Routledge
Taylor & Francis Group

Multivariate Behavioral Research

ISSN: 0027-3171 (Print) 1532-7906 (Online) Journal homepage: https://www.tandfonline.com/loi/hmbr20

A More Efficient Causal Mediator Model Without the No-Unmeasured-Confounder Assumption

Holger Brandt

To cite this article: Holger Brandt (2019): A More Efficient Causal Mediator Model Without the No-Unmeasured-Confounder Assumption, Multivariate Behavioral Research, DOI: 10.1080/00273171.2019.1656051

To link to this article: https://doi.org/10.1080/00273171.2019.1656051

Published online: 09 Sep 2019.

MULTIVARIATE BEHAVIORAL RESEARCH
https://doi.org/10.1080/00273171.2019.1656051

A More Efficient Causal Mediator Model Without the No-Unmeasured-Confounder Assumption

Holger Brandt

Department of Psychology, University of Zürich

ABSTRACT

Mediator models have been developed primarily under the assumption of no-unmeasured-confounding. In many situations, this assumption is violated and may lead to the identification of mediator variables that actually are statistical artifacts. The rank preserving model (RPM) is an alternative approach to estimate controlled direct and mediator effects. It is based on the structural mean models framework and a no-effect-modifier assumption. The RPM assumes that unobserved confounders do not interact with treatment or mediators. This assumption is often more plausible to hold than the no-unmeasured-confounder assumption. So far, models using the no-effect-modifier assumption have been rarely used, which might be due to its low power and inefficiency in many scenarios. Here, a semi-parametric nonlinear extension, the nRPM, is proposed that overcomes this inefficiency using thin plate regression splines that both increase the predictive power of the model and decrease the misspecification present in many situations. In a simulation study, it is shown that the nRPM provides estimates that are robust against the violation of the no-effect-modifier assumption and that are substantively more efficient than those of the RPM. The model is illustrated using a data set on CD4 cell counts in a context of the human immunodeficiency virus (HIV).

KEYWORDS

Mediator model; structural mean models; thin plate regression spline; generalized additive model; no-unmeasured-confounder; rank preserving model; no-effect-modifier

The identification of intermediate variables that give an indication about future treatment success has received increasing attention in social and education sciences as well as in clinical intervention studies (MacKinnon, Fairchild, & Fritz, 2007; Nock, 2007; Shrout & Bolger, 2002). In clinical trials, it is often of interest to identify variables that can indicate early on if a treatment is effective. For example, for the evaluation of HIV treatments, long-term outcomes such as survival rates or information on CD4 cell counts are typically assessed only after several years. Intermediate variables that indicate treatment success already at an earlier stage provide incremental knowledge on the treatment process, and may answer the question how a certain treatment works. Also, it makes intervention studies more cost efficient (Burzykowski, Molenberghs, & Buyse, 2006).

Different methods have been developed that try to identify such intermediate variables in empirical contexts (Baron & Kenny, 1986; Imai, Keele, & Tingley, 2010; Imai, Keele, & Yamamoto, 2010; Prentice, 1989; Ten Have et al., 2007; Van der Weele, 2010; Zheng & Zhou, 2015). Most of them are based on the concept of a mediator variable, which is assumed to be an intermediate variable that is part of a causal chain leading from the intervention (or treatment) to an outcome.

Different types of effects have been defined to evaluate the mediation process (Robins & Greenland, 1992). The controlled direct effect (CDE) is defined as the causal effect of the treatment on the outcome controlled for the mediator variable. It provides information about how strongly the treatment directly affects the outcome through mechanisms (or paths) other than the mediator investigated. Similarly, the controlled mediator effect (CME) is defined as the causal effect of the mediator on the outcome controlled for the treatment. The CME can be used to evaluate the relevance of mediator candidates as actual intermediate variables (Small, 2012; Ten Have et al., 2007). Controlled effects rely on the assumption that the mediator can at least hypothetically be directly manipulated. As an alternative, natural direct (NDE) and indirect effects (NIE) can be used to describe the

CONTACT Holger Brandt holger.brandt@uzh.ch Department of Psychology, University of Zürich, Binzmühlestrasse 14/28, Zürich, CH-8050 Switzerland.

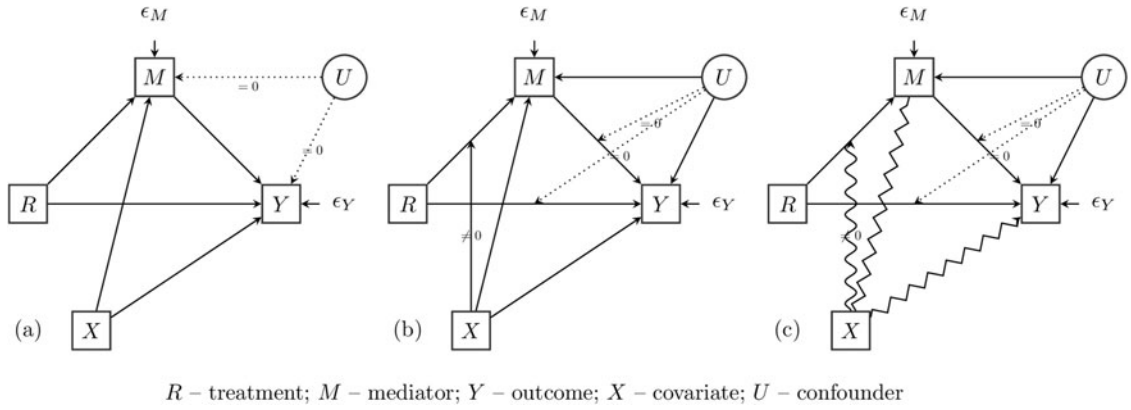

Figure 1. Path diagram for the mediator model using the no-confounder assumption (a), the RPM (b), and the nRPM (c). Dotted lines indicate paths that are assumed to be zero, and snake lines indicate semi-parametric spline functions. Arrows ending on other arrows indicate interaction effects.

mediation characteristics of intermediate variables. Natural effects use the levels of the mediator that naturally occur under the different treatment conditions (e.g., Van der Weele, 2009). An advantage of natural effects is that they sum up to the total effect; thus, they can be used to investigate the pathways of an intervention. Both natural and controlled effects can be related to each other under certain assumptions; the assumptions for natural effects are stronger than those for controlled effects as will be detailed below (cf. Van der Weele, 2009). However, this article focuses on controlled effects and their estimation. They can be identified primarily under two different types of assumptions: the no-unmeasured-confounder assumption and the no-effect-modifier assumption.

No-unmeasured-confounder assumption

Most applications that estimate controlled effects are based on two assumptions: First, it is assumed that there is (i) no unmeasured confounding of the relationship between outcome and treatment, and that there is (ii) no unmeasured confounding of the relationship between mediator and outcome (Van der Weele, 2009). The first assumption holds, for example, if the treatment is completely randomized. The second assumption is more critical because it can be violated even if treatment was randomized (post-treatment selection bias; Holland, 1988). A typical mediator model and its underlying assumptions are illustrated in Figure 1a: A treatment R affects both mediator (M) and outcome (Y). The no-unmeasured confounder assumption (ii) states that all relevant covariates are observed and subsumed under the observed covariates X; no further unobserved confounder (U) exists that influences the relationship between the mediator M and the outcome variable Y.[1]

One of the most prominent statistical models to identify mediator variables under these no-unmeasured-confounder assumptions using a regression analysis framework was proposed by Baron and Kenny (1986). Possible confounders can be incorporated in the model as covariates if they have been measured. Alternatively, the marginal structural modeling framework can be used to estimate CDE and CME. In this model class, covariates are used to first estimate weights that are then used for a weighted regression analysis, that is, covariates are not directly included in the regression of M and R on Y (details can be found in Van der Weele, 2009 and Robins, Hernán, and Brumback, 2000).

The no-unmeasured-confounder assumption particularly for the relationship between M and Y has repeatedly been criticized strongly because it cannot be tested and it is often implausible (Bullock, Green, & Ha, 2010; Holland, 1988; Joffe & Greene, 2009; Small, Joffe, Lynch, Roy, & Locali, 2014; Ten Have et al., 2007; Van der Weele, 2016). Violations of the assumption lead to severe bias for the controlled (as well as for the natural) effects and may result in the detection of spurious mediator variables (Zheng, Atkins, Zhou, & Rhew, 2015). Different consequences were drawn from this critique. Sensitivity analyses

[1]Note that the NDE is equivalent to the CDE as long as there is no interaction between treatment and mediator on the outcome ($R \times M \rightarrow Y$). The NIE can either be obtained as a the difference between total effect of R on Y and NDE or an additional model of the relationship between M and R is estimated (see details in Van der Weele, 2009 and Zheng and Zhou, 2015). Formally, the natural effects are identified with the additional assumptions of (iii) no unmeasured confounding of the relationship between M and R and (iv) no unmeasured confounding for the relationship between M and Y that are due to effects from the treatment. Together, assumptions (i) through (iv) form the so called "sequential ignorability" assumption (Van der Weele, 2009).

were proposed that can be used to investigate how controlled or natural effect parameter estimates change under different amounts of violation of the no-unmeasured-confounder assumption (Brumback, Hernán, Haneuse, & Robins, 2004; Ding & Van der Weele, 2016; Imai et al., 2010). Different weighting techniques (Hong, Qin, & Yang, 2018) or other adjustments (Kaufman, Kaufman, MacLehose, Greenland, & Poole, 2005) were developed to reduce the bias for natural or controlled effects. Alternative definitions were suggested to weaken the problematic assumption by defining natural effects for certain subgroups (e.g., for those exposed to the treatment; Vansteelandt & Van der Weele, 2012), and procedures on how to select relevant covariates (Van der Weele & Shpitser, 2011, 2013) or how to be more specific about the data generation process were developed (Pearl, 2012, 2014). These methods enhance the applicability of models using the no-unmeasured-confounder assumption. However, they still rely heavily on a complete and correct specification of all relevant covariates that affect outcome and mediator in the model to avoid bias. This specification may be difficult in most applied settings.

No-effect-modifier assumption

One alternative approach for the estimation of CME and CDE was introduced as the rank preserving model (RPM; Small, 2012; Ten Have et al., 2007; Zheng & Zhou, 2015). The RPM does not rely on the no-unmeasured-confounder assumption but is based on a "no-effect-modifier" assumption. Figure 1b illustrates this assumption (Zheng & Zhou, 2015): To provide unbiased estimates, it is assumed that no unobserved confounders exist that modify the effect between treatment, mediator, and outcome, that is, there are no interaction effects $R \times U \rightarrow Y$ or $M \times U \rightarrow Y$. The model allows unobserved confounders to be associated with both mediator and outcome. The model can be estimated via the G-estimation approach within the structural mean modeling framework (Robins, 1994).

Advantages of the RPM

The no-effect-modifier assumption can be viewed as a weaker assumption than the no-unmeasured-confounder assumption from several view points: First, confounders that interact with treatment or mediator typically also have a linear effect, violating both no-effect-modifier, and no-unmeasured-confounder assumption simultaneously (Zheng & Zhou, 2015). A violation of the no-unmeasured-confounder assumption, though, does not as strongly imply a violation of the no-effect-modifier assumption. Second, effects sizes of linear effects are often larger than those of interaction effects. Typical interaction effect sizes lie in an area of about 2% explained variance (Chaplin, 1991, 2007), and the actual severity of violating the no-effect-modifier assumption might be smaller in many situations (but see also Petersen, Sinisi, and van der Laan 2006 and Vansteelandt and Van der Weele 2012). Third, simulation studies showed that the RPM is rather robust against the violation of its no-effect-modifier assumption (Ten Have et al., 2007; Zheng et al., 2015; Zheng & Zhou, 2015). This is particularly important because this assumption is not testable.

Another advantage of the RPM compared to models formulated in the regression framework is that the consistency of the parameter estimates of treatment and mediator on the outcome does not rely on the correct specification of the relationship between covariates and outcome which makes model (mis)-specification more flexible (Ten Have et al., 2007).

Limitations of the RPM

However, the RPM's advantages come at the cost of a low efficiency and a low power to detect relevant mediator variables. Simulation results (Ten Have et al., 2007; Zheng et al., 2015; Zheng & Zhou, 2015) indicated that the actual power of the RPM depends on certain characteristics of the covariates, mainly on an interaction effect between treatment and covariate to predict the mediator variable ($X \times R \rightarrow M$; see Figure 1b). Some of the simulation studies might have been too optimistic because they used interaction effect sizes with up to 12% explained variance, which are untypical for reported effects in empirical settings (Chaplin, 1991, 2007).

In addition, even though consistency is not affected, the efficiency of the RPM depends on the correct specification of the relationship between covariates and outcome variable (Fischer-Lapp & Goetghebeur, 1999). So far, it has not been investigated how severely the efficiency of the RPM is affected if this relationship is misspecified. Nonetheless, it is a very important aspect of the RPM due to the overall low efficiency even if the relationship is correctly modeled. Furthermore, misspecification in this model part cannot be tested without additional strong assumptions (Ten Have et al., 2007).

Finally, applied researchers are often interested in more complex models that include, for example, interactions between treatment, covariates, and mediator variables in order to model differential relationships (e.g., for whom the mediation effect is most pronounced; Preacher, 2015; Preacher, Rucker, & Hayes, 2007). The RPM is capable to estimate these effects theoretically (Zheng & Zhou, 2015), its practical estimation, however, has not been investigated at all. Again, an application of the RPM might be problematic because its estimation routine can easily be impaired by multicollinearity problems (as discussed in the next section), resulting in non-identification or large standard errors.

Scope and outline

Taken together, the RPM is a promising model that potentially is more robust to the existence of unobserved confounders. Its practical use might be limited due to a general low power. Here, this limitation is overcome by extending the RPM to a nonlinear RPM (nRPM) using splines within the generalized additive models framework (GAM; Hastie, Tibshirani, & Friedman, 2009; Wood, 2017). GAMs are often used to increase the predictive power of a model. Here, GAMs are applied in a very specific way (see Figure 1c): while the structural part of the RPM referring to the mediation process remains unchanged, potential misspecification of the covariate function and the amount of residual variance is reduced using data-driven semi-parametric functions, which results in an increased efficiency for the model parameters of interest (e.g., CME or CDE). Furthermore, the model can be used to reduce multicollinearity in the estimation routine by drawing on any unspecific types of nonlinearity in the data that would be ignored by models with a parametric structural model. This might enhance the performance to detect, for example, interaction effects.

