Published in final edited form as:
Nurs Res. 2012 ; 61(3): 224–230. doi:10.1097/NNR.0b013e318254165c.

Causal Mediation of a Human Immunodeficiency Virus Preventive Intervention

Donna L. Coffman, Ph.D. [Research Assistant Professor] and
The Methodology Center & The College of Health and Human Development, The Pennsylvania State University

Kari C. Kugler, Ph.D. [Research Associate]
The Methodology Center, The Pennsylvania State University

Abstract

Background—Assessing mediation is important because most interventions are specifically designed to affect an intermediate variable or mediator; this mediator, in turn, is hypothesized to affect outcome behaviors. Although there may be randomization to the intervention, randomization to levels of the mediator is not generally possible. Therefore, drawing causal inferences about the effect of the mediator on the outcome is not straightforward.

Objectives—We introduce an approach to causal mediation analysis that uses the potential outcomes framework for causal inference, and then discuss this approach in terms of the scientific questions addressed and the assumptions needed for identifying and estimating the effects.

Method—We illustrate the approach using data from the Criminal Justice Drug Abuse Treatment studies: Reducing Risky Relationships HIV intervention (RRR-HIV) implemented with 243 incarcerated women reentering the community. The intervention was designed to affect various mediators at 30 days post-intervention including risky relationship thoughts, beliefs, and attitudes, which were then hypothesized to affect engagement in risky sexual behaviors such as unprotected sex at 90 days post-intervention.

Results—Using propensity score weighting, we found the intervention resulted in a significant decrease in risky relationship thoughts (−0.529, $p = .03$); risky relationship thoughts resulted in an increase in the odds of unprotected sex (.447, $p < .001$). However, the direct effect of the intervention on unprotected sex was not significant (0.388, $p = .479$).

Discussion—By reducing bias, propensity score models improve the accuracy of statistical analysis of interventions with mediators and allow researchers to determine not only *if* their intervention works, but also *how* it works.

Keywords

causal inference; potential outcomes framework; incarcerated women

Mediators, also known as mediating variables, play a significant role in most models of behavioral intervention: many programs target change in mediators in order to promote change

Address correspondence to, Donna L. Coffman, The Methodology Center, The Pennsylvania State University, 204 E. Calder Way, Ste. 400, State College, PA 16801, dlc30@psu.edu.

on the ultimate outcome of interest. Mediation occurs as part of a hypothesized causal chain of events: the independent variable (e.g., intervention) has an effect on the mediator (e.g., risky relationship thoughts), which then affects the outcome (e.g., sex without condom). Identification of mediators can reveal important pathways that lead to behavior change (MacKinnon, 2008).

An important step in making interventions more efficacious and cost-effective is to identify the mechanism(s) by which the intervention has its effect on the outcome. That is, researchers not only need to know that the intervention works but also need to know *how* it works. Thus, establishing the causal link between the mediator and the outcome is critical. The commonly used Baron and Kenny (1986) approach to mediation and its extensions (MacKinnon, 2008; see also, e.g., Krause et al., 2010 and Levy, Landerman, & Davis, 2011 for introductions to these extensions for nursing researchers) rely on an assumption of no unmeasured confounding, which holds if individuals are randomly assigned to *both* the intervention and the mediator. However, as noted in the statistics and methodology literature (e.g., Robins & Greenland, 1992; Rosenbaum, 1984), establishing this causal link can be difficult because randomization of individuals to levels of the mediator is usually impossible. Therefore, confounders may exist that influence both the mediator and the outcome and bias the estimate of the effect of the mediator on the outcome. Recently, the potential outcomes framework (Rubin, 1974) for causal inference has been proposed for assessing mediation (e.g., Coffman, 2011; Imai, Keele, & Tingley, 2010; Jo, 2008; VanderWeele, 2009). This framework clearly defines the causal effects to be estimated and states the assumptions needed to do so.

The goal of the present article is to introduce an approach to mediation that falls under the potential outcomes framework for causal inference. We will discuss the scientific question(s) that may be addressed and the assumptions required for identification and estimation of causal mediation effects. The article is organized as follows. First, we introduce a motivating example, the Reducing Risky Relationships Human Immunodeficiency Virus (RRR-HIV) intervention. Next, we introduce the potential outcomes framework, describe one way to define mediation using this framework, and discuss how to identify and estimate effects using this approach. Finally, we use this approach to assess mediation in our motivating example. Much has been written in the statistics and epidemiology literature about the various approaches to assessing mediation, but there are very few published applications of these approaches and, to our knowledge, no published applications in the nursing research literature.

Motivating Example: RRR-HIV Intervention

After three decades, the persistence of the HIV/acquired immune deficiency syndrome (AIDS) epidemic in the United States requires sustained attention. Distinct subgroups are disproportionately infected with HIV, such as individuals who enter the criminal justice system, especially females, with three to four times greater infection rates than the general population (Maruschak, 2009).

The risk factors underlying the disproportionate HIV rates among incarcerated women include exchange of sex for drugs or money, having a high-risk sexual partner, inconsistent condom use, and use of drugs or other substances (Cotton-Oldenburg, Jordan, Martin, & Kupper, 1999); these often occur at greater frequency than in the general population. The underlying causes for greater frequency of these risk factors is not clear, but it is important to intervene with this population because many incarcerated women return to the same behaviors, relationships, and financial context upon release (Adams et al., 2011). To date, very few interventions have focused on this population (Lichtenstein & Malow, 2010).

