}(0, 20)$, on k_0 and k_1 , respectively, with the constraint that $k_0 \leq k_1$. The additional uncertainty about the magnitude of k_0 and k_1 , as quantified by non-degenerate priors, can be seen in terms of the increased lengths of the credible intervals under the uniform priors.

8. Discussion

In this paper, we have proposed a Bayesian nonparametric (BNP) approach for the causal effects of mediation. We use BNP models for the observed data and two different sets of assumptions to identify and estimate the causal effects. Our Bayesian nonparametric method for estimating causal effects provides greater flexibility than parametric methods while still maintaining computational ease in addition to allowing uncertainty about assumptions via priors on sensitivity parameters.

For the STRIDE data, the effect of the telephone and print-based intervention was marginally significant for all subjects and the NDE was statistically significant for all subjects. The NIE was significant under mediator induction equivalence assumption for subjects younger than 45.

We can weaken Assumption 1 of the mediator induction equivalence assumption by replacing it with the set of two alternative sub-assumptions as in Daniels et al. (2012). In that case, the posterior means of the NIE and NDE will depend on sensitivity parameter ρ in Assumption 2. See the supplementary materials for details.

Future work will consider extensions of the model to allow multiple mediators. Although we explore the relationship between two sets of assumptions, the sequential ignorability assumptions and the mediator induction equivalence assumptions, in the supplementary materials, we are working on further understanding connections between the two sets of assumptions. In addition, we are working on extensions of our general approach to accommodate nonignorable missingness and different BNP models.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the following NIH grants: R01GM112327, R01CA183845, R01CA83295, and R01HL64342.

References

- Albert JM, Nelson S. Generalized causal mediation analysis. *Biometrics*. 2011; 67:1028–1038. [PubMed: 21306353]
- Albert JM, Wang W. Sensitivity analyses for parametric causal mediation effect estimation. *Biostatistics*. 2014:339–351. [PubMed: 25395683]
- Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*. 1986; 51:1173. [PubMed: 3806354]

- Daniels MJ, Roy JA, Kim C, Hogan JW, Perri M. Bayesian inference for the causal effect of mediation. *Biometrics*. 2012; 68:1028–1036. [PubMed: 23005030]
- Elliott MR, Raghunathan TE, Li Y. Bayesian inference for causal mediation effects using principal stratification with dichotomous mediators and outcomes. *Biostatistics*. 2010; 11:353–372. [PubMed: 20101045]
- Escobar MD, West M. Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*. 1995; 90:577–588.
- Imai K, Keele L, Yamamoto T. Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*. 2010; 25:51–71.
- Jara A, Hanson TE, Quintana FA, Müller P, Rosner GL. Dppackage: Bayesian non-and semi-parametric modelling in r. *Journal of statistical software*. 2011; 40:1–30.
- Johnson, VE., Albert, JH. Ordinal data modeling. Springer; 1999.
- MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. *Psychological methods*. 2002; 7:83. [PubMed: 11928892]
- Marcus BH, Napolitano MA, King AC, Lewis BA, Whiteley JA, Albrecht A, Parisi A, Bock B, Pinto B, Sciamanna C, Jakicic J, Papandonatos GD. Telephone versus print delivery of an individualized motivationally tailored physical activity intervention: Project stride. *Health Psychology*. 2007a; 26:401. [PubMed: 17605559]
- Marcus BH, Napolitano MA, King AC, Lewis BA, Whiteley JA, Albrecht AE, Parisi AF, Bock BC, Pinto BM, Sciamanna CA, Jakicic JM, Papandonatos GD. Examination of print and telephone channels for physical activity promotion: Rationale, design, and baseline data from project stride. *Contemporary Clinical Trials*. 2007b; 28:90–104. [PubMed: 16839823]
- Mattei A, Li F, Mealli F. Exploiting multiple outcomes in bayesian principal stratification analysis with application to the evaluation of a job training program. *The Annals of Applied Statistics*. 2013; 7:2336–2360.
- Müller P, Erkanli A, West M. Bayesian curve fitting using multivariate normal mixtures. *Biometrika*. 1996; 83:67–79.
- Neal RM. Markov chain sampling methods for dirichlet process mixture models. *Journal of computational and graphical statistics*. 2000; 9:249–265.
- Nelsen, RB. An introduction to copulas. Springer-Verlag Inc.; 1999.
- Pearl, J. Proceedings of the seventeenth conference on uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc.; 2001. Direct and indirect effects; p. 411–420.
- Petersen ML, Sinisi SE, van der Laan MJ. Estimation of direct causal effects. *Epidemiology*. 2006; 17:276–284. [PubMed: 16617276]
- Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*. 2004; 36:717–731.
- Richardson TS, Robins JM. Single world intervention graphs (swigs): a unification of the counterfactual and graphical approaches to causality. Center for the Statistics and the Social Sciences, University of Washington Series. Working Paper. 2013
- Robins, JM. Highly Structured Stochastic Systems, chapter Semantics of causal DAG models and the identification of direct and indirect effects. Oxford: Oxford Univ. Press; 2003. p. 70–81.
- Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992; 3:143–155. [PubMed: 1576220]
- Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*. 1974; 66:688.
- Rubin DB. Comment on “Randomization analysis of experimental data: The Fisher randomization test” by D. Basu. *Journal of the American Statistical Association*. 1980; 75:591–593.
- Schwartz SL, Li F, Mealli F. A Bayesian semiparametric approach to intermediate variables in causal inference. *Journal of the American Statistical Association*. 2011; 106:1331–1344.
- Sevick MA, Napolitano MA, Papandonatos GD, Gordon AJ, Reiser LM, Marcus BH. Cost-effectiveness of alternative approaches for motivating activity in sedentary adults: results of project stride. *Preventive medicine*. 2007; 45:54–61. [PubMed: 17573103]