In the next section, the model specification for the RPM is reviewed and its underlying assumptions are discussed; the extension to the nRPM is provided. A simulation study is used to demonstrate that the proposed extension increases the model efficiency and power substantively for both CME and interaction effects. The robustness of the RPM and the nRPM against a violation of the no-effect-modifier assumption is investigated and compared to the robustness of mediation models based on the no-unmeasured-confounder assumption. The nRPM is illustrated with an empirical data set on CD4 cell counts and its applicability will be discussed.

The rank preserving model

In this section, the main equations and underlying assumptions of the RPM are presented as they have been outlined in Ten Have et al. (2007) and Zheng and Zhou (2015). Then the extensions to the nRPM are presented. The main difference between the RPM and the nRPM is that the RPM uses a parametric specification for the structural relationship between the variables whereas the nRPM substitutes them with semi-parametric splines to incorporate the covariates as a standard specification.

Notation

For the remainder of the paper, it is assumed that R is a randomized treatment where $R_i = 0$ and $R_i = 1$ indicate that a subject i was randomized to either control or treatment group, and received some kind of intervention. Y is an outcome variable measured at a time point after the intervention. M is an intermediate variable that was measured in between the intervention and the outcome. Finally, $\mathbf{X}$ is a set of K baseline covariates that were measured at or before the intervention. Y_i^{rm} indicates a potential value for subject i for $R = r$ and $M = m$, where r, m are levels of the respective variables. An observed score in Y is indicated as $\tilde{Y}_i$. Similarly, $R = \tilde{r}, M = \tilde{m}$, and $\mathbf{X} = \tilde{\mathbf{x}}$ indicate observed values in the respective variables.

Model

The RPM and the nRPM are specified for the potential values Y_i^{rm}, and $j = 1 \ldots J$ mediators $\mathbf{M} = (M_1, \ldots M_J)'$ for each person $i = 1 \ldots N$ as follows (Zheng & Zhou, 2015):

$$Y_i^{r\mathbf{m}} = G(\tilde{\mathbf{x}}_i) + \boldsymbol{\theta} s(\mathbf{m}, r, \tilde{\mathbf{x}}_i) + \epsilon_{Yi}^{r\mathbf{m}}(\mathbf{u}_i, \tilde{\mathbf{x}}_i), \qquad (1)$$

where $G(\tilde{\mathbf{x}}_i)$ is an unknown function of the covariate vector and $s = (s_1, \ldots, s_P)'$ is a vector that specifies the structural model part referring to the mediator variables, the treatment, and possible interactions with the covariates, where each s_p satisfies $s(\mathbf{0}, 0, \tilde{\mathbf{x}}_i) = 0$. $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_P)$ is a parameter vector that includes the parameters of interest. $\epsilon_{Yi}^{r\mathbf{m}}(\mathbf{u}_i, \tilde{\mathbf{x}}_i)$ is a residual with unknown distributional form, and $\mathbf{u}$ is a set of unobserved confounders. It is not assumed that ϵ is uncorrelated with $\mathbf{M}$ (as in regression models).

The model includes two important parts: first, the mediating mechanism (given the covariates) is modeled via s. The actual specification of s is a parametric model that allows researchers to specify hypotheses that directly relate to, for example, CDE or CME as

parameters in the model. Second, G is a model used to account for the covariates. Overall model performance depends on both the selection of the covariates (e.g., with regard to the plausibility of the no-effect-modifier assumption) and the specific function used. For the RPM, parametric structural models were chosen so far (i.e., a standard regression type model); in the nRPM, G will be estimated via linear basis expansion (see subsection below). This part is considered the "noise" part of the model, and no specific interpretation for G is provided.

The simplest case for a single mediator M_1 and no interactions with $s_1 = r$ and $s_2 = m_1$ results in Ten Have's model (see Zheng & Zhou, 2015):

$$Y_i^{rm_1} = G(\tilde{\mathbf{x}}_i) + \theta_1 r + \theta_2 m_1 + \epsilon_{Yi}^{rm_1}(\mathbf{u}_i, \tilde{\mathbf{x}}_i), \quad (2)$$

from which the CDE and CME can directly be estimated as the parameters of interest:

$$CDE := E\left[Y_i^{1m_1} - Y_i^{0m_1} | \tilde{\mathbf{x}}_i\right] = \theta_1 \quad (3)$$

$$CME := E\left[Y_i^{rm_1^1} - Y_i^{rm_1^0} | \tilde{\mathbf{x}}_i\right] = \theta_2, \quad (4)$$

where m_1^1, m_1^0 are potential values in the mediator M_1 for $R = 1$ and $R = 0$.

More complex models can be specified, for example, for differential (interaction) effects: A model with a single mediator variable and its interaction with the treatment or the covariates is included by $s(\mathbf{m}, r, \tilde{\mathbf{x}}_i) = (m_1, r, m_1 \times r)'$ or $s(\mathbf{m}, r, \tilde{\mathbf{x}}_i) = (m_1, r, m_1 \times \tilde{x}_{1i}, \ldots, m_1 \times \tilde{x}_{Ki})'$, respectively. Multiple mediators can be tested simultaneously, for example, by $s(\mathbf{m}, r, \tilde{\mathbf{x}}_i) = (m_1, m_2, r, m_1 \times m_2)'$, where the last term specifies an interaction between the two mediator variables.

As soon as interactions between M and covariates $\mathbf{X}$ or the treatment R are specified, the CDE and CME are more complex and depend on the specific values of $\mathbf{X}$, M and R (more details and explanations about the relationship between natural and controlled effects can be found in, for example, Valeri & Van der Weele, 2013; Van der Weele & Vansteelandt, 2009; Zheng & Zhou, 2015). For example, for a model including interaction effects between mediator, treatment, and a covariate X_1 of the form

$$Y_i^{rm_1} = G(\tilde{\mathbf{x}}_i) + \theta_1 r + \theta_2 m_1 + \theta_3 m_1 r + \theta_4 m_1 \tilde{x}_{1i} + \theta_5 r \tilde{x}_{1i} + \epsilon_{Yi}^{rm_1}(\mathbf{u}_i, \tilde{\mathbf{x}}_i), \quad (5)$$

the CDE and the CME are given by

$$CDE = \theta_1 + \theta_3 m_1 + \theta_5 \tilde{x}_{1i} \quad (6)$$

$$CME = \theta_2 + \theta_3 r + \theta_4 \tilde{x}_{1i} \quad (7)$$

and depend on the actual levels of $R = r$, $M_1 = m_1$ and the observed levels in the covariate.

Model assumptions

The model is identified by two sets of assumptions (see details in Ten Have et al., 2007 and Zheng and Zhou, 2015). The first set includes standard assumptions in causal modeling, the second set includes specific assumptions for this mediator model. The assumptions are only partly testable.

Standard assumption

First, it is assumed that the stable unit treatment value assumption (SUTVA) holds (1). This assumption implies that the assignment of one subject does not affect other subjects' potential values ("no interference") and that there are no treatment changes that are not modeled ("no treatment variation"). Together, this implies a single set of potential values Y_i^{rm} for each subject i. Second, the observed outcome is a realization of the potential values (2) ("consistency"). And third, the treatment itself is randomized (3) (or, for observational studies, strong ignorability; Rosenbaum & Rubin, 1983).

These three assumptions are typically made in the context of treatment studies and are part of most approaches to mediator models (Gilbert & Hudgens, 2008; Zheng & Zhou, 2015). Assumptions (1) and (3) can be evaluated with regard to their plausibility in the respective context. Randomization can be ensured by the treatment design and can be tested at least partly using available covariates. Assumption (2) cannot be tested and should be viewed as technical assumption.

Specific assumptions

Estimation of the RPM is based on the following specific assumptions: (4) a constant conditional mean for the error term (no-effect-modifier assumption), (5) a correct model specification in s, and (6) a non-degenerate weight matrix for the estimation routine.

Assumption (4) states that the error term has a constant mean for all potential values, which implies that there are no unobserved confounders that interact with mediator or treatment. Formally, this is expressed as

$$E\left[\epsilon_{Yi}^{\mathbf{rm}}(\mathbf{u}_i, \tilde{\mathbf{x}}_i) | \mathbf{m}, r, \tilde{\mathbf{x}}_i\right] = F(\mathbf{m}, r, \tilde{\mathbf{x}}_i) \forall r, \mathbf{m} \quad (8)$$

(Small, 2012; Zheng & Zhou, 2015). This assumption is considerably weaker than the original rank preservation assumption in Ten Have et al. (2007), which also included non-interaction assumptions among the treatment, mediator and observed covariates. The assumption is not testable.

Assumption (5) states that the structural model is correctly specified in s. This implies that all relevant interactions between covariates, treatment, and mediator(s) are modeled. If the mediator variable is continuous this assumption extends to the correct specification of its relationship to the outcome variable. A misspecification of the covariate function G does not affect consistency of the estimation but it affects the efficiency of model estimation (Fischer-Lapp & Goetghebeur, 1999). The assumption is only testable under additional assumptions (such as no-unmeasured confounding) because it refers to potential values.

To identify the different parameters in the model, a non-degenerate weight matrix $\mathbf{W}(\tilde{\mathbf{x}}_i)$ needs to be specified (6). Optimal weights were derived for the model in Eq. (2) in Ten Have et al. (2007). A more general derivation of weights and specifically weights for interaction effects were presented in Zheng and Zhou (2015). The assumption is testable, for example, by calculating the eigenvalues of the covariance matrix of the weights. Details on specific weights will be presented below.

A general proof of model identification for Eq. (1) using assumptions (1) to (6) as well as the derivation of optimal weights and their properties are provided in theorems 1 to 3 in Zheng & Zhou (2015).

Estimating equations

Based on the model formulation and the assumptions, model estimation can be conducted using G-estimation equations (Robins, 1994; Ten Have et al., 2007; Zheng & Zhou, 2015). For a binary treatment, the estimating equations are given by

$$\sum_i (R_i - E[R_i])\mathbf{W}(\tilde{\mathbf{x}}_i)\Big(\tilde{Y}_i - \boldsymbol{\theta} s(\tilde{\mathbf{m}}_i, \tilde{r}_i, \tilde{\mathbf{x}}_i) - \tilde{G}(\tilde{\mathbf{x}}_i)\Big) = \mathbf{0} \tag{9}$$

with weight matrix $\mathbf{W}(\tilde{\mathbf{x}}_i) = (W_1(\tilde{\mathbf{x}}_i), ..., W_P(\tilde{\mathbf{x}}_i))$ and optimal weights

$$\begin{aligned} W_p(\tilde{\mathbf{x}}_i) &= E\big[s_p(\mathbf{m}, r, \tilde{\mathbf{x}}_i)|R = 1, \tilde{\mathbf{x}}_i\big] \\ &- E[s_p(\mathbf{m}, r, \tilde{\mathbf{x}}_i)|R = 0, \tilde{\mathbf{x}}_i] \text{ for } p = 1...P \end{aligned} \tag{10}$$

as derived in Zheng and Zhou (2015). For example, for the simple model specified in Eq. (2), weights are given by

$$\begin{aligned} W_1 &= E[R|R = 1, \tilde{\mathbf{x}}_i] - E[R|R = 0, \tilde{\mathbf{x}}] = 1 \\ W_2 &= E[M_1|R = 1, \tilde{\mathbf{x}}_i] - E[M_1|R = 0, \tilde{\mathbf{x}}_i] =: \eta_{1i1}(\tilde{\mathbf{x}}_i) \end{aligned} \tag{11}$$

(which is in line with Ten Have et al., 2007). The main technical assumption relevant for the derivation of the weights is assumption (6): The covariance matrix of $\mathbf{W}(\tilde{\mathbf{x}}_i)$ is positive definite (Zheng & Zhou, 2015). Here, this is the case as long as there is an interaction effect between treatment and at least one covariate on M_1 ($R \times X_k \rightarrow M_1$) because only in this situation $\eta_{1i1}(\tilde{\mathbf{x}}_i)$ is not constant.

A set of the most relevant optimal weights for $j, j' = 1...J$ mediator variables and $k = 1...K$ covariates are specified as follows (see Zheng & Zhou, 2015):

$$\begin{pmatrix} 1 \\ \eta_{1ij}(\tilde{\mathbf{x}}_i) \\ \eta_{2ij}(\tilde{\mathbf{x}}_i) \\ \tilde{x}_{ik} \\ \eta_{1ij}(\tilde{\mathbf{x}}_i) \cdot \tilde{x}_{ik} \\ \eta_{3ijj'}(\tilde{\mathbf{x}}_i) \end{pmatrix} \text{ corresponding to elements in s: } \begin{pmatrix} R \\ M_j \\ R \times M_j \\ R \times X_k \\ M_j \times X_k \\ M_j \times M_{j'} \end{pmatrix}, \tag{12}$$

with $\eta_{2ij}(\tilde{\mathbf{x}}_i)$ and $\eta_{3ijj'}(\tilde{\mathbf{x}}_i)$ defined as

$$\eta_{2ij}(\tilde{\mathbf{x}}_i) := E\big[M_j|R = 1, \tilde{\mathbf{x}}_i\big] \tag{13}$$

$$\eta_{3ijj'}(\tilde{\mathbf{x}}_i) := E\big[M_j M_{j'}|R = 1, \tilde{\mathbf{x}}_i\big] - E(M_j M_{j'}|R = 0, \tilde{\mathbf{x}}_i). \tag{14}$$

Again, it needs to be ensured that the weights are not collinear in order to identify the model parameters. For example, the identification of the interaction between treatment and k-th covariate ($R \times X_k \rightarrow Y$) implies that the respective covariate has some form of nonlinear interaction effect with the treatment on M_j (e.g., $R \times X_k^2 \rightarrow M_j$) to ensure that the weight is not collinear with η_{1ij} because η_{1ij} is a linear combination of the covariates (Zheng & Zhou, 2015). If the covariate is binary, this interaction is not identified in general.

Model estimation in the RPM and the nRPM

Estimation implies two parts (Ten Have et al., 2007): First, an estimate of the weight matrix needs to be obtained. Second, $\boldsymbol{\theta}$ and G are then estimated in an iterative procedure solving Eq. (9). For G a working model estimate $\tilde{G}$ is used. Both the RPM and the nRPM are semi-parametric approaches because they do not use parametric specifications of the error distribution but use a semi-parametric G-estimation procedure. The distinction made here and elsewhere for the RPM and the nRPM refers only to the structural specification (e.g., $X \rightarrow Y$).

