

Review

Assumptions Not Often Assessed or Satisfied in Published Mediation Analyses in Psychology and Psychiatry

Elizabeth A. Stuart*, Ian Schmid, Trang Nguyen, Elizabeth Sarker, Adam Pittman, Kelly Benke, Kara Rudolph, Elena Badillo-Goicoechea, and Jeannie-Marie Leoutsakos

* Correspondence to Dr. Elizabeth A. Stuart, Johns Hopkins Bloomberg School of Public Health, Hampton House Room 839, 624 N. Broadway, Baltimore, MD, 21205 (e-mail: estuart@jhu.edu)

Accepted for publication September 13, 2021.

Mediation analysis aims to investigate the mechanisms of action behind the effects of interventions or treatments. Given the history and common use of mediation in mental health research, we conducted this review to understand how mediation analysis is implemented in psychology and psychiatry and whether analyses adhere to, address, or justify the key underlying assumptions of their approaches. All articles ($n = 206$) were from top academic psychiatry or psychology journals in the PsycInfo database and were published in English from 2013 to 2018. Information extracted from each article related to study design, covariates adjusted for in the analysis, temporal ordering of variables, and the specific method used to perform the mediation analysis. In most studies, underlying assumptions were not adhered to. Only approximately 20% of articles had full temporal ordering of exposure, mediator, and outcome. Confounding of the exposure–mediator and/or mediator–outcome relationships was controlled for in fewer than half of the studies. In almost none of the articles were the underlying assumptions of their approaches discussed or causal mediation methods used. These results provide insights to how methodologists should aim to communicate methods, and motivation for more outreach to the research community on best practices for mediation analysis.

causal inference; mental health; statistical methods

Abbreviation: B&K, Baron and Kenny.

INTRODUCTION

Identifying the mechanisms behind exposures, treatments, or interventions is crucial, yet difficult. Examples of the sorts of research questions that are addressed by examining mechanisms include identifying which components of a complex behavioral intervention are most effective at reducing suicidal behavior, or examining whether a primary care service intervention reduces depression and suicidality in elderly adults by increasing depression diagnosis and treatment rates, treatment quality, or some or all of these, or whether adding cognitive behavior therapy on top of pharmacological treatment helps reduce anxiety symptoms by reducing negative self-talk, increasing coping efficacy, and/or improving medical adherence. Such knowledge could help fine-tune the treatment or intervention for enhanced effects. One strategy for assessing mechanisms is through what is known as mediation analysis, by which one broadly

aims to assess whether the effects of some exposure (A) on an outcome (Y) of interest go through (or are mediated by) an intermediate variable, the “mediator” (M). The idea of mediation is that the exposure influences the mediator, which in turn influences the outcome, and that there may also be direct effects of the exposure on the outcome. Interest in mediation analysis, which arguably originated in psychology in the 1980s and is still commonly used in that field, is increasing. For example, the number of articles in the PsycInfo database that mention mediation analysis in their abstracts increased sharply from 50 in 2007 to 917 in 2019. The importance of mediation analysis to understand what works and why is recognized and highlighted in the National Institute of Mental Health’s clinical trials request for proposals for effectiveness studies of treatment, prevention, and services interventions (1). The fields of psychology and psychiatry thus represent a useful case field

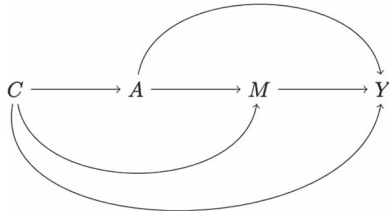


Figure 1. Single-mediator case. A, exposure; C, covariates; M, mediator; Y, outcome.

for examining the use of mediation analysis in practice, with implications for fields more broadly, including many areas within epidemiology.