To fill in the gap in the literature, the RRR-HIV intervention, as part of the Criminal Justice Drug Abuse Treatment Studies (CJ-DATS) cooperative agreement (Fletcher & Wexler, 2005), was developed, piloted, and tested among incarcerated women (Havens et al., 2009; Staton-Tindall et al., 2007). The intervention was designed to address the context of a relationship and change risky relationship thoughts with the aim of decreasing risky sex after release (Staton-Tindall et al., 2007). Thus, risky relationship thoughts mediate the effect of the intervention on risky sexual behavior. That is, the intervention was hypothesized to change risky relationship thoughts, which in turn were hypothesized to change risky sexual behaviors. These effects are denoted β_1 and β_2, respectively, in Figure 1. Figure 1 also illustrates the direct effect, β_3; that is, the effect of the intervention on risky sexual behavior that is not due to risky relationship thoughts. Next, we introduce the potential outcomes framework and how it can be used to assess mediation.

Potential Outcomes Framework for Causal Inference

In the potential outcomes framework (Rubin, 1974), each individual has a potential outcome for each possible treatment condition. For simplicity, consider a binary treatment indicator, T_i, where T_i=1 denotes the intervention condition and T_i=0 denotes the control condition for participant i, i=1,…,n. In this case, there are two potential outcomes for each individual: the potential outcome if the individual receives the intervention, denoted $Y_i(1)$, and the potential outcome if the individual is in the control condition, denoted $Y_i(0)$. The individual causal effect is the difference between these two potential outcomes. Because each participant is observed in only one of these conditions, only one potential outcome is observed; the other potential outcome is missing and, therefore, the individual causal effect cannot be computed. However, strategies have been implemented to estimate the causal effect averaged over participants in the study, called the average causal effect (ACE). The ACE is defined as $E[Y_i(1)-Y_i(0)]$; that is, the expected (or average) difference between the two potential outcomes. Introductions to the potential outcomes framework outside of the mediation context are provided by Little and Rubin (2000) and Schafer and Kang (2008).

A mediator is an outcome assessed within the context of the intervention and, therefore, there are also potential values for the mediator under each intervention condition for each individual: the potential value of the mediator under the intervention condition, denoted $M_i(1)$, and the potential value of the mediator under the control condition, denoted $M_i(0)$. The potential values for the outcome are then expanded to include potential values of the mediator. Thus, $Y_i(1,M_i(1))$ is the potential outcome if individual i receives the intervention, and $Y_i(0,M_i(0))$ is the potential outcome if individual i is in the control condition. As before, only one of these two potential outcomes is observed for each individual.

Throughout this article, we use Y_i to denote the observed value of unprotected sex (operationalized in this study as having had at least one sexual encounter without a condom in the previous 30 days; described in more detail in the Measures section below), M_i to denote the observed value for risky relationship thoughts, and $Y_i(t_i,M_i(t_i))$ to denote the potential outcomes where t_i is one of the levels of the intervention. We use X_i to denote measured confounders. We assume throughout that if an individual receives the intervention, then $Y_i=Y_i(1)=Y_i(1,M_i(1))$ and $M_i=M_i(1)$. Likewise, if an individual is in the control condition, then $Y_i=Y_i(0)=Y_i(0,M_i(0))$ and $M_i=M_i(0)$. This is usually referred to as the consistency assumption. Also implicit in this notation is that there is no interference among individuals because the potential outcomes are a function of only T_i and not T_j where i and j denote two different individuals. In other words, an individual's outcome does not depend on another individual's treatment assignment. Throughout this article, we make this no-interference assumption.

Using the Potential Outcomes Framework to Define Mediation Effects

Three different definitions of mediation using the potential outcomes framework have been introduced in the statistics literature: principal strata effects, natural effects, and controlled effects. We will focus on controlled effects, first introduced by Robins and Greenland (1992).

In our motivating example, the controlled direct effect is the causal effect of the RRR-HIV intervention on unprotected sex when risky relationship thoughts is set (i.e., held constant) to a specific value, m, for the entire population. That is, $E[Y_i(1,m)-Y_i(0,m)]$ where $Y_i(t,m)$ is the potential outcome when $T_i=t_i$ and $M_i(t_i)=m$. For the controlled direct effect, mediator is set to the same value for every individual. Also, for a binary intervention, such as RRR-HIV, there are as many controlled direct effects as there are possible values of the mediator. If the controlled direct effects are different across levels of the mediator, this implies that there is an interaction between the intervention and the mediator. If the controlled direct effects are equal across levels of the mediator, this implies that there is no interaction between the intervention and mediator.