In the original RPM formulation both weights and $\tilde{G}$ were based on fully parametric specifications (Ten Have et al., 2007; Zheng et al., 2015; Zheng & Zhou, 2015). These specific parametric models may be misspecified and they implied the need to derive specific

nonlinear terms (such as $R \times X_k^2 \rightarrow M_j$, see above). Furthermore, Zheng and Zhou (2015) proofed that consistency of the estimator holds even if the working model $\tilde{G}$ for G is misspecified or if $\mathbf{W}$ is not chosen optimal (as long as some general conditions hold). Its variance and the resulting standard errors, though, are affected by the amount of misspecification in $\tilde{G}$ (Fischer-Lapp & Goetghebeur, 1999) and if suboptimal weights are used (Zheng & Zhou, 2015). The correctness of $\tilde{G}$ cannot be tested because the model in Eq. (1) refers to potential values.

Standard errors for the relevant coefficients in $\boldsymbol{\theta}$ can be obtained after convergence using sandwich estimator formulas provided by Ten Have et al. (2007). They are based on the score vector $S_i = (r_i - q)(Y_i^{00}(\hat{\boldsymbol{\theta}}) - \tilde{G}(\mathbf{x}_i))\mathbf{W}(\mathbf{x}_i)$ and its gradient. These standard errors are likely to underestimate the sampling variability because they do not take into account that the weight matrix $\mathbf{W}$ is based on estimates (cf. Ten Have et al., 2007). Here, bootstrapped standard errors (Efron, 1979) are obtained via standard nonparametric resampling instead.

In the next subsections, the extension to the nRPM as a (nonlinear) model using thin plate regression splines will be presented. This extension increases the efficiency of the estimator and reduces potential misspecification.

A semi-parametric model for the weight scores

Most weights presented in Eq. (12) depend on a model to predict the mediator variable(s) (i.e., $E[M_j|R = r, \tilde{\mathbf{x}}_i]$). Here, a generalized additive modeling approach with thin plate regression splines is proposed:

$$\delta(\mu_{ji}) = F_j(\tilde{\mathbf{x}}_i, \tilde{r}_i) \tag{15}$$

(Wood, 2003, 2017), where δ is a link function depending on the variable type of M_j (e.g., an identity function for continuous M, or a logit function for binary M), and $\mu_{ji} := E[M_{ji}|\tilde{r}_i, \tilde{\mathbf{x}}_i]$. In addition, a distributional assumption for M_j is made such as a Gaussian distribution ($M_{ji} \sim N(\mu_{ji}, \sigma_j)$) for continuous variables, or a Bernoulli distribution ($M_{ji} \sim Bern(\mu_{ji})$) for binary variables.

Thin plate splines are a multivariate extension of smoothing splines for multiple dimensions, that is, for multiple predictors (Duchon, 1977; Wahba, 1990; Wood, 2003). Here, the function F_j is defined as:

$$\begin{aligned} F_j(\tilde{\mathbf{x}}_i, \tilde{r}_i) &= \alpha_{M_j} + f_{j1}(\tilde{x}_{i1}) + \ldots + f_{jK}(\tilde{x}_{iK}) \\ &+ f_{j12}(\tilde{x}_{i1}, \tilde{x}_{i2}) + \ldots + f_{j(K-1)K}(\tilde{x}_{i(K-1)}, \tilde{x}_{iK}) \\ &+ \Big(\beta_{Rj} + f'_{j1}(\tilde{x}_{i1}) + \ldots + f'_{jK}(\tilde{x}_{iK}) + f'_{j12}(\tilde{x}_{i1}, \tilde{x}_{i2}) \\ &+ \ldots + f'_{j(K-1)K}(\tilde{x}_{i(K-1)}, \tilde{x}_{iK})\Big) \cdot \tilde{r}_i, \end{aligned} \tag{16}$$

where α_{M_j} is an intercept and the f's are smooth functions. F_j is approximated by the function $\hat{H}_j$ that minimizes

$$||\delta(\boldsymbol{\mu_j}) - \mathbf{H}_j||^2 + \lambda \Im(\mathbf{H}_j) \tag{17}$$

with the vectors $\boldsymbol{\mu_j} = (\mu_{j1} \ldots \mu_{jN}), \mathbf{H}_j = (H_j(\tilde{\mathbf{x}}_1, \tilde{r}_1) \ldots H_j(\tilde{\mathbf{x}}_N, \tilde{r}_N))$, and the Euclidean norm $||\cdot||$. λ is a penalty parameter that controls the smoothness of the function $\mathbf{H}_j$ and $\Im$ is a penalty function that indicates the wiggliness of $\mathbf{H}_j$. $\mathbf{H}_j$ includes the basis functions $h_{jk\nu_k}$ ($\nu_k = 1 \ldots N_k$ nodes) for F_j (e.g., Wood, 2017). Each two-dimensional function $f_{jkk'}, k' > k = 1 \ldots K-1$ is specified as a tensor product basis of the form

$$h_{jkk'}(\tilde{x}_{ik}, \tilde{x}_{ik'}) = \sum_{\nu_k=1}^{N_k} \sum_{\nu_{k'}=1}^{N_{k'}} \beta_{j\nu_k\nu'_k} h_{jk\nu_k}(\tilde{x}_{ik}) h_{jk'\nu_{k'}}(\tilde{x}_{ik'}). \tag{18}$$

where $h_{jk\nu_k}$ and $h_{jk'\nu_{k'}}$ are the known basis functions for the covariates $\tilde{x}_K, \tilde{x}_{k'}$ and $\beta_{j\nu_k\nu'_k}$ are parameters to be estimated (see details in Wood, 2006, 2017). Tensor products allow for better multidimensional smoothness and are invariant to linear transformations of the covariates. In contrast to, for example, cubic basis splines, thin plate splines use as many basis functions as data points, so no specific decision about the number of nodes and the node placement are needed at this point. Further technical details on the construction of $\Im$, the basis functions, and further analytic results to solve Eq. (17) can be found in Wahba (1990); Wood (2003, 2017).

Estimates for F_j are obtained via regularized estimation using thin plate regression splines (Wood, 2003, 2011). Thin plate regression splines are an approximation of the thin plates that will find an optimal subset of the nodes that can approximate the function F_j (technically, an $n \times n$ matrix E that is composed of the basis functions for each data point is approximated by a lower rank eigen approximation E_q of dimension q; see also, e.g., Hastie, 1996). This reduces the computational cost from $O(n^3)$ to $O(n^2 q)$ (where n is the number of parameters and q depends on the chosen rank of the approximation; for technical details see Wood, 2003, 2017).

One of the advantages of thin plate regression splines is that they do not depend as heavily on the actual number and location of the nodes selected as

other splines. The actual dimension (and wiggliness of F_j) is controlled by the penalty parameter λ (Wood, 2003). The number of nodes (q) that are selected is an upper bound of the parameter space and needs to be large enough so that it does not restrict this space. This number is typically considerably smaller than the sample size. A number of nodes that is too large affects the computational costs but not the actual result (because the dimension is shrunken to the relevant dimension). In practical situations, the relevant dimension can be investigated with the effective number of degrees of freedom (EDF) that relates to the number of parameters that are not shrunken exactly to zero. The EDF should be (considerably) smaller than the chosen number of nodes (Wood, 2003, 2017).

The model is estimated using a REML estimator (Reiss & Ogden, 2009; Wood, 2011). This has the advantage that the penalty parameter λ is estimated and only depends on the nonlinearity present in the data (instead of choosing λ subjectively). The REML estimator can be viewed as an empirical Bayes approach for the marginal likelihood which is the same as the REML estimator for (generalized) multilevel models (Wood, 2017). Here, the coefficients are treated as random effects that are sampled from a normal distribution (i.e., similar to a ridge regression). Model comparisons can be conducted using likelihood ratio tests based on the approximation of the test statistic provided, for example, as in Wood (2017). Note that the test should be interpreted with caution because it might lead to too liberal significance test under certain conditions (cf. Scheipl, Greven, & Kuchenhoff, 2008).

To identify the basic model in Eq. (2), not all coefficients for the estimated functions f_j' from Eq. (16) are allowed to be zero to avoid collinearity in the weight matrix $\mathbf{W}(\tilde{\mathbf{x}}_i)$. Compared to the original RPM, any kind of interaction (i.e., $f_j'(X_k) \times R \rightarrow M_j$), and not just the linear or otherwise parametrically specified interaction ($X_k \times R \rightarrow M_j$) contributes to reducing this collinearity. The model also reduces the multicollinearity between η_{1k} and X_k, which are needed to identify the interaction effect $R \times X_k \rightarrow Y$ without the necessity to model specific nonlinear interactions. Instead, the model directly draws from any kind of (unspecific) nonlinearity that is present in the data. In principle, it is possible to include higher order interactions between the covariates. For binary or categorical predictor variables, identity functions are used for the covariate functions instead of basis expansions.

A semi-parametric model for G

To increase efficiency of the estimator and to reduce potential misspecification compared to a purely parametric specification of $\tilde{G}$, again thin plate regression splines are used to approximate G, which leads to a reduction of the misspecification of the function and a higher predictive power:

$$G(\tilde{\mathbf{x}}_i) = \alpha_Y + g_1(\tilde{x}_{i1}) + \ldots + g_K(\tilde{x}_{iK}) + g_{12}(\tilde{x}_{i1}, \tilde{x}_{i2}) + \ldots + g_{(K-1)K}(\tilde{x}_{i(K-1)}, \tilde{x}_{iK}), \quad (19)$$

where α_Y is an intercept. The functions $g_k, k = 1 \ldots K$ and $g_{kk'}, k' > k = 1 \ldots K-1$ for each continuous covariate X_k are estimated as described above.

In the next section, a simulation study is presented that will investigate how the semi-parametric formulation of the weight estimates and the covariate function G affects the performance of the rank preserving models.

Simulation study

This simulation study has three scopes: First, the efficiency and power of the nRPM is compared to the original RPM implementation. Second, the ability to estimate interaction effects in the nRPM and RPM framework is investigated. Third, the robustness of the nRPM and the RPM to both the violation of the no-effect-modifier and the no-unmeasured-confounder assumption is investigated and compared to the standard mediator model using a no-unmeasured-confounder assumption within a regression framework ("Baron-Kenny-Model"; BKM).

Design

Data were generated for Y, M and $K = 2$ covariates by

$$m_i = \sum_k f_k(x_{ik}, r_i) + 0.5 \cdot r_i + \gamma_{U0} u_i + \epsilon_{Mi} \quad (20)$$

$$y_i = \sum_k g_k(x_{ik}) + 0.125 \cdot r_i + \theta_M m_i + \theta_{RX_1} r_i x_{i1} + \theta_{MX_1} m_i x_{i1} + \theta_{MR} m_i r_i + \gamma_{UR} u_i r_i + \gamma_{UM} u_i m_i + \gamma_{U0} u_i + \epsilon_{Yi} \quad (21)$$

with normally distributed errors $\epsilon_{Mi}, \epsilon_{Yi}$, and variances such that M and Y had unit variances. The indices of the θ coefficients refer to the respective effects in model to facilitate presentation in this section; the γ coefficients indicate parts of the population model that are used to induce potential misspecifications. The covariates followed a standard normal distribution with a correlation of $r = .2$, which reflects a

Table 1. Simulation conditions for the data generation.

Factor	Levels			
Across all studies				
Nonlinearity	None	Medium	Strong	
Interaction β_{RX} ($R \times X \rightarrow M$)	0.05	0.10	0.15	
Sample size (N)	200	500	1,000	
Study 1				
CME θ_M	0.00	0.10	0.20	0.30
Study 2				
Type of interaction	θ_{RX}	θ_{MX}	θ_{MR}	
Size of interaction	0.10	0.15		
Study 3				
Type of misspecified interaction	γ_{UM}	γ_{UR}		
Size of misspecified interaction	0.10	0.20		
Size of linear confounder effect (γ_{U0})	0.00	0.30		

situation where the researcher uses variables with a small amount of overlapping information. The treatment variable was generated with a binomial distribution and probability $P(R = 1) = .5$. The unobserved confounder $U \sim N(0, 1)$ was uncorrelated with the covariates and treatment.

Variation in the covariate functions f_k, g_k was induced using random polynomial functions of order 5 for each of the $k = 1, 2$ covariates. Three conditions of nonlinearity were selected: None, medium, and strong nonlinearity. The resulting (non)-linear relationships were similar to those found in the empirical example presented in the next section. Under the condition of medium or strong nonlinearity, the average explained variance increased by $\Delta R^2 \approx 0.03$ or $\Delta R^2 \approx 0.07$ for a comparison between a linear vs. a correctly specified polynomial covariate function.

For all simulation studies, the interaction effect $R \times X \rightarrow M$ (β_{RX}) was varied on three levels: $0.05, 0.10, 0.15$, which reflect typical sizes for interaction effects in empirical settings (Chaplin, 1991, 2007). The no-unmeasured-confounder assumption was violated with $\gamma_{U0} = 0.3$, reflecting a typical situation where not all relevant covariates are included in the model. This resulted in a residual correlation of $\rho_{\epsilon_y \epsilon_m} \approx 0.1$ when U was omitted from the model. Under each scenario, data were generated for $R =$ 500 replications for sample sizes of $N =$ 200, 500, and 1,000. All conditions are summarized in Table 1.

For simulation study 1, no violation of the no-effect-modifier assumption was induced, that is, $\gamma_{UR} = \gamma_{UM} = 0$ and all interaction effects were zero, that is, $\theta_{RX_k} = \theta_{MX_k} = \theta_{MR} = 0$ ($k = 1, 2$). The CME was varied on the levels $\theta_M = 0, 0.1, 0.2, 0.3$ (standardized coefficients). In simulation study 2, the performance of the RPM and the nRPM to estimate interaction effects was investigated: six population models where specified by setting $\theta_{RX_1}, \theta_{MX_1}$, or θ_{MR} to 0.10 or 0.15 (in line with levels chosen for β_{RX}). θ_M was set to 0. In simulation study 3, the robustness of the (n)RPM was investigated by setting γ_{UR} or γ_{UM} to 0.1 or 0.2. In addition, a scenario was included where the no-unmeasured-confounder assumption was met ($\gamma_{U0} = 0.0$). The CME was set to zero ($\theta_M = 0$) in order to investigate the Type I error rates.