The majority of published mediation analyses are in a classic tradition (2–4) that started as a result of Baron & Kenny’s influential 1986 paper (2) (which amassed >98,000 citations as of January 16, 2021, according to Google Scholar), building on path analysis work going back to Wright’s 1934 paper (5). Central to this Baron and Kenny tradition is the definition of the indirect effect (i.e., the effect of the exposure on the outcome that goes through the mediator) as the product of 2 regression coefficients. In a simple case with a single mediator (Figure 1), denote the exposure, outcome, mediator, and covariates by A, Y, M, and C, respectively. A classic analysis would fit 2 models for the mediator and outcome; the following linear models are examples:

$$E[M|C, A] = \alpha_0 + \alpha_1 A + \alpha_2 C$$

and

$$E[Y|C, A, M] = \beta_0 + \beta_1 A + \beta_2 M + \beta_3 C.$$

The indirect effect would be defined as $\alpha_1 \beta_2$, the product of coefficient α_1 (representing the effect of exposure A on mediator M) and coefficient β_2 (representing the effect of M on outcome Y); the direct effect would be defined as β_1 . In this review, we use the term “B&K” (referring to Baron and Kenny) to mean to all analyses that define the effect as the product of coefficients; this also includes methods that use structural equation models to fit the models that then multiply the relevant coefficients. See MacKinnon (3), Hayes (6), and Hayes and Preacher (7) for overviews of this broad approach.

Although the B&K approach is simple and intuitive, an important limitation is that its effect definitions are highly model dependent. Methods in this tradition generally assume no interaction of exposure and the mediator in influencing the outcome (8), which is restrictive because there are cases where such an interaction is expected (9). Also, if nonlinear models are used, the product $\alpha_1 \beta_2$ generally does not match, and may not be on the same scale as, the causal effect of interest to be estimated.

Inherent to mediation analysis is an assumption that the exposure influences the mediator and the exposure and mediator both influence the outcome. That is, mediation is a causal hypothesis, which is the common starting point regardless of which methodological approaches are used to

evaluate the hypothesis. This causal hypothesis implies a temporal ordering of these variables that, at the minimum, should be reflected in the data. Without temporal ordering in the data, there is no way to establish that the temporal ordering is, in fact, exposure first, then mediator, and finally outcome, rather than another order. Yet authors of past reviews have documented the common use of cross-sectional data in mediation analysis. Over a decade ago, in a review of mediation analyses published in psychology journals in 2005, authors reported that in more than half of the reviewed articles, authors used cross-sectional data, and most of the remaining articles reported on studies in which exposure and mediator or mediator and outcome were measured at the same time (8). This prompted the publication of several articles by quantitative psychologists such as Maxwell and Cole (9) and Mitchell and Maxwell (10) in which they clarified and warned of bias of mediation effects when using cross-sectional data. More recently, Vo et al., in their review (11) of mediation analyses using randomized clinical trial data published in 2017–2018, found that although study design provides for temporal precedence of the exposure, more than half of the studies used concurrent mediator and outcome. Their review, however, was not specific to psychology or psychiatry research.

A second broad assumption of mediation analysis (because it is an analysis of causal effects) is no uncontrolled confounding. Confounding is a well-known problem for the simple 2-variable exposure–outcome analysis, where a common remedy taught in most elementary data analysis courses is to adjust for covariates (i.e., confounders) in regression models. For mediation analysis, confounding is a more complicated problem, because instead of 1, there are 3 relationships (exposure–outcome, exposure–mediator, and mediator–outcome) that can be confounded. Randomizing the exposure (e.g., in a clinical trial) takes care of the first 2 types of confounding but not the third. The mediator can almost never be randomized; this nuance is often underappreciated by researchers conducting mediation analyses in randomized trials (12). Bias due to uncontrolled mediator–outcome confounding was pointed out in 1981 by Judd and Kenny (13) but unfortunately was not mentioned in Baron and Kenny’s 1986 paper (2). As shown by Judd and Kenny (13), Maxwell et al. (9), and Valente et al. (14), it is particularly important for bias reduction to control for baseline measures of the mediator and outcome.