Next, consider defining the effect, $E[Y_i(1,m)-Y_i(1,m')]$ for two different values of m and m'. This is the effect of, for example, a one-unit change in risky relationship thoughts on unprotected sex when $T_i=1$. Similarly, the difference, $E[Y_i(0,m)-Y_i(0,m')]$, defines the effect of a one-unit change in risky relationship thoughts on unprotected sex when $T_i=0$. These differences define the effect of risky relationship thoughts on unprotected sex at each level of the intervention. The effect of the intervention on risky relationship thoughts is defined as $E[M_i(1)-M_i(0)]$. Finally, if there is no interaction between the intervention and mediator, such that there is only one controlled direct effect, the controlled direct effect may be subtracted from the total effect to obtain the indirect effect. The total effect is defined as the effect of the intervention on unprotected sex, $E[Y_i(1)-Y_i(0)]$.

Identification and Estimation

Assumptions

Much less attention has been given to estimating mediation via controlled effects than via natural effects or principal strata effects. However, as discussed by VanderWeele (2009), if there is no interaction between the intervention and the mediator, then the controlled direct effect is the same for every level of the mediator. In this case, the controlled direct effect can be subtracted from the total effect to obtain the indirect effect. This approach requires Assumption (a): that there are no unmeasured confounders of the intervention and unprotected sex; Assumption (b): that there are no unmeasured confounders of the intervention and risky relationship thoughts; Assumption (c): that there are no unmeasured confounders of risky relationship thoughts and unprotected sex; and Assumption (d): that there are no interactions between the intervention and risky relationship thoughts. If individuals are randomized to levels of the intervention, as they are in the RRR-HIV study, then Assumptions (a) and (b) hold. However, this randomization *does not* imply that Assumption (c) holds. Randomization guarantees that the intervention groups are equivalent on pre-randomization variables, but it does not preclude the possibility of differences between the intervention groups on post-randomization variables. Further, randomization to levels of the intervention does not mean that individuals are randomized to levels of the mediator. Without randomization to levels of the mediator, there is no guarantee that there are no confounders of the mediator and outcome.

Estimation

Coffman and Zhong (2011) proposed to define and estimate all of the effects given above using marginal structural models (MSMs) with an inverse propensity weighted (IPW) estimator using identifying assumptions (a) – (c). MSMs are models for the *potential outcomes* and have been previously described, along with IPW estimation, in the prevention literature (e.g., Bray, Almirall, Zimmerman, Lynam, & Murphy, 2006; Coffman, Caldwell, & Smith, in press) and epidemiology literature (e.g., Cole et al., 2003; Robins, Hernan, & Brumback, 2000). For example, for continuous outcomes, the MSMs may be given as

$$E[M(t)]=\beta_{0M}+\beta_1 t \quad (1)$$

and

$$E[Y(t,m)]=\beta_{0Y}+\beta_2 m+\beta_3 t\,, \quad (2)$$

where $\beta_3 = E[Y(1,m)-Y(0,m)]=(\beta_{0Y}+\beta_2 m+\beta_3)-(\beta_{0Y}+\beta_2 m)$ is the controlled direct effect defined above, $\beta_1 = E[M(1)-M(0)]=(\beta_{0M}+\beta_1)-\beta_{0M}$ is the effect of the intervention on risky relationship thoughts, and $\beta_2 = E[Y(t,m)-Y(t,m')]$ is the effect of risky relationship thoughts on unprotected sex, holding constant intervention condition. MSMs are fit by choosing an appropriate model for the *observed outcome* (e.g., linear regression, logistic regression, survival model) and using the IPW estimator rather than the usual ordinary least squares (OLS) or maximum likelihood (ML) estimator. We describe how to obtain the weights for the IPW estimator in the Empirical Demonstration section. If individuals are randomized to levels of the intervention, then Assumption (a) holds and weights are unnecessary for estimating Equation (1). The models for the observed data, as opposed to the potential outcomes, are given as

$$E[M_i|T_i=t_i]=\beta_{0M}+\beta_1 t_i \quad (3)$$

and

$$E[Y_i|T_i=t_i, M_i=m]=\beta_{0Y}+\beta_2 m+\beta_3 t_i. \quad (4)$$

Coffman and Zhong (2011) also proposed a null hypothesis test of no mediation. Specifically, the null hypothesis is that either the effect of the intervention on risky relationship thoughts or the effect of risky relationship thoughts on unprotected sex, holding constant intervention condition, is zero. If Assumption (d) holds, then an estimate of the indirect effect may be obtained as described above by subtracting the controlled direct effect from the total effect. However, it is not necessary to make Assumption (d). Although an estimate of the indirect effect itself requires Assumption (d), the null hypothesis test of no mediation is still valid and we can obtain unbiased estimates of the causal effect of the intervention on risky relationship thoughts and of risky relationship thoughts on unprotected sex, holding constant intervention condition regardless of whether Assumption (d) holds.

Empirical Demonstration: RRR-HIV Intervention

Study design and participants

The data for these analyses came from the CJ-DATS RRR-HIV intervention. Briefly, the intervention includes five group sessions in prison (prior to release) and one individual telephone or face-to-face session after release. The intervention attempts to change risky relationship thoughts, such as “Having sex without a condom will strengthen my relationship,”

with the intent of reducing HIV-risk behaviors. Participants were recruited from correctional facilities in Connecticut, Delaware, Kentucky, and Rhode Island. Women were eligible to participate if there were at least 18 years of age, scheduled to go before the parole board in the next six weeks, had at least weekly substance use before incarceration, and were willing to be randomized to study condition. Intervention details, including consent procedures, are described in Havens et al. (2009) and Staton-Tindall et al. (2007).