Analysis

In study 1, data were analyzed with the RPM and nRPM that only included θ_R, θ_M as CDE and CME (see Eq. (2)). The nRPM used thin plate regression splines with a REML estimator and the RPM used a linear covariate function. In a preliminary study, the effect of different node numbers was tested for the nRPM (10, 20, 30, and 40). Results were very similar across conditions with average EDFs lying between 2.00 and 2.45 for linear, between 4.60 and 11.61 for medium, and between 5.75 and 18.12 for strong nonlinearity. Hence, for all studies, 20 nodes were used which allowed the basis space to be large enough for any kind of imposed nonlinearity in this simulation study and to keep the computational burden to a reasonable amount. Standard errors were bootstrapped for the RPM and the nRPM with 200 bootstrap replications for each data set.

In study 2, a correctly specified model including the respective type of interaction effect was used. For example, for a population model including $\theta_{RX_1} \neq 0$, both covariate by treatment interactions simultaneously ($\theta_{RX_1}, \theta_{RX_2}$) were included in addition to θ_R and θ_M.

In study 3, each data set was analyzed with four different models: the RPM, the nRPM (as specified in study 1), a standard mediator model using the no-unmeasured-confounder assumption (given the covariates; BKM) and a mediator model using a thin plate regression spline to account for the nonlinear relationships of the covariates in the model (nBKM). Here, the focus lies on the effect of M on Y conditional on R, which is indicative for the CME in this model specification.

Data were generated in R (R Core Team, 2018). All models were implemented in R using the packages rootSolve (for the estimating equations; Soetaert, 2009), and mgcv (for the estimation of the splines; Wood, 2011).

Results

Study 1

The results will focus on the CME θ_M that allows one to investigate if the intermediate variable functions as

Table 2. Study 1: percent significant results for the CME θ_M in the RPM and the nRPM.

		$N = 200$			$N = 500$			$N = 1{,}000$		
θ_M	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
					No nonlinearity					
0.0	RPM	2.0	2.2	3.8	1.8	2.4	4.0	0.8	3.4	6.2
0.0	nRPM	4.4	1.8	3.6	1.4	2.0	3.2	1.4	4.2	6.4
0.1	RPM	4.2	7.0	7.0	4.8	7.0	10.2	4.8	8.8	13.0
0.1	nRPM	6.0	4.2	4.4	4.6	7.4	10.4	5.2	8.6	14.8
0.2	RPM	7.6	10.2	17.4	7.0	15.4	29.4	9.4	24.4	46.2
0.2	nRPM	8.8	6.2	16.0	9.8	17.0	27.0	9.2	27.0	49.2
0.3	RPM	10.0	16.6	30.4	18.0	32.8	59.8	22.2	51.8	82.8
0.3	nRPM	10.2	15.2	26.8	18.0	34.8	57.2	23.6	54.8	84.4
					Medium nonlinearity					
0.0	RPM	1.8	3.0	2.6	3.2	1.2	2.4	1.8	4.2	5.4
0.0	nRPM	4.2	2.6	3.8	4.0	3.4	4.4	2.8	5.2	4.6
0.1	RPM	4.4	4.0	5.6	4.2	6.2	10.0	2.4	6.2	12.2
0.1	nRPM	5.2	3.8	6.0	4.4	7.0	12.0	6.4	9.2	17.2
0.2	RPM	4.8	6.2	10.4	8.4	7.8	17.0	7.6	15.8	29.6
0.2	nRPM	7.8	8.6	12.4	11.6	16.8	25.4	15.6	32.2	47.6
0.3	RPM	12.4	16.6	22.6	10.8	26.4	39.0	17.0	34.4	57.4
0.3	nRPM	17.4	22.8	27.2	23.6	43.8	58.0	32.4	57.8	83.0
					Strong nonlinearity					
0.0	RPM	5.8	4.0	5.4	3.4	4.2	3.0	3.0	2.2	2.4
0.0	nRPM	3.2	2.8	2.2	2.8	3.2	4.8	4.2	5.4	5.2
0.1	RPM	3.4	2.8	4.4	6.4	3.6	3.8	5.8	5.2	4.6
0.1	nRPM	4.2	5.4	4.8	8.6	6.6	10.6	9.6	13.8	17.6
0.2	RPM	7.8	6.8	6.4	7.2	7.0	8.4	6.0	6.8	10.4
0.2	nRPM	10.4	11.2	13.0	14.4	17.4	31.0	18.6	31.0	47.6
0.3	RPM	12.0	14.0	13.2	9.6	11.6	14.4	14.6	12.4	22.6
0.3	nRPM	20.0	23.6	30.6	28.6	40.0	56.0	36.8	56.2	74.6

For $\theta_M = 0$, they can be interpreted as Type I error rates, for $\theta_M \neq 0$, they indicate power.
N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$, RPM: rank preserving model, nRPM: nonlinear rank preserving model.

a mediator. Tables for the results can be found in the appendix. Coverage rates for the CME lay between 93.6% and 99.2% for the RPM and between 91.8% and 99.2% for the nRPM. On average, coverage rates were very similar for the RPM (96.6%) than for the nRPM (96.4%).

Table 2 shows the results for the percent of significant parameter estimates (based on the Wald-test for the parameter estimates and their bootstrapped standard errors). Type I error rates (for $\theta_M = 0$) were mostly kept at the nominal rate of 5% across all conditions. Results were very similar for both the RPM (on average 3.1% and maximal 6.2%) and for the nRPM (on average 3.6% and maximal 6.4%). The power to detect an effect $\theta_M \neq 0$ increased with sample size, effect size, and interaction effect β_{RX}. Under the condition of a linear covariate function, the RPM had a power between 4.2% and 13.0% for $\theta_M = 0.1$, between 7.0% and 46.2% for $\theta_M = 0.2$, and between 10.0% and 82.8% for $\theta_M = 0.3$. This power strongly decreased if the covariate function was nonlinear (and G was misspecified in the RPM) with a power for large sample sizes ($N = 1{,}000$) below 57.4% and below 22.6% for medium and strong nonlinearity, respectively. The power for the nRPM was very

Table 3. Study 1: relative efficiency of the CME θ_M for RPM vs. nRPM.

		$N = 200$			$N = 500$			$N = 1{,}000$		
θ_M	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
					No nonlinearity					
0.0		1.08	1.04	0.94	1.01	0.99	1.06	1.07	1.05	0.99
0.1		1.13	1.12	0.88	1.01	1.00	0.94	1.09	1.00	0.97
0.2		1.17	1.02	1.02	1.18	1.00	1.12	1.10	0.99	1.03
0.3		1.06	1.02	0.93	1.02	0.96	0.95	1.10	1.08	1.13
					Medium nonlinearity					
0.0		1.28	1.38	1.29	1.41	1.40	1.37	1.49	1.36	1.40
0.1		1.36	1.20	1.14	1.41	1.42	1.37	1.33	1.37	1.40
0.2		1.50	1.35	1.15	1.43	1.40	1.26	1.48	1.35	1.55
0.3		1.55	1.41	1.15	1.57	1.38	1.38	1.56	1.47	1.32
					Strong nonlinearity					
0.0		1.49	1.66	1.89	1.80	2.51	2.02	1.97	2.54	2.45
0.1		1.65	1.66	1.93	1.89	2.02	2.37	2.52	2.47	2.35
0.2		1.69	1.64	1.78	2.19	2.08	2.82	2.25	2.84	2.39
0.3		2.12	1.99	1.81	2.00	2.43	2.35	2.25	2.99	2.73

Values above 1 indicate smaller MAD for the nRPM.
N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$.

similar across the different amounts of nonlinearity in the covariate function. Under the condition of a linear covariate function, it lay between 4.2% and 14.8% for $\theta_M = 0.1$, between 6.2% and 49.2% for $\theta_M = 0.2$, and between 10.2% and 84.4% for $\theta_M = 0.3$. While the power for nRPM and the RPM were similar for a linear covariate function, a maximal difference of 52.0% was found for $N = 1000, \theta_M = 0.3, \beta_{RX} = 0.15$, and strong nonlinearity.

To assess estimation efficiency, the Monte Carlo median absolute deviations (MAD)[2] were investigated. The bootstrapped standard errors recovered the empirical distribution (Monte Carlo MAD) very well with a rate of average SE vs. MAD of 1.04 for both estimators across conditions. Table 3 shows the relative efficiency of the RPM vs. the nRPM (i.e., a ratio between MAD(RPM)/MAD(nRPM) under each condition). Across conditions with the linear covariate function, the RPM and nRPM had a similar efficiency, as indicated with values close to 1 (ranging from 0.88 to 1.18 and an average of 1.04). For the nonlinear covariate functions, the nRPM was always more efficient than the RPM, which was indicated by values above 1. The ratios lay between 1.14 and 1.57 with an average of 1.38, and between 1.49 and 2.99 with an average of 2.15 for medium and strong nonlinearity, respectively. The ratios were generally larger with increasing sample size and nonlinearity but were fairly

[2]The MAD is the (standardized) median absolute deviation (Huber & Ronchetti, 2009). It is defined for a parameter γ and $r = 1...R$ replications as $MAD(\gamma) = Mdn|\hat{\gamma}_r - Mdn(\hat{\gamma})|$ (where Mdn is the median) and it is used as a robust measure of the variability of the parameter estimates (instead of a Monte Carlo standard deviation). The advantage of the MAD in this simulation is that replications with extreme estimates had not to be deleted – which in many situations is based on a subjective decision and might skew the performance of the estimator.

Table 4. Study 2: percent significant results for the interaction effects that were non-zero in the population (i.e., power) and their relative efficiency.

Nonlinearity		$N = 200$		$N = 500$		$N = 1{,}000$	
	$\theta_{..} =$	0.10	0.15	0.10	0.15	0.10	0.15
		Power					
		$\theta_{RX} \neq 0$					
None	RPM				n.i.		
	nRPM	10.9	29.1	38.0	63.3	51.9	75.6
Medium	RPM				n.i.		
	nRPM	16.2	39.0	46.4	82.5	74.7	94.5
Strong	RPM				n.i.		
	nRPM	17.2	42.2	57.4	90.9	87.9	99.3
		$\theta_{MX} \neq 0$					
None	RPM	8.3	9.0	7.2	10.9	10.9	17.5
	nRPM	5.1	6.7	6.2	9.6	12.2	19.3
Medium	RPM	10.1	11.0	9.7	10.8	9.9	14.5
	nRPM	5.3	5.8	6.0	11.7	11.2	21.8
Strong	RPM	10.3	10.6	10.3	11.5	10.8	13.8
	nRPM	4.3	4.0	6.7	11.6	14.3	28.8
		$\theta_{MR} \neq 0$					
None	RPM	8.2	11.6	11.3	18.8	20.6	34.1
	nRPM	6.1	8.9	9.5	18.8	19.4	33.9
Medium	RPM	8.5	9.0	9.4	12.5	11.3	18.5
	nRPM	4.4	6.7	10.2	21.0	22.9	39.5
Strong	RPM	6.7	9.5	5.9	7.5	4.9	7.4
	nRPM	3.6	4.9	10.7	19.9	20.2	39.1
		Relative efficiency RPM vs. nRPM					
		$\theta_{MX} \neq 0$					
None		1.04	0.98	1.01	1.07	1.08	1.00
Medium		1.27	1.23	1.41	1.53	1.63	1.50
strong		1.53	1.44	2.08	2.00	1.99	2.10
		$\theta_{MR} \neq 0$					
None		1.08	1.02	1.15	1.14	1.11	1.10
Medium		1.73	1.63	1.91	1.85	1.80	1.94
Strong		3.04	3.07	3.29	3.47	3.17	3.43

Values for the relative efficiency above 1 indicate a higher efficiency for the nRPM (Results are averaged across conditions of β_{RX}).

N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M, \theta_{RX}, \theta_{MX}, \theta_{MR}$: interaction effects $R \times X \rightarrow Y, M \times X \rightarrow Y, M \times R \rightarrow Y$, RPM: rank preserving model, nRPM: nonlinear rank preserving model, n.i.: not identified.

similar across different interaction effects β_{XR} or θ_M. The RMSE showed very similar results because the bias of the estimators was negligible (see Tables A1 and A2 in the Appendix). The RMSE and the MAD were correlated at .999, indicating that the found higher efficiency of the nRPM also translated into a higher estimation accuracy.

Study 2

In study 2, average Type I error rates for θ_M for the RPM and the nRPM were 9.3% and 5.2% (when θ_{MX} was part of the model), or 7.8% and 4.5% (when θ_{MR} was part of the model), respectively. For the nRPM, the average Type I rate for θ_M in θ_{RX} models lay at 2.7% (this interaction was not identified for the RPM due to the linear model for the weight estimation). Average Type I error rates for the interaction effects in the nRPM (based on second covariate X_2) were below 6.8% in θ_{RX} models and below 8.0% for θ_{MX} models. The RPM provided higher Type I error rates for the interaction effect in the θ_{MX} models between 1.4% and 12.8%.

Table 4 illustrates the power to detect interaction effects with the RPM and nRPM. Results were summarized across conditions of β_{RX} because they were virtually identical. The power increased with sample size and effect size (as expected). For the nRPM, the power also increased with the amount of nonlinearity. For example, for the treatment by covariate interaction, it increased from 75.6% to 99.3% for $N =$ 1,000 and $\theta_{RX} = 0.15$. The power to detect the other interaction effects was lower. For smaller interaction effects of $\theta_{MX} = 0.10$ (or $\theta_{MR} = 0.10$), the nRPM produced between 11.2% and 14.3% (19.4% and 22.9%) significant results at $N =$ 1,000. For larger interaction effects, the power for $N =$ 1,000 lay between 19.3% and 28.8% (33.9% and 39.5%). For small sample sizes ($N =$ 200), the power to detect any larger interaction effect lay between 4.0% (θ_{MX}) and 42.2% (θ_{RX}). The power of the RPM to detect interaction effects lay below the one of the nRPM when sample size was larger than $N =$ 200. For large sample sizes ($N =$ 1,000), it was closest to the power of the nRPM for a linear covariate function with a maximum of 17.5% (or 34.1%) to detect a large interaction effect of $\theta_{MX} = 0.15$ (or $\theta_{MR} = 0.15$). When the covariate function was nonlinear in the population, the power of the RPM to detect any of the interaction effects lay below 14.5% ($\theta_{MX} = 0.15$) or 18.5% ($\theta_{MR} = 0.15$) for $N =$ 1,000. For small sample sizes (N=200), the power was slightly larger than for the nRPM, but the power was still low and lay between 6.7% and 11.6%.