The recent incorporation of a causal inference framework (15) in mediation research (16), which results in the causal mediation approach, addresses these issues. The causal mediation approach starts with model-free definitions of causal effects and clarifies the assumptions that need to hold for mediation analysis (16–22). The sorts of questions for which these methods are used to answer include, for example, “If we could fix M at a specific value m, how would that change outcomes and what would be the effect of A on Y?” or “What would be the effect of A on Y if we could block the direct path from A to M in some manner?” The first question corresponds to an effect type termed “controlled direct effects,” and the second question corresponds to 1 of several types of indirect effects, termed “natural” or “interventional,” depending on the manner of the blocking;

different effect types are relevant to different real-world research questions (22). This framework provides new insight on confounding. In broad strokes, the assumption of no unmeasured exposure–outcome and mediator–outcome confounding is required for mediation analyses targeting all effect types, and if interventional direct or indirect effects are of interest, it is also required that there be no unmeasured exposure–mediator confounding. These 3 assumptions are an intuitive extension from the 2-variable exposure–outcome setting. What is discovered due to the causal inference framework is that the natural direct or indirect effects (the closest match to the common interest in decomposing the total effect into direct and indirect components) require a fourth assumption that no mediator–outcome confounders (observed or not) are influenced by exposure (often called “no post-treatment confounders”). There is now a substantial body of statistical literature with methods for estimating these different effects on the basis of relevant assumptions.

Despite the recent methodological advances for mediation analysis, it is unclear whether those methods have been adopted by applied researchers or if the key assumptions underlying mediation analyses, in general, are recognized. In this article, we present results of a review of research articles published between 2013 and 2018 that include mediation analyses. We focus on the methods used and whether they addressed key assumptions that underlie basically all mediation analyses, such as temporal ordering of the variables and the handling of confounding. We found that authors generally did not discuss or examine the key underlying assumptions in the paper, many did not deal with potential confounding, and many did not even have temporal ordering in their studies. This provides motivation for further dissemination of best practices for mediation analysis.

METHODS

Those assumptions underlying mediation analysis are not trivial. In this literature review, we sought to understand how mediation analysis is being implemented in the psychology and psychiatry literature and, in particular, whether key assumptions are addressed.

To assess the use of methods in what should be the top-quality manuscripts, we extracted information from articles published in top journals in psychology and psychiatry. In particular, we took the top 15 psychology and 15 psychiatry journals as ranked by Google Scholar based on their h5-index (see listing in Web Table 1). Using the PsycInfo database, we searched these journals for articles published in 2013–2018 that contained the phrase “mediation analysis*” in the title, abstract, or keywords. We excluded those that did not report on a mediation analysis. The resulting 206 articles were included in our study.

Two unique reviewers were randomly assigned to extract information from each article. In this report, we focus in particular on the aspects we described previously, which can be considered foundational assumptions underlying nearly all mediation analyses:

1. Temporal ordering
2. Whether the exposure was randomized

3. Confounding adjustment

- a. Whether the analysis adjusted for confounders and explained which relationships were deemed confounded by which variables
 - b. Whether the variables adjusted for included baseline measures of the mediator and/or outcome
- ### 4. Which mediation method was used; in particular, whether it was a method in the B&K tradition or a causal mediation approach
- a. If a B&K approach was used, whether the authors assessed the plausibility of no exposure–mediator interaction in the outcome model

Note that regarding confounding, to be fair to all articles, we did not use the assumptions clarified in the causal mediation approach as evaluation standards. Specifically, practitioners of the B&K approach may not be familiar with the assumption of no post-treatment confounders. We only tracked whether the articles considered exposure–mediator, exposure–outcome, and mediator–outcome confounders, because this notion of confounding of the 3 relationships is a natural extension of the well-known confounding issue in the 2-way exposure–outcome analysis.

RESULTS

In nearly all the articles (97%), the researchers used a B&K approach. Across the period examined, causal mediation approaches were used in only 3% of studies and with no dramatic change over time; the methods used for causal mediation were from either Imai et al. (18) or Valeri and VanderWeele (23).

Of the 206 articles, 21.5% reported on a study with randomized exposure. Figure 2 indicates that analyses often did not have temporal ordering; cross-sectional data were used in slightly more than 50% of the studies. Only 23.8% had temporal ordering of all 3 variables: exposure to mediator to outcome; this is about the same level found by Maxwell and Cole in their review of psychology research published 15 years ago (8). Temporal ordering was more common in randomized trials, with about 80% having at least temporal ordering of exposure before mediator or mediator before outcome (and almost 50% having full temporal ordering of all 3 variables); this is similar to the finding reported by Vo et al. (11).