Data were collected at three time points: baseline, 30 days post-release, and 90 days post-release. The participants include 243 women who were present at each of the three time points. Characteristics of this sample can be found in Table 1.

Measures

All study measures were based on participant self-report data collected by trained research staff. Each of the measures reports on behavior in the previous 30 days.

Outcome—The outcome for these analyses, measured at 90 days post-release, is unprotected sex, a binary variable defined as 1 = having had sex at least once without a condom in the previous 30 days and 0 =no sex without a condom or no sex in the previous 30 days.

Mediator—The mediator, *risky relationship thoughts*, is a scale score computed by taking the mean of six items. The scale of each item ranged from 1=never to 10=every day; higher scores correspond to more frequent risky relationship thoughts. Coefficient alpha for the items was . 68 (95% confidence interval (CI): .61, .74; Maydeu-Olivares, Coffman, & Hartmann, 2007). The items were based on focus groups by the study investigators prior to implementation of the intervention (Staton-Tindall et al., 2007).

Potential confounders—Eighteen characteristics were included in the propensity score model to control for the potential confounding between the mediator and the outcome. They included baseline socio-demographic characteristics such as race (Black vs. not), housing (own home vs. other), live with spouse/partner (yes/no), marital status (married, previously married, single), high school graduate (yes/no), employment status (full time, part time, unemployed), financial support from job (# of months), and have medical insurance (yes/no). They also included numerous cognitions and behaviors at baseline such as arrested (yes/no), spent time in jail (yes/no), been physically, sexually, or emotionally abused in previous six months (yes/no), drank alcohol (yes/no), alcohol use problems (mean of 8-item scale, 0 = never to 4 = always), received drug treatment since locked up (yes/no), drug use problems (mean of 8-item scale, 0=never to 4=always), substance use problems (mean of 3-item scale, 1=not at all to 4=extremely), substance use treatment important after release (yes/no), and sex without a condom at baseline intake (yes/no).

Statistical analysis

We make Assumptions (a) – (c) and use MSMs and an IPW estimator. IPW relies on propensity scores, defined as the probability that an individual receives a particular level of the intervention or exposure variable given measured confounders (Rosenbaum & Rubin 1983). Given that propensity score methods are based on an assumption of no unmeasured confounding, it is best to err on the side of being overly inclusive when adding potential confounders to the propensity score model. However, confounders that occur after the hypothesized causal variable or exposure should not be included in the propensity model as this will bias the causal effect estimate (Rosenbaum, 1984).

Estimating propensity scores and creating weights for continuous variables, such as risky relationship thoughts, is only slightly more difficult than it is for a binary variable. The

propensity score may be obtained from the probability density function (p.d.f.) of risky relationship thoughts given the measured confounders and treatment history, $\phi(M_i|X_i,T_i)$, (Robins et al., 2000). The propensity scores are obtained by a linear regression of M_i on X_i and T_i and a probability is obtained by inserting the fitted values from the regression, denoted $\hat{m}$, in the normal p.d.f. (denoted $\phi()$),

$$\phi(m)=\frac{1}{\sqrt{2\pi\hat{\sigma}^2}}\,e^{-\frac{(m-\hat{m})^2}{2\hat{\sigma}^2}},$$

where $\hat{\sigma}$ is the residual standard error from the regression of M_i on X_i and T_i.

Essentially, this approach treats the intervention-mediator sequence as a time-varying treatment. Therefore, the probability for the numerator of the weights for risky relationship thoughts is given by the p.d.f. of M given treatment history, $\phi(M_i|T_i)$. The probability for the denominator of the weights for the mediator is given by the p.d.f. of M given treatment history and the measured confounders, $\phi(M_i|T_i,X_i)$. Thus, the weights for risky relationship thoughts are $\phi(M_i|T_i)/\ \phi(M_i|T_i,X_i)$. For further details about creating weights and the numerator and denominator models for the weights, see Cole and Hernan (2008) and Robins et al. (2000). After the weights are created, they are incorporated into Equation (4) in the same manner as survey weights using SAS PROC GENMOD. We used SAS PROC GLM for fitting Equation (3) because no weights were needed since the intervention was randomized. Example SAS code is provided in the appendix.

Results

Before presenting results of the mediation analysis, we assess balance. Figure 2 illustrates what is commonly referred to as balance on the measured confounders and presents the correlations between risky relationship thoughts and the confounders included in the propensity model before and after weighting. It is recommended that these differences be less than .1 (in absolute value), which is considered a small effect size (Cohen, 1988). As shown in Figure 1, before weighting, several of the correlations were greater than |.1|, but after weighting, they were all less than |.1|. If they were not all less than |.1|, the propensity model should be revisited and interaction or quadratic terms added to the propensity model until balance is achieved.