The relative efficiency for θ_{MX} and θ_{MR} showed MAD ratios of the RPM vs. the nRPM between 0.98 and 1.14 (where values above 1 indicated smaller MAD's for the nRPM). For nonlinear covariate functions, the relative efficiency increased from 1.23 to 2.10 for θ_{MX}, which implied that the nRPM's MAD were as small as about half the size of the RPM's MAD. Similarly, the interaction effect θ_{MR} ratios lay between 1.63 and 3.47 when the covariate function was nonlinear.

Study 3

The final simulation investigated the robustness of the (n)RPM. Table 5 shows the Type I error rates for a situation, where an unobserved confounder is present in the data that interacts with the treatment (violation of no-effect-modifier assumption). Results for medium or strong interaction ($\gamma_{UR} = 0.1$ or 0.2) as well as those for an interaction between confounder and

Table 5. Study 3: percent significant estimates for the CME $\theta_M = 0$ (Type I error rate) for misspecified models with $\gamma_{UR} = 0.2$ (strong violation of no-effect modifier assumption).

		$N = 200$			$N = 500$			$N = 1{,}000$		
Nonlinearity	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
		$\gamma_{U0} = 0.0$								
None	RPM	0.4	0.0	0.2	0.2	0.8	3.0	0.0	2.0	3.8
	nRPM	0.2	0.4	0.2	0.0	1.0	2.2	0.4	2.4	3.4
	BKM	3.0	5.6	7.0	5.6	6.6	5.2	6.0	5.2	4.0
	nBKM	3.2	5.6	7.4	5.8	6.6	5.2	6.0	5.4	3.8
Medium	RPM	0.0	0.4	0.8	0.0	0.4	1.2	0.6	1.0	2.8
	nRPM	0.2	0.4	0.4	1.0	0.8	4.0	0.0	2.2	3.8
	BKM	7.8	10.0	6.2	12.2	12.4	13.4	15.4	16.2	12.8
	nBKM	6.0	7.0	4.4	5.4	6.2	5.6	6.0	7.0	6.2
Strong	RPM	0.0	0.4	0.4	0.0	0.2	0.8	0.0	0.2	0.6
	nRPM	0.4	0.4	0.2	0.4	1.0	2.4	1.4	2.6	4.2
	BKM	18.8	17.2	23.0	33.2	34.8	28.4	45.2	43.6	47.4
	nBKM	7.4	8.6	8.4	8.0	5.6	8.6	8.0	7.4	9.8
		$\gamma_{U0} = 0.3$								
None	RPM	0.0	0.2	1.2	0.0	1.4	3.2	0.2	1.8	3.6
	nRPM	0.0	0.2	0.8	0.4	0.6	3.4	0.0	2.4	3.8
	BKM	46.8	40.6	45.2	81.4	82.0	81.8	99.6	98.2	98.4
	nBKM	45.8	40.8	46.2	82.0	82.4	82.0	99.6	98.0	98.4
Medium	RPM	0.0	0.2	0.6	0.0	0.6	1.4	0.0	0.6	2.8
	nRPM	0.2	0.2	1.2	0.4	0.6	2.0	1.0	2.6	3.0
	BKM	40.8	38.2	36.8	73.8	69.4	73.2	90.4	91.0	87.6
	nBKM	43.0	39.8	38.4	80.4	77.0	80.8	97.8	96.4	96.4
Strong	RPM	0.8	0.6	0.2	0.6	0.4	0.6	0.2	0.4	0.4
	nRPM	0.2	0.4	0.6	0.6	0.8	2.2	0.8	3.2	5.8
	BKM	34.6	42.2	42.0	58.8	60.6	57.0	69.2	67.2	75.0
	nBKM	37.6	39.0	41.2	77.2	78.6	75.6	96.0	94.0	96.0

N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$, γ_{UR}: interaction effect of the confounder $U \times R \rightarrow Y$, γ_{U0}: linear effect $U \rightarrow M, Y$, RPM: rank preserving model, nRPM: nonlinear rank preserving model, BKM: Baron-Kenny model, nBKM: nonlinear Baron-Kenny model.

Table 6. Study 3: relative accuracy for the CME $\theta_M = 0$ compared to the nRPM for misspecified models with $\gamma_{UR} = 0.2$.

		$N = 200$			$N = 500$			$N = 1{,}000$		
Nonlinearity	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
		$\gamma_{U0} = 0.0$								
None	RPM	1.22	1.01	1.01	1.12	1.04	0.94	1.01	0.93	1.05
	BKM	0.13	0.16	0.28	0.09	0.19	0.23	0.09	0.14	0.27
	nBKM	0.13	0.17	0.28	0.09	0.19	0.23	0.09	0.14	0.27
Medium	RPM	1.40	1.21	1.27	1.38	1.39	1.21	1.49	1.32	1.35
	BKM	0.15	0.20	0.23	0.13	0.21	0.29	0.14	0.24	0.34
	nBKM	0.14	0.17	0.22	0.12	0.18	0.25	0.11	0.19	0.28
Strong	RPM	1.62	1.82	1.68	1.98	2.17	2.13	2.26	2.60	2.43
	BKM	0.22	0.28	0.34	0.22	0.36	0.38	0.25	0.40	0.60
	nBKM	0.16	0.20	0.24	0.12	0.19	0.23	0.10	0.18	0.28
		$\gamma_{U0} = 0.3$								
None	RPM	1.05	1.10	0.99	1.24	1.02	0.97	1.05	0.95	1.03
	BKM	0.23	0.36	0.49	0.31	0.48	0.72	0.36	0.69	1.02
	nBKM	0.23	0.36	0.50	0.31	0.48	0.72	0.36	0.69	1.02
Medium	RPM	1.23	1.19	1.26	1.37	1.48	1.20	1.40	1.35	1.26
	BKM	0.25	0.34	0.44	0.30	0.46	0.73	0.36	0.64	0.98
	nBKM	0.25	0.34	0.45	0.30	0.47	0.74	0.36	0.66	0.99
Strong	RPM	1.61	1.59	1.80	1.69	2.09	2.51	2.29	2.31	2.28
	BKM	0.26	0.34	0.49	0.34	0.50	0.73	0.42	0.67	0.96
	nBKM	0.25	0.32	0.45	0.32	0.45	0.71	0.38	0.60	0.86

Values for the relative accuracy above 1 indicate a higher accuracy for the nRPM.

N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$, γ_{UR}: interaction effect of the confounder $U \times R \rightarrow Y$, γ_{U0}: linear effect $U \rightarrow M, Y$, RPM: rank preserving model, nRPM: nonlinear rank preserving model, BKM: Baron-Kenny model, nBKM: nonlinear Baron-Kenny model.

mediator (γ_{UM}) were very similar (see Tables A3 and A4 in the Appendix). Here, results are presented only for the larger misspecification $\gamma_{UR} = 0.2$.

Both the RPM and the nRPM were rather robust against the violations imposed on the model. The RPM's Type I error rates lay between 0.0% and 3.8% and kept the nominal level across all conditions. The nRPM had slightly higher Type I error rates ranging from 0.0% to 5.8% with only a single condition having a Type I error rate above 5% ($\gamma_{U0} = 0.3, N = 1000$, strong nonlinearity). Both the BKM and nBKM were strongly affected by the misspecification due to $\gamma_{U0} = 0.3$, as it was expected. For a linear covariate function, the Type I error rates for both BKM and nBKM lay between 40.6% and 46.8% for $N = 200$, between 81.4% and 82.4% for $N = 500$, and between 98.0% and 99.6% for $N = 1{,}000$. For the nBKM, similar rates could be observed for medium and strong nonlinear covariate functions while they were somewhat lower for the BKM (especially with larger sample sizes). For the BKM, inflated Type I error rates were observed even when the no-unmeasured-confounder assumption was met ($\gamma_{U0} = 0.0$): if the covariate functions were nonlinear, they increased to between 6.2% and 16.2% (medium nonlinearity), and between 17.2% and 47.4% (strong nonlinearity).

Finally, the relative accuracy was investigated. For the accuracy, the RMSE was calculated which comprises estimation bias and efficiency, and is more informative for model performance than efficiency in situations where estimators are biased. The relative accuracy was calculated by dividing the respective RMSE by the nRPM's RMSE (values larger than 1 indicate a more accurate performance of the nRPM). The relative accuracy is illustrated in Table 6. The RPM was less accurate than the nRPM when the covariate function was nonlinear with ratios ranging between 1.19 and 2.60; they were mainly influenced by the amount of nonlinearity in the covariate function. For a linear covariate function, ratios ranged between 0.93 and 1.24. Both the BKM and the nBKM produced more accurate results than the nRPM as long as the no-unmeasured-confounder assumption was not violated with ratios ranging between 0.09 and 0.60 for the BKM and between 0.09 and 0.28 for the nBKM. Ratios increased for the BKM when the covariate function was nonlinear especially for larger sample sizes. When the no-unmeasured-confounder assumption was violated, the nRPM produced results with a similar accuracy for large sample sizes ($N =$ 1,000) and strong interaction effect $\beta_{RX} = 0.15$, with ratios ranging between 0.96 and 1.02 or between 0.86

and 1.02 compared to the BKM or the nBKM, respectively.

Empirical example

In this section, an empirical example is presented to illustrate the application of the nRPM. Results are compared to the original RPM and to the (n)BKM. The data set used is based on the AIDS Clinical Trials Group Study 175, a publicly available data set from the R package `speff2trial` (Hammer et al., 1996; Juraska et al., 2012). Patients received either a treatment of zidovudine only or an alternative treatment (zidovudine and didanosine, zidovudine and zalcitabine, or didanosine treatment). The main outcome variable was the CD4 cell count at week 96 (±5). Here, it is investigated if the CD4 and the CD8 cell count at week 20 (±5) mediated the effect of the treatment on the outcome. Several covariates were measured at baseline, which included binary variables (hemophilia, homosexual activity, history of intravenous drug use, if the patient had received an earlier non-zidovudine antiretroviral therapy, if patients had used zidovudine in the 30 days or directly prior to treatment initiation, race, gender, antiretroviral history, a symptomatic indicator), ordinal variables (Karnofsky score), and continuous variables (age, weight, the number of days of previously received antiretroviral therapy ["days"], CD4, and CD8 cell count at baseline). Data were used for complete cases only ($N =$ 1,342).

Initial analyses indicated significant treatment effects for both CD4 cell counts at week 96 ($t = -4.99, df = 560.87, p<.001$) and week 20 ($t = -4.57, df = 600.67, p<.001$). For the CD8 cell counts at week 20, no treatment effect was found ($t = -0.01, df = 544.03, p = .99$).

Models

Three different models were tested for each mediator candidate.[3] Model 1 used a linear structural model that included all covariates only in the covariate function G (see Eq. (2)). Model 2 additionally included interaction effects between treatment and all five continuous covariates ($R \times X \rightarrow Y$). Model 3 included only the subset of continuous covariates found irrelevant (no interaction or linear effects in models 1 or 2) and their interactions with the treatment (i.e., age, weight, and days). Model 3 was misspecified because it did not include relevant covariates or their interactions; it represented a situation where both no-unmeasured-confounder and no-effect-modifier assumptions were (actively) violated.

All three models were estimated using the nRPM, the RPM, the BKM, and the nBKM. Semi-parametric nonlinear functions were defined using thin plate regression splines with 60 nodes for the nRPM and nBKM (due to the model with five continuous covariates, the dimension of the basis space needed to be larger than in the simulation study). The EDF were 55 for the models 1 and 2 with five covariates and 9 for the model 3 with three covariates. All models were also run with different number of initial nodes (between 10 and 90). As expected, the estimates were virtually identical (because in each scenario, the EDF were again always about 55 and about 9) as is illustrated in Figure B1 in the Appendix for the CME in the nRPM. Standard errors were bootstrapped for the RPM and nRPM with 1,000 bootstrap replications per model. For the RPM and BKM, linear relationships were assumed. R code for all models can be downloaded from the first author's website.

Weight matrix

The (nonlinear) spline functions in the nRPM for the continuous covariates in model 1 increased the explained variance by $\Delta R^2 = .014$ ($\Delta\chi^2 = 80.69, df = 100, p = .017$)[4] and $\Delta R^2 = .059$ ($\Delta\chi^2 = 114.59, df = 104.34, p<.001$) for the CD4 and CD8 cell counts at week 20 compared to the linear functions used in the RPM. Removing all interaction effects in the nRPM between treatment and covariates did not change the model fit for the CD4 cell counts ($\Delta\chi^2 = 34.52, df = 65, p = .764$) but for the CD8 cell counts ($\Delta\chi^2 = 61.46, df = 69.34, p<.001$). For the RPM both CD4 ($\Delta\chi^2 = 4.58, df = 15, p = .948$) and CD8 ($\Delta\chi^2 = 2.55, df = 15, p = .981$) cell count models did not include any significant interactions.

The minimal eigenvalues of the weight covariance matrices for all models lay between 0.039 and 0.643 for the nRPM, and between 0.003 and 0.038 for the RPM, indicating that the weights were closer to collinear for all RPM's due to the non-existing interaction effects. This implied that the nRPM should have a higher efficiency and power in the next step.

[3] A model including both mediator candidates simultaneously was also tested but due to the increased model complexity in relation to the sample size, all confidence intervals became very wide for the (n)RPM and the results lost their illustrative character.

[4] Model comparisons are conducted using an approximation of the likelihood ratio test statistic that is similar to tests for multilevel models (see details in Wood, 2017).

Table 7. Results for the empirical example (Y = CD4 cell count at week 96).