- Stan Development Team. Rstan: the r interface to stan, version 2.5.0. 2014
- Taddy, MA. PhD thesis. Santa Cruz: University of California; 2008. Bayesian Nonparametric analysis of conditional distributions and inference for poisson point processes.
- Tchetgen Tchetgen EJ, Shpitser I, et al. Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis. *The Annals of Statistics*. 2012; 40:1816–1845. [PubMed: 26770002]
- Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. Mediation: R package for causal mediation analysis. 2014
- Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with sas and spss macros. *Psychological methods*. 2013; 18:137. [PubMed: 23379553]
- VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*. 2009; 20:18–26. [PubMed: 19234398]
- VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology (Cambridge, Mass.)*. 2010; 21:540.
- VanderWeele TJ, Chiba Y. Sensitivity analysis for direct and indirect effects in the presence of exposure-induced mediator-outcome confounders. *Epidemiology, Biostatistics and Public Health*. 2014
- Yuan Y, MacKinnon DP. Bayesian mediation analysis. *Psychological methods*. 2009; 14:301. [PubMed: 19968395]

Figure 1.

The horizontal line from Z to Y captures the direct effect and the lines from Z to M and M to Y captures the indirect effect. And dashed line emitting from X implies conditioning on covariate values.

Table 1

Results for point estimators of NIE, NDE and TE over 500 replications. The columns correspond to bias and MSE to the true values of NIE, NDE and TE under three methods: (a) a frequentist semiparametric model; (b) a Bayesian parametric model; (c) a Bayesian nonparametric model

	Truth	(a) Frequentist			(b) Parametric			(c) BNP		
		Bias	MSE		Bias	MSE		Bias	MSE	
Case 1	NIE	0.616	0.012	(0.16)	-0.020	(0.34)		-0.004	(0.42)	
	NDE	1.3	0.060	(2.22)	-0.097	(1.99)		0.045	(2.03)	
	TE	1.916	0.073	(2.19)	-0.118	(2.12)		0.041	(2.16)	
	Time (Sec.)	-	108.8 sec.		157.4 sec.			287.7 sec.		
Case 2	NIE	1.836	-0.387	(0.90)	-0.142	(2.60)		-0.099	(1.21)	
	NDE	1.976	0.195	(3.39)	0.228	(3.46)		0.065	(2.04)	
	TE	3.812	-0.191	(3.14)	0.085	(2.46)		-0.035	(2.59)	
	Time (Sec.)	-	112.6 sec.		159.1 sec.			291.1 sec.		
Case 3	NIE	0.752	0.024	(0.09)	0.179	(0.19)		0.049	(0.12)	
	NDE	2.5	-0.002	(0.00)	-0.167	(0.03)		-0.045	(0.01)	
	TE	3.252	0.022	(0.09)	0.012	(0.14)		0.004	(0.11)	
	Time (Sec.)	-	105.1 sec.		323.3 sec.			268.8 sec.		

Table 2

Posterior means (95% credible intervals) of NIE, NDE and TE for each combination of covariates values (bold font) under 2 models: 1) the BNP model with sequential ignorability; 2) the BNP model with mediator induction equivalence. Other covariates are integrated out.