To determine whether the effect of the intervention was mediated by risky relationship thoughts, we estimated each of the effects as described above. The effect of the intervention on risky relationship thoughts was significant (−0.529, p = .03, 95% CI: −1.00, −.06) such that the intervention resulted in a decrease in risky relationship thoughts. The effect of risky relationship thoughts on unprotected sex was also statistically significant (.447, p < .001, 95% CI: .22, .68). A one-unit increase in risky relationship thoughts resulted in a 1.56 times increase in the odds of unprotected sex. The effect of the intervention on unprotected sex, holding constant risky relationship thoughts, was not significant (0.388, p = .479, 95% CI: −.69, 1.46). We also included a term for the interaction between the intervention and risky relationship thoughts; however, it was not statistically significant (−.112, p = .495, 95% CI: −.43, .21). The total effect of the RRR-HIV intervention on unprotected sex (i.e., risky relationship thoughts were not in the model) was not significant (−.092, p = .731, 95% CI: −.61, .43).1 Following

[1]The requirement of a significant total effect is controversial because there can be a mediated effect even if the total effect is not significant. For example, this may happen if there is another mediator and the effect through that mediator cancels out the effect through the original mediator. It may also happen if there are moderators and there is positive mediated effect in one subgroup and a negative mediated effect in another subgroup. It is now generally accepted in the statistical mediation literature that a significant total effect is not required (MacKinnon, 2008).

the procedures in Coffman and Zhong (2011), we rejected the null hypothesis that either the effect of the intervention on risky relationship thoughts or the effect of risky relationship thoughts on unprotected sex, holding constant intervention condition, was zero ($p = .028$). Based on this information, we conclude that the intervention's effect on unprotected sex was indeed mediated by risky relationship thoughts.

Discussion

In this paper, we sought to describe an approach using the potential outcomes framework for causal inference to assess causal mediation. This is important because mediation is, by definition, a question about causal pathways. Even if individuals are randomly assigned to levels of the intervention and randomization does not fail (e.g., no non-compliance), this *does not* imply that individuals are randomly assigned to levels of the mediator. In fact, confounders of the mediator and outcome almost always exist. Without proper control of these confounders, the estimate of the effect of the mediator on the outcome and the estimate of the direct effect of the intervention on the outcome will be biased. IPW is one approach to controlling for confounders. Another approach would be to control for all confounders using regression adjustment (i.e., ANCOVA); however, propensity scores are advantageous because they reduce a potentially large number of confounders into a single-number summary. Furthermore, regression adjustment may still result in biased estimates of the direct effect if post-treatment confounders are included in the regression model (Robins et al., 2000).

The RRR-HIV intervention was used as a motivating example because of the public health significance of HIV/AIDS, especially among incarcerated women. However, the public use data set posed challenges because access to all the potential mediators and confounders of this study were unavailable. In particular, there was no measure of the mediator at baseline. The baseline measure of the mediator is an obvious potential confounder of the mediator at 30-day follow-up and of the outcome at 90-day follow-up. The method described above relies on the assumption that all confounders are measured and properly controlled. This is a very strong assumption that cannot be tested in practice. However, the more potential confounders that are included in the propensity model, the more plausible the assumption becomes. Thus, it is imperative that researchers measure as many potential confounders as possible. In addition, the impact of the unmeasured confounder is mitigated if a measured potential confounder is highly correlated with the unmeasured confounder. Nevertheless, a next step with approaches that rely on the assumption of no unmeasured confounding is to conduct a sensitivity analysis, which attempts to determine how strongly influential an unmeasured confounder would need to be in order to change the estimate in a meaningful way (e.g., change the estimate from statistically significant to not, reverse the sign of the estimate). Sensitivity analysis is still being developed for continuous exposures/mediators. Despite challenges of the data set, we were able to demonstrate how this approach can be used to assess mediation. It should be noted that the traditional regression approach to mediation, which is likely familiar to most readers, also assumes no unmeasured confounding of the mediator and outcome, although this assumption is rarely stated explicitly. Furthermore, in the traditional approach, researchers typically control for only a few demographic variables, if they control for any potential confounders at all. As shown by Steiner, Cook, Shadish, and Clark (2010), controlling for only a few demographic variables is generally insufficient to obtain unbiased estimates.

For comparison, we fit the model without the weights; that is, a logistic regression of unprotected sex on the intervention and risky relationship thoughts. As with the IPW estimates, the effect of the intervention on unprotected sex, holding constant risky relationship thoughts, was not statistically significant. The effect of risky relationship thoughts on unprotected sex, holding constant the intervention condition, was statistically significant (.352, $p < .001$, 95% CI: .20, .50). Thus, a one-unit increase in risky relationship thoughts resulted in a 1.42 times

increase in the odds of unprotected sex. This estimate is not dramatically different from the IPW estimate (the CIs overlap); however, it is a 14% difference in the odds ratio point estimates.

We believe that the controlled effects approach is valuable in nursing research. Controlled effects differ from the traditional regression approach to mediation primarily in that their estimation uses weights, which account for potential confounding. The controlled effects approach addresses questions such as, "What is the effect of the RRR-HIV intervention on unprotected sex, holding constant the level of risky relationship thoughts?", "What is the effect of the RRR-HIV intervention on risky relationship thoughts?", and "What is the effect of risky relationship thoughts on unprotected sex, holding constant intervention status?" If the no-interaction assumption holds, then this approach also addresses the question, "What is the effect of the RRR-HIV intervention on unprotected sex that is due to risky relationship thoughts?" The indirect effect itself is not identified unless there is no interaction between the intervention and mediator. Nevertheless, the two effects that make up the indirect effect (i.e., the effect of the intervention on the mediator, and the effect of the mediator on the outcome, holding constant the intervention condition) are identified regardless of whether or not there is an interaction between the intervention and mediator.