		M = CD4 at 20 weeks			M = CD8 at 20 weeks		
		$\hat{\theta}$ (SE)	CI	p	$\hat{\theta}$ (SE)	CI	p
				Model 1			
nRPM	M	0.37 (0.18)	[0.02;0.73]	.04	0.12 (0.29)	[−0.45;0.68]	.69
	R	0.24 (0.07)	[0.10;0.39]	<.01	0.36 (0.07)	[0.22;0.50]	<.01
RPM	M	0.42 (0.32)	[−0.21;1.04]	.19	0.67 (0.56)	[−0.43;1.76]	.23
	R	0.23 (0.12)	[0.00;0.46]	.05	0.36 (0.06)	[0.24;0.49]	<.01
BKM	M	0.50 (0.02)	[0.45;0.55]	<.01	0.17 (0.03)	[0.11;0.24]	<.01
	R	0.20 (0.05)	[0.11;0.29]	<.01	0.37 (0.05)	[0.27;0.47]	<.01
				Model 2			
nRPM	M	0.36 (0.18)	[0.01;0.71]	.04	0.11 (0.28)	[−0.43;0.65]	.69
	R	0.25 (0.07)	[0.10;0.39]	.01	0.36 (0.07)	[0.22;0.50]	<.01
	$R \times$ age	0.00 (0.06)	[−0.12;0.12]	.98	0.00 (0.08)	[−0.17;0.16]	.98
	$R \times$ weight	0.03 (0.05)	[−0.07;0.14]	.54	0.03 (0.08)	[−0.12;0.18]	.69
	$R \times$ days	−0.04 (0.06)	−0.15;0.07]	.46	−0.05 (0.08)	−0.21;0.11]	.55
	$R \times CD40$	0.03 (0.06)	[−0.09;0.15]	.61	0.04 (0.08)	[−0.12;0.20]	.59
	$R \times CD80$	0.11 (0.06)	[−0.01;0.23]	.06	0.11 (0.09)	[−0.06;0.28]	.22
RPM	M	0.46 (0.34)	[−0.21;1.13]	.18	0.43 (0.95)	[−1.43;2.29]	.65
	R	0.22 (0.12)	[−0.01;0.45]	.06	0.37 (0.07)	[0.24;0.50]	<.01
	$R \times$ age	0.00 (0.06)	[−0.11;0.11]	.98	0.02 (0.09)	[−0.15;0.19]	.84
	$R \times$ weight	0.04 (0.05)	[−0.06;0.13]	.43	0.02 (0.08)	[−0.14;0.19]	.80
	$R \times$ days	−0.03 (0.05)	[−0.13;0.07]	.56	−0.01 (0.08)	[−0.16;0.14]	.89
	$R \times CD40$	0.00 (0.06)	[−0.11;0.11]	.96	0.02 (0.08)	[−0.13;0.18]	.78
	$R \times CD80$	0.10 (0.04)	[0.02;0.18]	.02	0.08 (0.08)	[−0.09;0.24]	.35
BKM	M	0.50 (0.02)	[0.45;0.55]	<.01	0.17 (0.03)	[0.10;0.24]	<.01
	R	0.20 (0.05)	[0.11;0.29]	<.01	0.37 (0.05)	[0.26;0.47]	<.01
	$R \times$ age	0.00 (0.05)	[−0.09;0.09]	.98	0.01 (0.05)	[−0.10;0.11]	.92
	$R \times$ weight	0.04 (0.05)	[−0.06;0.13]	.44	0.04 (0.06)	[−0.07;0.14]	.52
	$R \times$ *days*	−0.03 (0.05)	[−0.12;0.06]	.48	−0.02 (0.05)	[−0.12;0.08]	.69
	$R \times CD40$	0.00 (0.05)	[−0.09;0.09]	.98	0.01 (0.05)	[−0.09;0.12]	.81
	$R \times CD80$	0.10 (0.05)	[0.01;0.19]	.03	0.09 (0.05)	[−0.01;0.19]	.09
				Model 3			
nRPM	M	0.47 (0.33)	[−0.17;1.12]	.15	−0.05 (0.34)	[−0.71;0.61]	.88
	R	0.20 (0.10)	[0.00;0.40]	.05	0.33 (0.07)	[0.20;0.46]	<.01
	$R \times$ age	0.01 (0.06)	[−0.11;0.13]	.85	0.00 (0.09)	[−0.17;0.17]	.99
	$R \times$ weight	0.10 (0.06)	[−0.02;0.23]	.10	0.16 (0.07)	[0.02;0.29]	.03
	$R \times$ days	−0.04 (0.06)	[−0.17;0.08]	.49	−0.06 (0.07)	[−0.20;0.08]	.39
RPM	M	0.30 (0.34)	[−0.36;0.96]	.37	−0.11 (0.32)	[−0.74;0.52]	.73
	R	0.24 (0.11)	[0.01;0.46]	.04	0.32 (0.06)	[0.20;0.44]	<.01
	$R \times$ age	−0.01 (0.07)	[−0.14;0.13]	.91	−0.03 (0.08)	[−0.19;0.13]	.74
	$R \times$ weight	0.11 (0.07)	[−0.03;0.25]	.11	0.15 (0.07)	[0.02;0.29]	.03
	$R \times$ days	−0.05 (0.06)	[−0.16;0.07]	.42	−0.05 (0.07)	[−0.19;0.08]	.44
BKM	M	0.65 (0.02)	[0.60;0.69]	<.01	0.04 (0.03)	[−0.02;0.09]	.17
	R	0.14 (0.05)	[0.04;0.23]	<.01	0.32 (0.06)	[0.19;0.44]	<.01
	$R \times$ age	0.00 (0.05)	[−0.09;0.10]	.98	−0.01 (0.06)	[−0.14;0.11]	.84
	$R \times$ weight	0.07 (0.05)	[−0.03;0.17]	.15	0.14 (0.07)	[0.02;0.27]	.03
	$R \times$ days	−0.04 (0.05)	[−0.13;0.05]	.41	−0.05 (0.06)	[−0.18;0.07]	.38

$\hat{\theta}$: parameter estimates, SE: Standard error, CI: 95% confidence interval, p: significance, RPM: rank preserving model, nRPM: nonlinear rank preserving model, BK M: Baron-Kenny model.

Results for the CD4 cell counts as mediator

Results for the mediator analysis can be found in Table 7. For all three models, the nRPM provided very similar results for the CME of the CD4 cell counts at week 20: They provided significant CME's for models 1 and 2, but not for model 3 with a size of 0.37 (and 95% confidence interval of $[0.02; 0.73]$), 0.36 $[0.01; 0.71]$, and 0.47 $[-0.17; 1.12]$ for models 1 through 3. In model 2, a close to significant interaction between baseline CD8 and treatment was identified ($\hat{\theta}_{R\times CD80} = 0.11, SE = 0.06, p = .06$). Omitting this interaction effect and the CD8 baseline measure in model 3 resulted in a slightly different estimate for the CME, but mainly affected its SE estimates that increased from 0.18 in model 2 to 0.33 in model 3.

The RPM provided fairly similar parameter estimates compared to the nRPM for all three models with a CME of 0.47 $[-0.21; 1.04]$, 0.46 $[-0.21; 1.13]$, and 0.30 $[-0.36; 0.96]$. Its standard errors were considerably larger and none of the CME's were significant, while the interaction effect $R \times CD8$ was significant in model 2.

Results for the BKM and nBKM were virtually identical and only those for the BKM are presented in Table 7. The CME was significant across all three models ($p < .01$). The CME for the CD4 cell counts was larger than in the (n)RPM with $\theta_M =$

0.50[0.45; 0.55] in models 1 and 2. Both models included the same set of covariates and hence controlled for confounders in the same way. This CME increased to 0.65 when CD4 and CD8 baseline cell counts were excluded from the model and the no-unmeasured-confounder assumption was intentionally violated; its confidence interval of [0.60; 0.68] did not overlap with those of models 1 and 2.

CD8 cell counts as mediator

Again, the results for the nRPM were similar across models. In contrast to the CD4 cell counts at week 20, the CD8 cell count at week 20 did not mediate the effect of the treatment on the CD4 cell count at week 96 with CME's of 0.12 [−0.45; 0.68], 0.11 [−0.43; 0.65], and −0.05 [−0.71; 0.61] in models 1 through 3. The CME varied more strongly for the RPM across models and resulted in wider confidence intervals with estimates of 0.67 [−0.43; 1.76], 0.43 [−1.43; 2.29], and −0.11 [−0.74; 0.52]. Finally, contrary to the nRPM, the BKM provided significant CME's for models 1 and 2, but not for model 3, again with confidence intervals for the CME that did not overlap between the first two and the last model.

Further analyses

Figure 2 illustrates the estimated relationships between the five continuous covariates and the two mediator variables, as well as the potential values $\hat{Y}_{00}$ based on nRPM's estimates for model 2 with the CD4 cell counts as mediator. The functions show that the set of irrelevant covariates have a close to zero relationship with the mediator candidates and the potential outcomes. For the two relevant variables, baseline CD4 and baseline CD8, some indications of nonlinearity can be observed. For the two mediator variables, these relationships varied across the two treatment groups.

Discussion

In this article, a specific nonlinear extension of the RPM was introduced and tested under a variety of empirically relevant scenarios. This extension allows one to approximate unknown functions between covariates and mediators or outcomes using a flexible nonlinear spline model. It overcomes the limitations of the original RPM, for which a misspecification of the covariate function impacts its efficiency. Simulation studies showed that the nRPM has more efficient estimates and an increased power compared to the RPM. This advantage could also be seen in the empirical example, where the nRPM was more efficient even though the deviations from linearity in the variables' relationships were small.

In simulation study 1, the RPM and nRPM were compared under ideal conditions, that is, under the no-effect-modifier assumption. One of the main results was that the nRPM led to more efficient estimates and to an increased power that was up to 52% higher than the power of the RPM. The advantages of specified the nRPM increased with the amount of nonlinearity. This was expected because in these situations, the linear function used in the RPM was misspecified. In a situation, where the RPM was correctly specified, the nRPM performed as well as the RPM. The more complex model did not lead to a substantial loss of power compared to the RPM even if the parametric RPM was the correctly specified model. Both the nRPM and the RPM kept the nominal type error rate, but the RPM provided a very low power as soon as the linear relationship assumption was violated.[5] In practical situations, other functions than the linear function can be chosen for the RPM. However, any choice could imply another misspecification, which is difficult to detect because the relationship is formulated for the unobserved potential values of the outcome variable. The nRPM allows for a data-driven approximation of the relationship, which reduces the misspecification without the need to formulate specific parametric models.

In simulation study 2, the possibility to detect interaction effects was investigated. Again, the nRPM provided more efficient estimates than the RPM. The interactions between treatment and covariates were not identified in the RPM using this specific linear model. In empirical situations, the inclusion of additional nonlinear terms (such as a moderated quadratic effect as suggested in Zheng and Zhou, 2015) will allow identification, though, the decision on which nonlinear term is included might be challenging. The nRPM can overcome this specific problem by estimating the nonlinear interactions with a semi-parametric function, which can reduce the multicollinearity in the weight matrix. Here, this was sufficient to identify the model even when the relationships in the data were generated as linear. The power to detect the other interaction effects using the RPM were close to the Type I error rate for small to medium sample sizes (200 or 500) for most

[5]Results not presented here showed that the original sandwich estimator SE were underestimated and led to inflated Type I error rates for the nRPM under some conditions (strong nonlinearity).

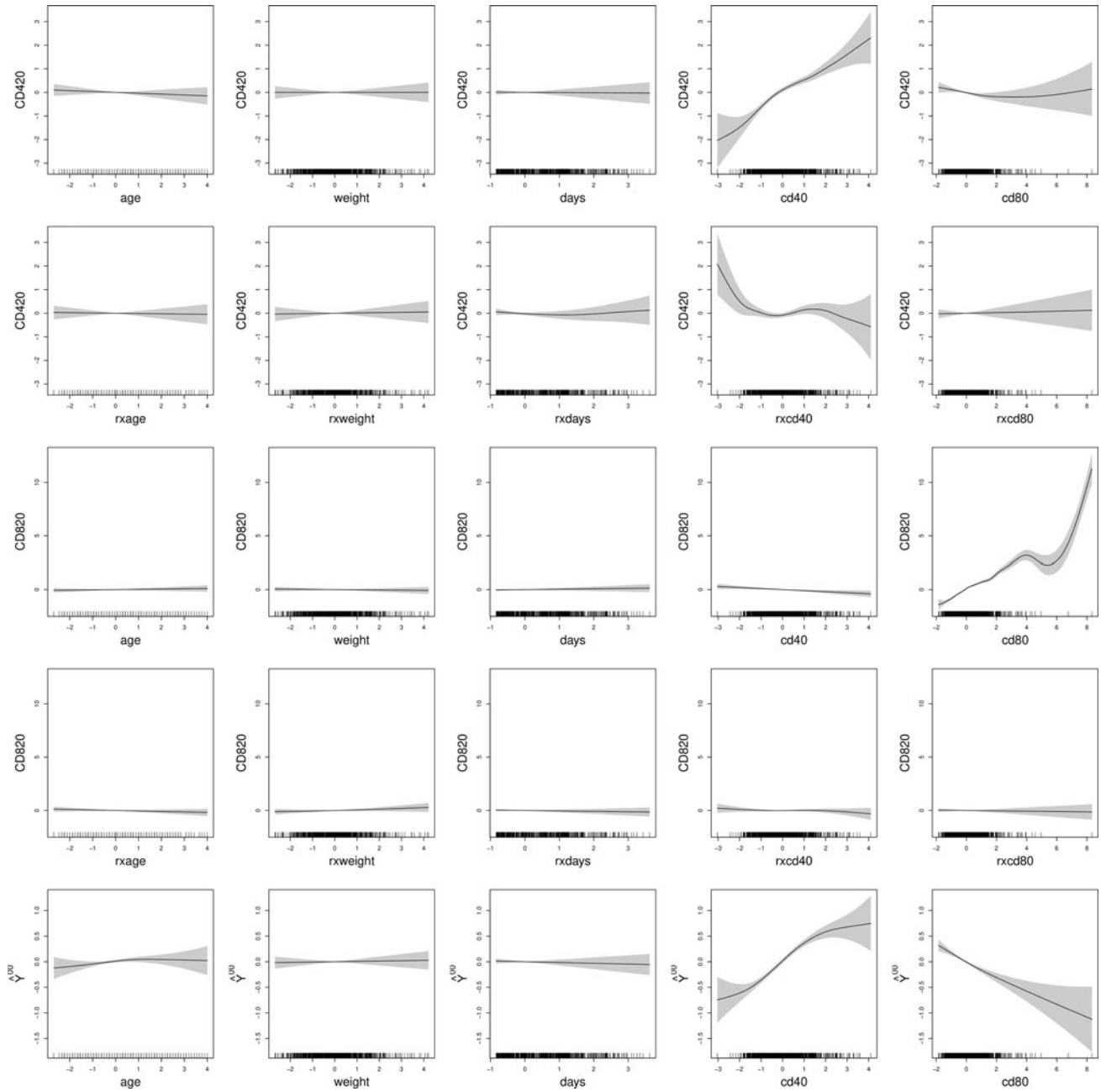

Figure 2. Illustration of the estimated (nonlinear) covariate functions for the intermediate variables (CD4 and CD8 at week 20) and potential values of the outcome variable $\hat{Y}^{00}$ in model 2 (for CD4 cell counts at week 20).

conditions. This implies that even though interaction effects are identified in the RPM, its practical application is limited to large sample sizes or effect sizes that are rare in empirical settings. The nRPM has a substantively higher power under all conditions, which makes it useful for practical applications.