In terms of controlling confounding, in only 60.7% of the articles did the authors indicate they adjusted for some covariates (Table 1), meaning that in more than a third of the articles, the researchers did not consider confounding an issue. The covariates were adjusted to include baseline measures of the mediator and outcome in only a small fraction of the studies (13.6% for the mediator, 16% for the outcome, and 11.7% for both). In addition, it is important to note that most authors simply listed the set of covariates controlled for, without any specific justification for those as potential confounders of any of the relationships of interest (i.e., exposure–mediator, exposure–outcome, or mediator–outcome).

Table 1. Proportion of Articles in This Review in Which Certain Mediation Assumptions Were Addressed

Study Characteristic	% of Articles (n = 206)
Temporal ordering of exposure and mediator	43.7
Temporal ordering of mediator and outcome	30.6
Temporal ordering of exposure, mediator, and outcome	23.8
Adjusted for covariates	60.7
Adjusted for covariates that include baseline measure of mediator	13.6
Adjusted for covariates that include baseline measure of outcome	16.0
Adjusted for covariates that include baseline measures of mediator and outcome	11.7
Discussed mediation assumptions at all	42.2
Examined assumption of no interaction of exposure and mediator	2.5
Sensitivity analysis performed to mediation assumptions	1.5

In only 2.5% of the articles in which authors indicated a B&K approach was used did they include or justify their exclusion of an exposure–mediator interaction in the outcome model. Fewer than half of the articles (42.2%) were scored as having some discussion of mediation assumptions. We note that in most cases, this discussion was simply an acknowledgement that the data were cross-sectional and thus results should be interpreted with caution. Separate from this point, a sensitivity analysis for mediation assumptions was conducted in only 1.5% of the studies. These general patterns of results were similar across the psychology and psychiatry journals (Web Table 2).

DISCUSSION

It is clear that mediation is of great interest in psychology and psychiatry. However, current practice does not seem to acknowledge key limitations of all mediation analyses, including the potential for confounding and the need for

temporal ordering. In addition, specific approaches have their own assumptions, which are often not addressed (e.g., approaches in the B&K tradition generally assume no mediator–exposure interaction, which is not always realistic). Causal methods can relax that assumption but have been used rarely in psychology or psychiatry.

We note that neglecting to justify the inclusion of covariates in terms of confounding control or to assess for an exposure–mediator interaction on the outcome could be issues of reporting rather than practice. It could be the case that the investigators considered confounders of all the different relationships but just listed them all in a brief sentence in the article or withheld that they tested for, but did not find evidence, of an exposure–mediator interaction. However, that these types of considerations are not reported across the board suggests they may not be routinely practiced.

There are many potential reasons why the fields of psychology and psychiatry have not adapted a causal mediation approach, including lack of awareness of the methods and/or software to use them. Another reason may be that the B&K approach easily allows for the incorporation of measurement models for latent constructs. Such constructs are somewhat rare in medicine and public health but quite common in psychology and psychiatry.

These findings should provide additional justification and motivation for further bridging of the statistical methods and their use in practice, and to ensure that applied users of methods understand the underlying assumptions. Although we focused within the fields of psychology and psychiatry in this review, given the long history of mediation analysis in those fields, we expect the findings likely carry over to other fields using mediation analyses.

Causal mediation analysis offers the potential to answer highly relevant questions in public health and epidemiology, but it is a particularly challenging area. Applied users need tools to help them assess the validity of the underlying assumptions and sensitivity of results to violation of those assumptions. Doing so will ensure that the results obtained are accurate and lead to improvements in our knowledge of interventions, treatments, and risk factors across many areas of public health and epidemiology.