Conclusions

In summary, an advantage of the potential outcomes framework is that it allows for the careful definition of causal effects and of the assumptions needed for identification and estimation of the causal effects. From a substantive standpoint, the results of this study provide support for interventions that target aspects of relationship thoughts and beliefs, consistent with findings from other studies (Knudsen et al., 2008; Pulerwitz, Amaro, DeJong, & Gortmaker, 2002).

Acknowledgments

Preparation of this article was supported by National Institute on Drug Abuse (NIDA) Center Grant P50 DA100075-15 and National Cancer Institute (NCI) Center Grant P50 CA143188. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIDA, NCI, or the National Institutes of Health (NIH). We thank Amanda Applegate for editorial assistance.

References

Adams J, Nowels C, Corsk K, Long J, Steiner JF, Binswanger IA. HIV risk after release from prison: A qualitative study of former inmates. JAIDS: Journal of Acquired Immune Deficiency Syndromes. 2011; 57:429–434.

Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. Journal of Personality and Social Psychology. 1986; 51:1173–1182. [PubMed: 3806354]

Bray BC, Almirall D, Zimmerman RS, Lynam D, Murphy SA. Assessing the total effect of time-varying predictors in prevention research. Prevention Science. 2006; 7:1–17. [PubMed: 16489417]

Coffman DL. Estimating causal effects in mediation analysis using propensity scores. Structural Equation Modeling. 2011; 18:357–369. [PubMed: 22081755]

Coffman DL, Caldwell L, Smith E. Introducing the at-risk average causal effect with application to HealthWise South Africa. Prevention Science. (in press).

Coffman DL, Zhong W. Assessing mediation using marginal structural models in the presence of confounding and moderation. 2011 Manuscript submitted for publication.

Cohen, J. Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: LEA; 1988.

Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. American Journal of Epidemiology. 2008; 168:656–664. [PubMed: 18682488]

Cole SR, Hernan MA, Robins JM, Anastos K, Chmiel J, Detels R, Munoz A. Effect of highly active antiretroviral therapy on time to aquired immunodeficiency syndrome or death using marginal structural models. American Journal of Epidemiology. 2003; 158:687–694. [PubMed: 14507605]

Cotton-Oldenburg NU, Jordan BK, Martin SL, Kupper L. Women inmates' risky sex and drug behaviors: Are they related? American Journal of Drug and Alcohol Abuse. 1999; 25:129–149. [PubMed: 10078982]

Fletcher BW, Wexler HK. National Criminal Justice Drug Abuse Treatment Studies (CJ-DATS): Update and progress. Justice Research and Statistics Association Forum. 2005; 23:1–7.

Havens JR, Leukefeld CG, Oser CB, Staton-Tindall M, Knudsen HK, Mooney J, Inciardi JA. Examination of an interventionist-led HIV intervention among criminal justice-involved female prisoners. Journal of Experimental Criminology. 2009; 5:245–272. [PubMed: 20090928]

Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychological Methods. 2010; 15:309–334. [PubMed: 20954780]

Jo B. Causal inference in randomized experiments with mediational processes. Psychological Methods. 2008; 13:314–336. [PubMed: 19071997]

Knudsen HK, Leukefeld C, Havens JR, Duvall JL, Oser CB, Staton-Tindall M, Inciardi JA. Partner relationships and HIV risk behaviors among women offenders. Journal of Psychoactive Drugs. 2008; 40:471–481. [PubMed: 19283951]

Krause MR, Serlin RC, Ward SE, Rony RYZ, Ezenwa MO, Naab F. Testing mediation in nursing research: Beyond Baron and Kenny. Nursing Research. 2010; 59(4):288–293. [PubMed: 20467337]

Levy JA, Landerman LR, Davis LL. Advances in mediation analysis can facilitate nursing research. Nursing Research. 2011; 60(5):333–339. [PubMed: 21873916]

Lichtenstein B, Malow R. A critical review of HIV-related interventions for women prisoners in the United States. Journal of the Association of Nurses in AIDS Care. 2010; 21:380–394. [PubMed: 20350816]

Little RJA, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: Concepts and analytical approaches. Annual Review of Public Health. 2000; 21:121–145.

MacKinnon, DP. Introduction to statistical mediation analysis. New York: LEA; 2008.

Maruschak, LM. HIV in Prisons, 2007–08. Bureau of Justice Statistics; 2009.

Maydeu-Olivares A, Coffman DL, Hartmann WM. Asymptotically distribution-free interval estimation for coefficient alpha. Psychological Methods. 2007; 12(2):157–176. [PubMed: 17563170]

Pulerwitz J, Amaro H, DeJong W, Gortmaker SL. Relationship power, condom use and HIV risk among women in the USA. AIDS Care. 2002; 14:789–800. [PubMed: 12511212]

Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology. 1992; 3:143–155. [PubMed: 1576220]

Robins JM, Hernan MA, Brumback BA. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000; 11:550–560. [PubMed: 10955408]

Rosenbaum PR. The consequences of adjustment for a concomitant variable that has been affected by the treatment. Journal of the Royal Statistical Society, Series A (General). 1984; 147:656–666.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. Journal of Educational Psychology. 1974; 66:688–701.