In simulation study 3, the no-effect-modifier assumption was violated in order to investigate the robustness of the RPM and the nRPM. Neither the no-effect-modifier nor the no-unmeasured-confounder assumption can be tested, which makes a robustness against its violation essential. Both the RPM and the nRPM kept the nominal Type I error rate under most conditions. These conditions were chosen to reflect empirically relevant scenarios. More extreme conditions would necessarily lead to a stronger bias, but these scenarios might not be informative for the applicability of the model as they do not represent empirically relevant settings (Fleishman, 1978). Both the BKM and nBKM showed that even a minor violation of the no-unmeasured-confounder assumption ($\rho_{\epsilon_y \epsilon_m} = 0.1$) may lead to Type I error rates up to 100%, which will result in statistical artifacts in most situations.

Limitations and future directions

As any simulation study, it is necessary to limit the design and to include reasonable conditions. Here, the design focused on situations that can be expected to reflect empirical data. More extreme conditions, for example, with larger interaction effects will lead to a better performance of the RPM but might only be of theoretical value. As another example, the violation of the assumptions was kept at a lower boundary because in many situations, researchers will know at least some relevant covariates. Results concerning the relative performance may be even more leaning toward the nRPM compared to models using the no-unmeasured-confounder assumption if models are kept minimal (i.e., with only few or maybe even no covariates).

Zheng and Zhou (2015) proofed under very general conditions that the model provides consistent estimates. However, their findings with regard to convergence rate and asymptotic normality depend on a parametric working model G. The smoothing functions of the nRPM imply a data-driven selection of smoothing parameters and EDFs, which might impact these latter properties, for example, by resulting in slower convergence rates than the RPM. Future work is needed to extend the results from Zheng and Zhou (2015) to the nRPM. Though, it should be noted that the simulation results suggest that the model already performs well for sample sizes of 200 and that the bootstrap standard errors provide reliable estimates for the sampling variation.

In a different line of research, alternative concepts to causal mediators have been developed in the surrogate framework using principal stratification (Frangakis & Rubin, 2002; Gallop et al., 2009; Gilbert & Hudgens, 2008) or meta-analytic approaches (Joffe & Greene, 2009). These models do not assume that the intermediate variable (necessarily) is part of the causal chain, which might result in weaker assumptions to identify the model. The nRPM was not compared to these models here because they are conceptually very different (e.g., no CME is defined). Future research, however, could evaluate if clinically relevant information on intermediate variables can similarly be obtained in both frameworks.

The no-effect-modifier assumption cannot be tested, so two important implications need to be considered: First, one can test in general how robust the nRPM is against a violation of its assumptions using Monte Carlo simulation studies; here, encouraging results were found under realistic amounts of violation and data conditions that are relevant for empirical situations. Second, sensitivity analyses could be conducted to provide more information in single empirical data sets. These analyses provide more information about how reliable the model will work under the specific conditions a researcher encountered (and thus, results are more specific for the characteristics of a given data set than those from a Monte Carlo study). Zheng and Zhou (2015) developed formulas that can be used for single data sets for the RPM; however, they also stated that in many scenarios the resulting confidence bands were very conservative and might only be of limited practical relevance. The reason for this is that the whole multivariate information of mediator M, treatment R, covariates $\mathbf{X}$, and the confounder U needs to be considered (in contrast to a single parameter in the sensitivity analysis for the no-unmeasured-confounder assumption, cf. Imai et al., 2010). More work is needed to extend the sensitivity analysis for the nRPM.

Finally, an investigation of intermediate variables could be conducted on a latent variable level, which can attenuate problems of unreliable measures. While in medical research (e.g., CD 4 cell counts), measurement error might not be too problematic, social, and behavioral sciences often encounter more noise in the data. Here, an extension of the nRPM to semi-parametric structural equation models (e.g., Kelava & Brandt, 2014, 2019; Song, Lu, Cai, & Ip, 2013) might enhance its applicability.

Article information

Conflict of Interest Disclosures: The author signed a form for disclosure of potential conflicts of interest. The author did not report any financial or other conflicts of interest in relation to the work described.

Ethical Principles: The author affirms having followed professional ethical guidelines in preparing this work. These guidelines include obtaining informed consent from human participants, maintaining ethical treatment and respect for the rights of human or animal participants, and ensuring the privacy of participants and their data, such as ensuring that individual participants cannot be identified in reported results or from publicly available original or archival data.

Funding: This work was not supported.

Role of the Funders/Sponsors: None of the funders or sponsors of this research had any role in the design and conduct of the study; collection, management, analysis, and interpretation of data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Acknowledgments: The ideas and opinions expressed herein are those of the author alone, and endorsement by the author's institution is not intended and should not be inferred.

References

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. doi:10.1037//0022-3514.51.6.1173

Brumback, B. A., Hernán, M. A., Haneuse, S. J. P. A., & Robins, J. M. (2004). Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Statistics in Medicine*, *23*(5), 749–767. doi:10.1002/sim.1657

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology*, *98*(4), 550–558. doi:10.1037/a0018933

Burzykowski, T., Molenberghs, G., & Buyse, M. (2006). *The evaluation of surrogate endpoints*. New York, NY: Springer.

Chaplin, W. (1991). The next generation of moderator research in personality psychology. *Journal of Personality*, *59*(2), 143–178. doi:10.1111/j.1467-6494.1991.tb00772.x

Chaplin, W. (2007). Moderator and mediator models in personality research: A basic introduction. In R. W. Robins, C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 602–632). New York, NY: Guilford.

Ding, P., & Van der Weele, T. J. (2016). Sensitivity analysis without assumptions. *Epidemiology*, *27*(3), 277–368. doi:10.1097/EDE.0000000000000457

Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In W. Schempp & K. Zeller (Eds.), *Constructive theory of functions of several variables. Lecture notes in mathematics* (Vol. 571, pp. 85–100). Berlin: Springer.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, *7*, 1–26. doi:10.1214/aos/1176344552

Fischer-Lapp, K., & Goetghebeur, E. (1999). Practical properties of some structural mean analyses of the effect of compliance in randomized trials. *Controlled Clinical Trials*, *20*(6), 531–546. doi:10.1016/S0197-2456(99)00027-6

Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, *43*(4), 521–532. doi:10.1007/BF02293811

Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, *58*(1), 21–29. doi:10.1111/j.0006-341X.2002.00021.x

Gallop, R., Small, D. S., Lin, J. Y., Elliott, M. R., Joffe, M., & Ten Have, T. R. (2009). Mediation analysis with principal stratification. *Statistics in Medicine*, *28*(7), 1108–1130. doi:10.1002/sim.3533

Gilbert, P. B., & Hudgens, M. G. (2008). Evaluating candidate principal surrogate endpoints. *Biometrics*, *64*(4), 1146–1154. doi:10.1111/j.1541-0420.2008.01014.x

Hammer, S. M., Katzenstein, D. A., Hugher, M. D., Gundacker, H., Schooley, R. T., & Haubrich, R. H. (1996). A trial comparing nucleoside monotherapy with combination therapy in HIV-infected adults with cd4 cell counts from 200 to 500 per cubic millimeter. *The New England Journal of Medicine*, *336*, 1081–1090. doi:10.1056/NEJM199610103351501

Hastie, T. (1996). Pseudosplines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *58*, 379–396. doi:10.1111/j.2517-6161.1996.tb02088.x

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). New York, NY: Springer.

Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. *Sociological Methodology*, *18*, 449–484. doi:10.2307/271055

Hong, G., Qin, X., & Yang, F. (2018). Weighting-based sensitivity analysis in causal mediation studies. *Journal of Educational and Behavioral Statistics*, *43*(1), 32–56. doi:10.3102/1076998617749561

Huber, P. J., & Ronchetti, E. M. (2009). *Robust statistics* (2nd ed.). New York, NY: Wiley.

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–334. doi:10.1037/a0020761

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference, and sensitivity analysis for causal mediation effects. *Statistical Science*, *25*(1), 51–71. doi:10.1214/10-STS321

Joffe, M. M., & Greene, T. (2009). Related causal frameworks for surrogate outcomes. *Biometrics*, *65*(2), 530–538. doi:10.1111/j.1541-0420.2008.01106.x

Juraska, M., Gilbert, P. B., Lu, X., Zhang, M., Davidian, M., & Tsiatis, A. A. (2012). speff2trial: Semiparametric efficient estimation for a two-sample treatment effect. R package version 1.0.4 [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=speff2trial

Kaufman, S., Kaufman, J. S., MacLehose, R. F., Greenland, S., & Poole, C. (2005). Improved estimation of controlled direct effects in the presence of unmeasured confounding of intermediate variables. *Statistics in Medicine*, *24*(11), 1683–1702. doi:10.1002/sim.2057

Kelava, A., & Brandt, H. (2019). A Nonlinear Dynamic Latent Class Structural Equation Model. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(4), 509–528. doi:10.1080/10705511.2018.1555692

Kelava, A., & Brandt, H. (2014). A general nonlinear multilevel structural equation mixture model. *Frontiers in Quantitative Psychology and Measurement*, *5*(748). doi:10.3389/fpsyg.2014.00748

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*(1), 593–614. doi:10.1146/annurev.psych.58.110405.085542

Nock, M. K. (2007). Conceptual and design essentials for evaluating mechanisms of clinical change. *Alcoholism: Clinical and Experimental Research*, *31*(s3), 4S–12S. doi:10.1111/j.1530-0277.2007.00488.x

Pearl, J. (2012). The causal mediation formula: A guide to the assessment of pathways and mechanisms. *Prevention Science*, *13*(4), 426–436. doi:10.1007/s11121-011-0270-1

Pearl, J. (2014). Identification and interpretation of causal mediation. *Psychological Methods*, *19*(4), 459–481. doi:10.1037/a0036434

Petersen, M., Sinisi, S., & van der Laan, M. (2006). Estimation of direct causal effects. *Epidemiology*, *17*(3), 276–284. doi:10.1097/01.ede.0000208475.99429.2d

Preacher, K. J. (2015). Advances in mediation analysis: A survey and synthesis of new developments. *Annual Review of Psychology*, *66*(1), 825–852. doi:10.1146/annurev-psych-010814-015258

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, *42*(1), 185–227. doi:10.1080/00273170701341316

Prentice, R. L. (1989). Surrogate endpoints in clinical trials: Definition and operational criteria. *Statistics in Medicine*, *8*(4), 431–440. doi:10.1002/sim.4780080407

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. R Core Team: Vienna. Retrieved from https://www.R-project.org

Reiss, P. T., & Ogden, R. T. (2009). Smoothing parameter selection for a class of semiparametric linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology))*, *71*(2), 505–523. doi:10.1111/j.1467-9868.2008.00695.x

Robins, J. M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics - Theory and Methods*, *23*(8), 2379–2412. doi:10.1080/03610929408831393

Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, *3*(2), 143–155. doi:10.1097/00001648-199203000-00013

Robins, J. M., Hernán, M. A., & Brumback, B. A. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, *11*(5), 550–560. doi:10.1097/00001648-200009000-00011

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. doi:10.1093/biomet/70.1.41

Scheipl, F., Greven, S., & Kuchenhoff, H. (2008). Size and power of tests for a zero random effect variance or polynomial regression in additive and linear mixed models. *Computational Statistics & Data Analysis*, *52*, 3283–3299. doi:10.1016/j.csda.2007.10.022

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, *7*(4), 422–445. doi:10.1037//1082-989X.7.4.422

Small, D. S. (2012). Mediation analysis without sequential ignorability: Using baseline covariates interacted with random assignment as instrumental variables. *Journal of Statistical Research*, *46*, 91–103.

Small, D. S., Joffe, M. M., Lynch, K. G., Roy, J. A., & Locali, A. R. (2014). Tom Ten Have's contributions to causal inference and biostatistics: Review and future research directions. *Statistics in Medicine*, *33*(20), 3421–3433. doi:10.1002/sim.5708

Soetaert, K. (2009). rootSolve: Nonlinear root finding, equilibrium and steady-state analysis of ordinary differential equations [Computer software manual]. R package 1.6.

Song, X.-Y., Lu, Z.-H., Cai, J.-H., & Ip, E. H.-S. (2013). A Bayesian approach for generalized semiparametric structural equation models. *Psychometrika*, *78*(4), 624–647. doi:10.1007/s11336-013-9323-7

Ten Have, T. R., Joffe, M. M., Lynch, K. G., Brown, G. K., Maisto, S. A., & Beck, A. T. (2007). Causal mediation analyses with rank preserving models. *Biometrics*, *63*(3), 926–934. doi:10.1111/j.1541-0420.2007.00766.x

Valeri, L., & Van der Weele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and 7implementation with sas and spss macros. *Psychological Methods*, *18*(2), 137–150. doi:10.4310/10.1037/a0031034

Van der Weele, T. J. (2009). Marginal structural models for the estimation of direct and indirect effects. *Epidemiology*, *20*(1), 18–26. doi:10.1097/EDE.0b013e31818f69ce

Van der Weele, T. J. (2010). Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology*, *21*, 540–551. doi:10.1097/EDE.0b013e3181df191c

Van der Weele, T. J. (2016). Mediation analysis: A practitioner's guide. *Annual Review of Public Health*, *37*(1), 17–32. doi:10.1146/annurev-publhealth-032315-021402

Van der Weele, T. J., & Shpitser, I. (2011). A new criterion for confounder selection. *Biometrics*, *67*(4), 1406–1413. doi:10.1111/j.1541-0420.2011.01619.x

Van der Weele, T. J., & Shpitser, I. (2013). On the definition of a confounder. *The Annals of Statistics*, *41*(1), 196–220. doi:10.1214/12-AOS1058

Van der Weele, T. J., & Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and Its Interface*, *2*(4), 457–468. doi:10.4310/SII.2009.v2.n4.a7

Vansteelandt, S., & Van der Weele, T. J. (2012). Natural direct and indirect effects on the exposed: Effect decomposition under weaker assumptions. *Biometrics*, *68*(4), 1019–1027. doi:10.1111/j.1541-0420.2012.01777.x

Wahba, G. (1990). *Spline models for observational data*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Wood, S. N. (2003). Thin-plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(1), 95–114. doi:10.1111/1467-9868.00374

Wood, S. N. (2006). Low-rank scale-invariant tensor product smooths for generalized additive mixed models. *Biometrics*, *62*(4), 1025–1036. doi:10.1111/j.1541-0420.2006.00574.x

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology))*, *73*(1), 3–36. doi:10.2307/41057423

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Boca Raton, FL: Chapman & Hall, CRC.