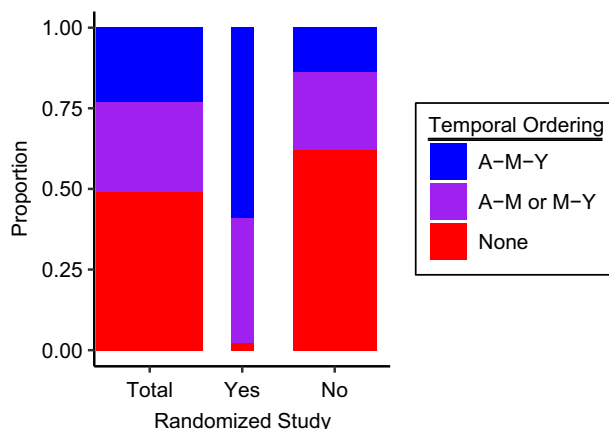


Figure 2. Proportion of studies with temporal ordering, stratified by whether exposure (A) was randomized. Horizontal bar size reflects the relative number of randomized and nonrandomized studies. M, mediator; Y, outcome.

ACKNOWLEDGMENTS

Author affiliations: Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States (Elizabeth A. Stuart, Ian Schmid, Trang Nguyen, Kelly Benke, Elena Badillo-Goicoechea); Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States (Elizabeth Sarker, Adam Pittman); Department of Epidemiology, Columbia University, New York, New York, United States (Kara Rudolph); Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, Maryland, United States (Jeannie-Marie Leoutsakos).

This work was funded by the National Institute of Mental Health (grants R01MH115487 and T32MH122357).

Conflict of interest: none declared.

REFERENCES

- National Institutes of Health. Clinical trials to test the effectiveness of treatment, preventive, and services interventions (R01 Clinical Trial Required). Reissue of RFA-MH-17-608. <https://grants.nih.gov/grants/guide/rfa-files/RFA-MH-18-701.html>. Accessed April 22, 2020.
- Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6):1173–1182.
- MacKinnon DP. *Introduction to Statistical Mediation Analysis*. New York, NY: Taylor & Francis; 2008.
- Preacher KJ, Rucker DD, Hayes AF. Addressing moderated mediation hypotheses: theory, methods, and prescriptions. *Multivar Behav Res*. 2007;42(1):185–227.
- Wright S. The method of path coefficients. *Ann Math Stat*. 1934;5:161–215.
- Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. New York, NY: The Guilford Press; 2013.
- Hayes AF, Preacher KJ. Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivar Behav Res*. 2010;45(4):627–660.
- Maxwell SE, Cole DA. Bias in cross-sectional analyses of longitudinal mediation. *Psychol Methods*. 2007;12(1):23–44.
- Maxwell SE, Cole D, Mitchell MA. Bias in cross-sectional analyses of longitudinal mediation: partial and complete mediation under an autoregressive model. *Multivar Behav Res*. 2011;46(5):816–841.
- Mitchell MA, Maxwell SE. A comparison of the cross-sectional and sequential designs when assessing longitudinal mediation. *Multivar Behav Res*. 2013;48(3):301–339.
- Vo TT, Superchi C, Boutron I, et al. The conduct and reporting of mediation analysis in recently published randomized controlled trials: results from a methodological systematic review. *J Clin Epidemiol*. 2020;117:78–88.
- Imai K, Keele L, Yamamoto T. Identification, inference and sensitivity analysis for causal mediation effects. *Stat Sci*. 2010;25(1):51–71.
- Judd CM, Kenny DA. Process analysis: estimating mediation in treatment evaluations. *Eval Rev*. 1981;5(5):602–619.
- Valente MJ, MacKinnon DP, Mazza GL. A viable alternative when propensity scores fail: evaluation of inverse propensity weighting and sequential G-estimation in a two-wave mediation model. *Multivar Behav Res*. 2020;55(2):165–187.
- Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol*. 1974;66(5):688–701.
- Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992;3(2):143–155.
- Pearl J. Direct and indirect effects. *Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence*, 2001;411–420.
- Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods*. 2010;15(4):309–334.
- Tchetgen EJ, Shpitser I. Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness and sensitivity analysis. *Ann Stat*. 2012;40(3):1816–1845.
- Didelez V, Dawid AP, Geneletti, S. Direct and indirect effects of sequential treatments. *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence*. 2006;138–146.
- VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology*. 2014;25(2):300–306.
- Nguyen T, Schmid I, Stuart EA. Clarifying causal mediation analysis for the applied researcher: defining effects based on what we want to learn. *Psychol Methods*. 2021;26(2):255–271.
- Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods*. 2013;18:137–150.