Schafer JL, Kang J. Average causal effects from non-randomized studies: A practical guide and simulated example. Psychological Methods. 2008; 13:279–313. [PubMed: 19071996]

Staton-Tindall M, Leukefeld C, Palmer J, Oser C, Kaplan A, Krietemeyer J, Surratt HL. Relationships and HIV risk among incarcerated women. The Prison Journal. 2007; 87:143–165.

Steiner PM, Cook TD, Shadish WR, Clark MH. The importance of covariate selection in controlling for selection bias in observational studies. Psychological Methods. 2010; 15(3):250–267. [PubMed: 20822251]

VanderWeele TJ. Marginal structural models for the estimation of direct and indirect effects. Epidemiology. 2009; 20:18–26. [PubMed: 19234398]

Appendix

```
*estimate propensity scores and create weights;
*denominator model;
proc reg data=cjdats.final;
model RRA_30 = cond and list all confounders;
output out=cjdats.final student=rden;
run;
quit;
*numerator model;
proc reg data=cjdats.final;
model RRA_30 = cond;
output out=cjdats.final student=rnum;
run;
quit;
data cjdats.weights;
set cjdats.final;
pnum = exp(-.5*(rnum**2))/2.506;
pden = exp(-.5*(rden**2))/2.506;
wt = pnum/pden;
run;
*assess balance;
proc corr data=cjdats.weights;
var RRA_30;
with names of all confounders;
run;
proc corr data=cjdats.weights;
var RRA_30;
with names of all confounders;
weight wt;
run;
*outcome analysis;
proc genmod data=cjdats.weights descending;
model condomyn= cond RRA_30 cond*RRA_30 /link=logit dist=binomial;
weight wt;
run;
quit;
proc genmod data=cjdats.weights;
model RRA_30 = cond;
run;
quit;
```

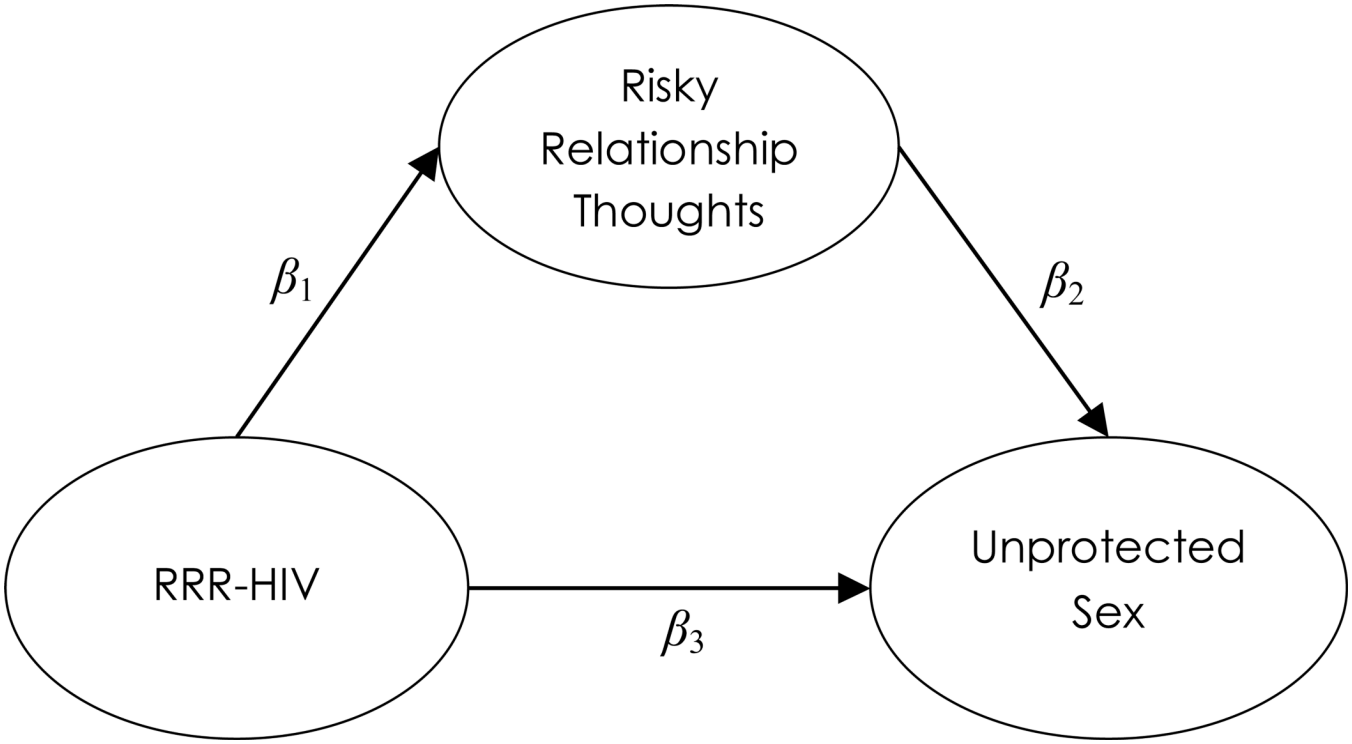

Figure 1.
Theoretical mediation model. RRR-HIV = Reducing Risky Relationships – Human Immunodeficiency Virus intervention.