Zheng, C., Atkins, D. C., Zhou, X.-H., & Rhew, I. C. (2015). Causal models for mediation analysis: An introduction to structural mean models. *Multivariate Behavioral Research*, *50*(6), 614–631. doi:10.1080/00273171.2015.1070707

Zheng, C., & Zhou, X.-H. (2015). Causal mediation analysis in the multilevel intervention and multicomponent mediator case. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *77*(3), 581–615. doi:10.1111/rssb.12082

Appendix A

Additional tables for the simulation study.

Table A1. Study 1: relative bias (in percent) of the CME θ_M for RPM and nRPM.

		$N = 200$			$N = 500$			$N = 1{,}000$		
θ_M	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
					No nonlinearity					
0.0	RPM	5.76	3.64	1.50	2.11	2.13	−0.24	3.02	−1.31	1.39
0.0	nRPM	6.09	2.24	2.95	3.37	2.68	0.53	3.20	−1.63	1.03
0.1	RPM	7.94	3.89	2.23	6.44	1.55	1.72	3.40	−0.25	0.21
0.1	nRPM	8.45	5.61	0.89	6.09	1.43	1.59	3.01	0.40	0.23
0.2	RPM	6.34	1.47	1.22	4.70	1.55	−0.04	4.40	−0.63	0.82
0.2	nRPM	10.21	0.79	2.37	4.39	1.79	0.23	3.55	−0.23	0.63
0.3	RPM	4.00	−4.59	0.66	1.47	−0.27	−0.19	−0.65	−0.63	0.90
0.3	nRPM	7.55	−0.89	0.73	2.72	0.87	0.38	0.94	0.12	1.28
					Medium nonlinearity					
0.0	RPM	9.95	2.66	−2.92	4.16	3.71	0.66	0.53	0.81	0.30
0.0	nRPM	8.75	4.33	2.75	7.06	3.11	1.64	2.83	1.42	1.74
0.1	RPM	9.42	2.82	1.30	5.84	3.68	1.10	2.43	0.82	−0.90
0.1	nRPM	11.74	8.53	4.55	6.82	3.28	1.62	5.35	1.58	1.77
0.2	RPM	7.55	2.03	2.13	3.64	−0.19	0.56	2.63	1.85	0.76
0.2	nRPM	8.33	4.24	3.12	6.72	2.22	0.07	4.37	2.42	1.37
0.3	RPM	10.24	2.21	4.01	7.22	4.12	1.33	6.07	0.70	0.39
0.3	nRPM	12.14	6.98	3.20	8.78	4.80	2.52	5.76	2.74	2.48
					Strong nonlinearity					
0.0	RPM	6.41	5.29	5.08	6.61	3.19	2.38	13.74	1.03	3.41
0.0	nRPM	10.37	9.57	5.26	8.59	6.06	3.21	8.12	4.94	2.93
0.1	RPM	9.91	2.54	3.02	8.99	3.81	3.14	6.50	1.18	5.15
0.1	nRPM	14.97	5.94	9.21	9.83	5.09	2.64	6.51	4.76	2.64
0.2	RPM	4.94	4.17	4.47	9.04	8.38	0.76	5.02	1.82	2.60
0.2	nRPM	10.12	8.60	6.60	4.98	5.45	3.88	6.99	5.05	3.87
0.3	RPM	5.29	8.90	7.75	6.52	2.27	3.45	6.23	2.69	3.50
0.3	nRPM	12.54	8.53	6.02	6.75	5.32	2.98	9.98	2.60	2.05

N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$, RPM: rank preserving model, nRPM: nonlinear rank preserving model.

Table A2. Study 1: relative RMSE of the CME θ_M for RPM vs. nRPM.

		$N = 200$			$N = 500$			$N = 1{,}000$		
θ_M	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
					No nonlinearity					
0.0		1.08	1.04	0.94	1.01	0.99	1.06	1.07	1.05	1.00
0.1		1.13	1.11	0.88	1.01	1.00	0.94	1.09	1.00	0.97
0.2		1.15	1.02	1.02	1.18	1.00	1.12	1.10	0.99	1.03
0.3		1.05	1.03	0.93	1.02	0.96	0.95	1.10	1.09	1.12
					Medium nonlinearity					
0.0		1.28	1.38	1.29	1.40	1.40	1.36	1.49	1.36	1.39
0.1		1.34	1.17	1.13	1.40	1.42	1.37	1.32	1.36	1.39
0.2		1.48	1.34	1.14	1.41	1.39	1.26	1.47	1.34	1.54
0.3		1.51	1.38	1.15	1.53	1.36	1.36	1.54	1.44	1.29
					Strong nonlinearity					
0.0		1.47	1.63	1.88	1.78	2.44	2.00	1.96	2.46	2.41
0.1		1.58	1.64	1.86	1.85	1.99	2.35	2.47	2.40	2.33
0.2		1.65	1.60	1.74	2.19	2.06	2.75	2.20	2.70	2.29
0.3		2.01	1.95	1.79	1.97	2.36	2.33	2.13	2.95	2.70

Values above 1 indicate smaller RMSE for the nRPM.
N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$.

Table A3. Study 3: percent significant estimates for the CME $\theta_M = 0$ (type I error rate) for misspecified models with $\gamma_{UM} = 0.2$ (strong violation of no-effect modifier assumption).

		$N = 200$			$N = 500$			$N = 1{,}000$		
Nonlinearity	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
					$\gamma_{U0} = 0.0$					
None	RPM	0.0	0.0	0.6	0.0	1.2	2.2	0.0	2.0	1.6
	nRPM	0.0	0.2	0.2	0.2	0.8	3.2	0.2	1.6	2.4
	BKM	3.4	4.8	5.4	6.4	8.4	3.4	5.6	6.4	5.0
	nBKM	3.6	4.6	5.4	6.4	8.2	3.6	5.4	6.4	5.0
Medium	RPM	0.4	0.6	0.0	0.2	0.6	0.8	0.0	1.4	3.4
	nRPM	0.0	0.2	0.6	0.8	0.6	1.6	0.2	2.8	4.0
	BKM	7.8	7.2	8.6	12.6	13.8	10.6	18.8	17.2	20.4
	nBKM	6.2	6.2	5.2	6.4	7.0	6.0	7.8	7.2	9.2
Strong	RPM	0.4	0.0	0.2	0.0	0.2	1.0	0.4	0.6	1.0
	nRPM	0.4	0.2	0.6	0.8	1.2	2.0	1.0	1.4	6.2
	BKM	16.4	18.8	20.0	32.2	32.2	32.2	47.4	44.6	43.4
	nBKM	7.4	7.8	8.8	10.0	9.2	8.8	9.0	8.8	11.4
					$\gamma_{U0} = 0.3$					
None	RPM	0.0	0.2	1.0	0.2	1.2	2.8	0.4	1.4	4.2
	nRPM	0.6	0.2	0.2	0.2	0.8	2.2	0.6	2.0	4.2
	BKM	39.6	43.6	44.4	78.6	79.0	78.6	96.8	98.0	98.0
	nBKM	40.2	43.4	45.0	78.6	79.0	78.6	96.8	98.0	98.0
Medium	RPM	0.0	0.2	0.6	0.0	0.6	1.6	0.2	1.4	2.6
	nRPM	0.4	0.4	1.0	0.6	0.8	3.6	0.6	2.6	3.0
	BKM	41.4	40.8	43.8	70.2	72.2	73.6	89.0	86.8	88.6
	nBKM	45.8	43.4	45.6	78.6	77.8	79.4	97.2	96.6	96.4
Strong	RPM	0.2	0.2	0.8	0.4	0.4	0.0	0.2	1.0	0.0
	nRPM	0.4	0.8	0.4	1.0	1.6	3.8	0.8	3.6	3.6
	BKM	39.6	40.0	39.2	56.8	61.2	61.6	72.6	73.0	76.8
	nBKM	41.0	44.6	39.2	72.8	75.8	75.2	94.4	95.0	96.6

N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$, γ_{UR}: interaction effect of the confounder $U \times R \rightarrow Y$, γ_{U0}: linear effect $U \rightarrow M, Y$, RPM: rank preserving model, nRPM: nonlinear rank preserving model, BKM: Baron-Kenny model, nBKM: nonlinear Baron-Kenny model.

Table A4. Study 3: relative accuracy for the CME $\theta_M = 0$ compared to the nRPM for misspecified models with $\gamma_{UM} = 0.2$.

		$N = 200$			$N = 500$			$N = 1{,}000$		
Nonlinearity	$\beta_{RX} =$	0.05	0.10	0.15	0.05	0.10	0.15	0.05	0.10	0.15
					$\gamma_{U0} = 0.0$					
None	RPM	1.28	0.98	1.01	1.04	1.05	0.95	1.14	1.01	1.04
	BKM	0.13	0.16	0.24	0.11	0.19	0.23	0.10	0.16	0.27
	nBKM	0.13	0.16	0.24	0.11	0.18	0.23	0.10	0.17	0.27
Medium	RPM	1.33	1.35	1.19	1.62	1.46	1.29	1.39	1.49	1.31
	BKM	0.13	0.20	0.29	0.15	0.21	0.29	0.12	0.24	0.34
	nBKM	0.13	0.19	0.25	0.13	0.18	0.25	0.10	0.18	0.25
Strong	RPM	1.71	1.52	1.95	2.25	2.08	2.36	2.14	2.39	2.27
	BKM	0.22	0.24	0.35	0.25	0.35	0.38	0.25	0.38	0.50
	nBKM	0.16	0.17	0.25	0.13	0.20	0.26	0.11	0.18	0.26
					$\gamma_{U0} = 0.3$					
None	RPM	1.12	1.01	0.89	1.10	0.95	0.95	1.04	1.03	1.06
	BKM	0.22	0.38	0.50	0.27	0.51	0.72	0.38	0.76	1.15
	nBKM	0.22	0.38	0.50	0.27	0.50	0.72	0.38	0.76	1.15
Medium	RPM	1.22	1.43	1.32	1.46	1.29	1.14	1.39	1.25	1.31
	BKM	0.25	0.37	0.55	0.34	0.50	0.64	0.39	0.65	0.99
	nBKM	0.26	0.38	0.54	0.34	0.50	0.64	0.39	0.65	1.00
Strong	RPM	1.65	1.88	2.06	1.87	2.15	2.29	2.02	2.36	2.20
	BKM	0.30	0.38	0.48	0.31	0.48	0.65	0.40	0.66	0.95
	nBKM	0.28	0.35	0.44	0.33	0.47	0.62	0.36	0.61	0.87

Values for the relative accuracy above 1 indicate a higher accuracy for the nRPM.

N: sample size, β_{RX}: interaction effect $R \times X \rightarrow M$, γ_{UR}: interaction effect of the confounder $U \times R \rightarrow Y$, γ_{U0}: linear effect $U \rightarrow M, Y$, RPM: rank preserving model, nRPM: nonlinear rank preserving model, BKM: Baron-Kenny model, nBKM: nonlinear Baron-Kenny model.

Appendix B

Additional results for the empirical example.

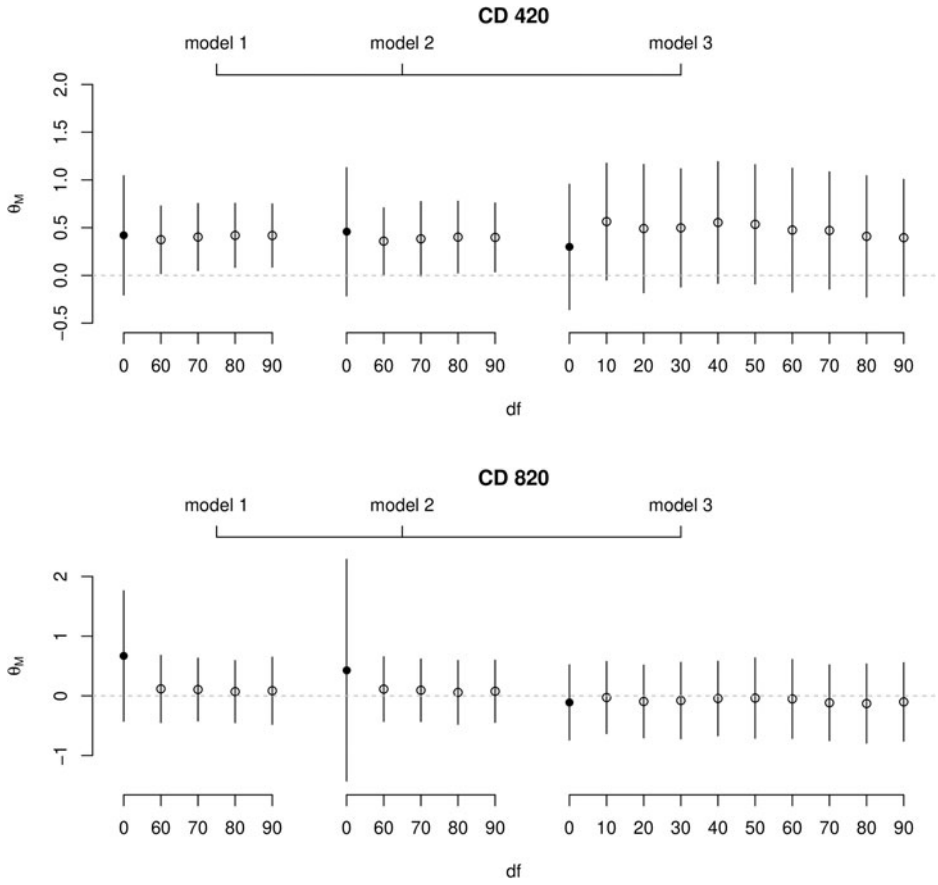

Figure B1. The CME and its 95% confidence intervals (CIs) across the different models 1 through 3 and different numbers of nodes. The RPM is indicated with a filled dot and nRPM with empty dots. Models 1 and 2 needed a minimum dimension of 60 nodes. As can be seen, parameter estimates and CI's are virtually identical across different specifications of the nRPM. The RPM provides larger CI's especially for models 1 and 2 because the two additional covariates in these models had nonlinear relationships with M and Y^{00} that increased the nRPM's relative efficiency.