Covariate	Assumption	NIE	NDE	TE
AGE	(1)	0.49 (−0.02, 1.05)	0.77 (−0.30, 1.85)	1.26 (0.36, 2.19)
BMI	(2)	0.57 (0.10, 1.25)	0.69 (−0.45, 1.81)	
AGE	(1)	0.57 (0.08, 1.19)	0.75 (−0.31, 1.71)	1.32 (0.60, 2.12)
BMI	(2)	0.68 (0.12, 1.26)	0.82 (−0.32, 1.73)	
AGE	(1)	0.59 (0.06, 1.39)	0.44 (−0.95, 1.61)	1.04 (0.12, 1.98)
BMI	(2)	0.66 (0.07, 1.41)	0.37 (−1.08, 1.62)	
AGE	(1)	0.38 (−0.04, 0.80)	1.39 (0.68, 2.14)	1.78 (1.18, 2.37)
BMI	(2)	0.44 (0.07, 0.87)	1.33 (0.61, 2.07)	
AGE	(1)	0.51 (0.08, 1.01)	1.05 (0.29, 1.70)	1.56 (1.05, 2.07)
BMI	(2)	0.59 (0.19, 1.06)	0.96 (0.22, 1.64)	
AGE	(1)	0.61 (0.06, 1.28)	0.38 (−0.84, 1.45)	0.99 (0.07, 1.88)
BMI	(2)	0.73 (0.11, 1.43)	0.27 (−1.11, 1.39)	
AGE	(1)	0.29 (−0.17, 0.78)	1.42 (0.59, 2.16)	1.72 (0.95, 2.37)
BMI	(2)	0.36 (−0.08, 0.85)	1.36 (0.49, 2.18)	
AGE	(1)	0.44 (−0.08, 1.04)	0.79 (−0.21, 1.68)	1.23 (0.46, 2.07)
BMI	(2)	0.56 (0.01, 1.15)	0.67 (−0.28, 1.63)	
AGE	(1)	0.59 (−0.14, 1.40)	−0.11 (−1.59, 1.22)	0.48 (−0.79, 1.61)
BMI	(2)	0.80 (−0.04, 1.67)	−0.32 (−1.76, 1.11)	

Table 3

Posterior means and credible intervals for marginal NDE, NIE, and TE under different combinations of sensitivity parameters k_0 and k_1 or uniform priors on them: $k_0 \sim \text{Unif}(0, 10)$, $k_1 \sim \text{Unif}(k_0, 20)$.

Sign	k_0	k_1	NIE	NDE	TE
+	5	5	0.42 (−0.03, 0.97)	0.78 (0.19, 1.26)	1.20 (0.95, 1.45)
		10	0.61 (0.16, 1.16)	0.59 (0.01, 1.07)	1.20 (0.95, 1.45)
		15	0.76 (0.31, 1.31)	0.45 (−0.13, 0.92)	1.20 (0.95, 1.45)
	10	10	0.42 (−0.03, 0.97)	0.78 (0.19, 1.26)	1.20 (0.95, 1.45)
		15	0.56 (0.12, 1.12)	0.64 (0.05, 1.11)	1.20 (0.95, 1.45)
		20	0.69 (0.24, 1.24)	0.64 (−0.07, 0.99)	1.20 (0.95, 1.45)
	Unif(k_0 , 20)		0.69 (0.13, 1.29)	0.51 (−0.24, 1.12)	1.20 (0.95, 1.45)
	5	5	0.42 (−0.03, 0.97)	0.78 (0.19, 1.26)	1.20 (0.95, 1.45)
		10	0.23 (−0.22, 0.78)	0.97 (0.38, 1.45)	1.20 (0.95, 1.45)
		15	0.08 (−0.37, 0.64)	1.12 (0.52, 1.60)	1.20 (0.95, 1.45)
−	10	10	0.42 (−0.03, 0.97)	0.78 (0.19, 1.26)	1.20 (0.95, 1.45)
		15	0.27 (−0.18, 0.82)	0.93 (0.34, 1.41)	1.20 (0.95, 1.45)
		20	0.15 (−0.30, 0.70)	1.05 (0.46, 1.53)	1.20 (0.95, 1.45)
	Unif(k_0 , 20)		0.14 (−0.43, 0.77)	1.06 (0.43, 1.66)	1.20 (0.95, 1.45)