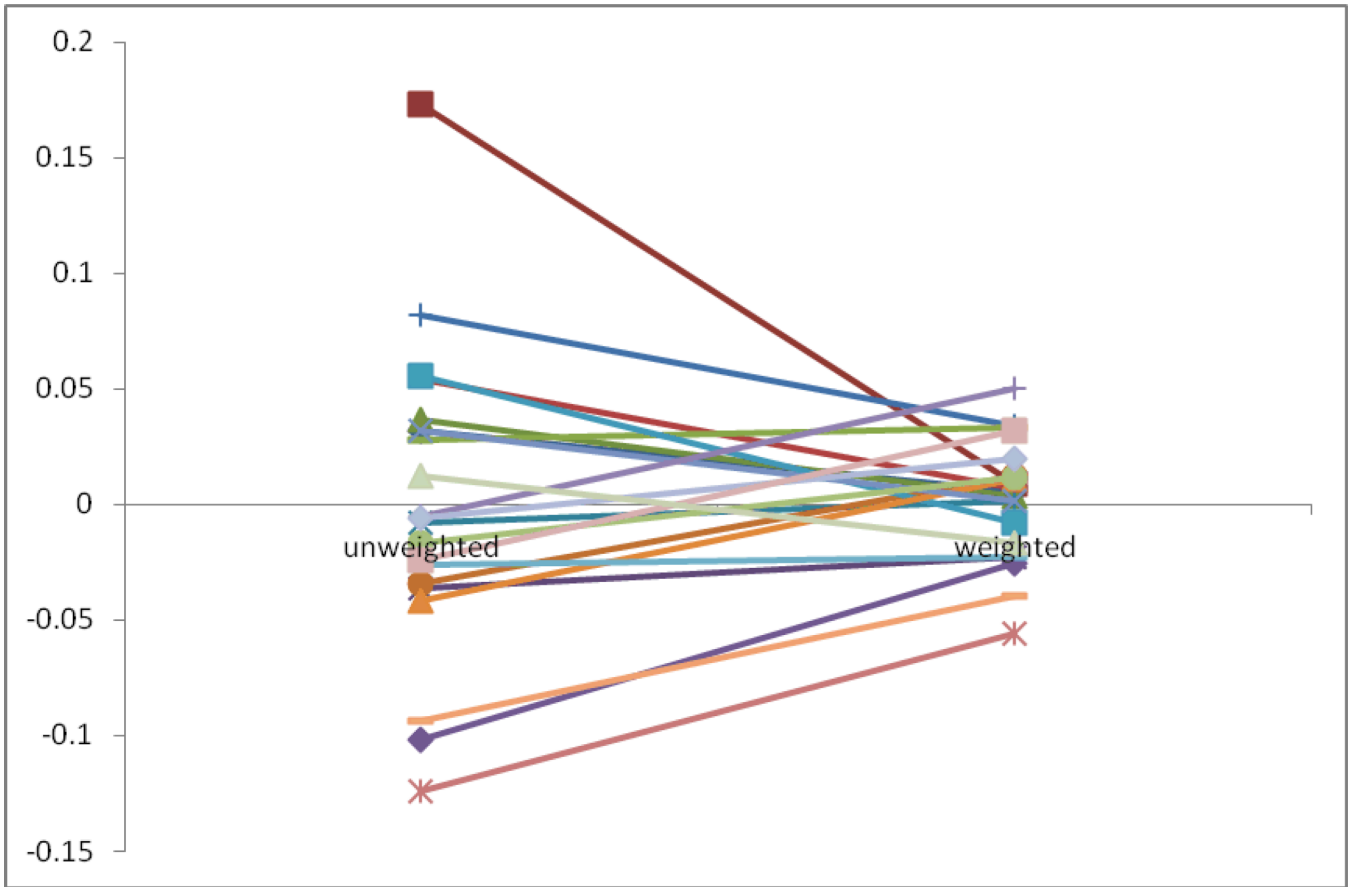

Figure 2.
Correlations between risky relationship thoughts and each confounder before and after weighting. Each line represents one of the potential confounders included in the propensity model.

Table 1

Demographic characteristics and HIV-risk related cognitions and behaviors (*N*=243)

Characteristic	*N*	% (or mean)
Race		
White	168	69.1
African American	59	24.3
Hispanic	11	4.5
Other	16	6.6
Marital status		
Married/cohabitating	33	13.6
Single, never married	113	46.5
Separated/divorced	97	39.9
High school graduate (% Yes)	111	45.7
Employment		
Full-time	74	30.5
Part-time	38	15.6
Unemployed	130	53.5
Arrested (past 30 days)	177	72.8
Abused (past 6 months)	43	17.7
Unprotected sex (past 30 days)	188	77.4