Journal of Clinical Epidemiology 117 (2020) 78−88

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

The conduct and reporting of mediation analysis in recently published randomized controlled trials: results from a methodological systematic review

Tat-Thang Vo[a,b,*], Cecilia Superchi[b,c], Isabelle Boutron[b], Stijn Vansteelandt[a,d]

[a]*Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281-S9, 9000, Ghent, Belgium*
[b]*Université de Paris, CRESS, INSERM, INRA, F-75004, Paris, France*
[c]*Department of Statistics and Operations Research, Barcelona-Tech, UPC, c/ Jordi Girona 1-3, 08034, Barcelona, Spain*
[d]*Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK*

Accepted 1 October 2019; Published online 5 October 2019

Abstract

Objectives: To describe the methodological characteristics of mediation analyses (MAs) reported in recent randomized controlled trials (RCTs) and to propose recommendations on the planning, conduct, and reporting of MAs in practice.

Study Design and Setting: We conducted a systematic review by searching MEDLINE (January 1, 2017, to December 1, 2018) for all reports of RCTs or secondary analyses of previously published RCTs that reported a MA. Two reviewers independently screened the title, abstracts, and full texts of the identified reports and extracted the data from the 98 eligible studies.

Results: MAs were nearly always (96%) based on a traditional mediation approach. Most studies did not report a sample size calculation for the MA (96%) or assess potential treatment-by-mediator interactions (96%). In 53% of studies, mediators and outcomes were simultaneously measured. In 57% of studies, mediator-mediator and mediator-outcome confounders were adjusted for in the analysis, although adjustment was often limited to few potential confounders. About 30% of studies discussed the assumptions underlying the MA.

Conclusion: The conduct and reporting of MAs remained quite heterogeneous in practice. Future MAs could benefit from a consensus-based planning, conduct, and reporting guideline for MA.

Keywords: Mediation analysis; Traditional mediation approaches; Counterfactual-based mediation approaches; Reporting; Systematic review; Randomized controlled trial

1. Introduction

Mediation analysis (MA) is a very common type of statistical analysis in psychology, sociology, epidemiology, and medicine [1−3]. Such analysis aims at assessing the relative magnitude of different pathways and mechanisms by which an exposure may affect an outcome [1−5]. As an example, Stensrud and Strohmaier conducted an MA to investigate whether the intent-to-treat (ITT) effect of intensive vs. standard systolic blood pressure treatment on cardiovascular risk was mediated by an indirect, potentially harmful effect through too low diastolic blood pressure [6]. If results of this MA show that the intensive treatment may lead to a diastolic hypotension, which in turn is associated with unfavorable cardiovascular outcomes, then caution should be taken when

Funding: This study is part of the January 2019 Journal Club, organized by T.-T.V., C.S., and S.V. within the Methods in Research on Research (MiRoR) project. This project received funding from the European Union's Horizon 2020 research and innovation program, under the Marie Sklodowska-Curie grant agreement (grant no.: 676207).

Availability of data and materials: The data set supporting the conclusions of the research reported in this article will be available in the Zenodo repository in the Methods in Research on Research (MiRoR) community (https://zenodo.org/communities/miror/?page=1&size=20)

Author's contributions: T.-T.V., C.S., and S.V. contributed to conceptualizing this review. T.-T.V. and C.S. conducted the data extraction and analysis. T.-T.V. and S.V. wrote the recommendations. T.-T.V. and C.S. drafted the manuscript. I.B., and S.V. made revisions. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate: Not required.

Consent for publication: Not required.

Conflict of interest: The authors declare no conflict of interest.

* Corresponding author. Department of Applied Mathematics, Computer Science and Statistics, Ghent University, Krijgslaan 281-S9, 9000, Ghent, Belgium. Tel.: +329 264 47 73.

E-mail address: tatthang.vo@ugent.be (T.-T. Vo).

https://doi.org/10.1016/j.jclinepi.2019.10.001

What is new?

Key findings?

- Mediation analysis (MA) is a very common type of statistical analysis in psychology, sociology, epidemiology, and medicine. This type of analysis aims to discover pathways and mechanisms by which an exposure may affect an outcome. This review describes how MA was conducted and reported in recent randomized controlled trials.
- Reported MAs were mostly based on a traditional approach. The counterfactual-based approach, which generalizes the traditional approaches to nonlinear models and linear models that include interactions, was rarely used in recent clinical research practice.
- Half of the studies considered adjusting for the mediator-outcome (M-OC) confounders. The set of M-OC confounders adjusted for often only consisted of the baseline value of the outcome and of the mediators.

What this adds to what is known?

- The conduct and reporting of MA are quite heterogeneous. This might be related to the complexity of the methodological literature on mediation and the lack of practical consensus guidelines for the conduct and reporting of MA.

What is the implication and what should change now?

- Future MAs could benefit from a consensus-based planning, conduct, and reporting of MA guideline. Hands-on tutorials would be useful to facilitate the use of counterfactual-based mediation techniques for the analysis of dichotomous and survival endpoints, common in RCTs.

prescribing intensive medication therapies to manage hypertension in practice [6]. The MA, therefore, is useful to improve patients' clinical outcomes.

The most traditional approaches of mediation date back to the 1980s and were inspired by the popular articles of Judd and Kenny [7] and of Baron and Kenny [8] in psychological research. These first seminal works on mediation proposed a series of tests of links in the causal chain to assess the presence of an indirect effect. The proposed tests are conservative and have limited statistical power because of the unnecessary requirement of a nonzero ITT effect to investigate mediation [3,9,10]. Several extensions have been made to overcome this limitation and to provide an effect-size measure for the so-called direct and indirect effects [3,9,10]. In some of these, the direct effect is estimated as the residual association between outcome and treatment after regression adjustment for the mediators [3,4,9]. Its difference with the ITT effect is then viewed as the indirect effect of treatment. In other extensions, the indirect effect (with respect to a single mediator) is calculated as the product of the ITT effect on the mediator times the residual association between outcome and mediator after regression adjustment for treatment assignment [3,4,9,11]. These so-called difference- and product-of-coefficient methods are equivalent and justified when both models for the outcome and mediator are linear with no interaction. They differ and raise validity concerns when one or both of these models is/are nonlinear [4,5]. For instance, when a logistic outcome model is used for a binary endpoint, then the difference-of-coefficient approach has a systematic tendency to find indirect effects in settings where treatment has no effect on mediator; the product-of-coefficient approach instead delivers direct and indirect effects which may not add up to the total causal effect, thereby failing to provide an effect decomposition [4,5]. To accommodate this, a counterfactual-based framework to MA has recently been introduced, which has the aforementioned approaches as special cases [1,4,5,12]. The literature on this framework is often of a more technical nature, partly because of its focus on nonlinear models and partly because it makes more explicit the unverifiable assumptions on which an MA relies. For instance, when there are multiple mediators, the counterfactual-based approaches have made clear that the effects along specific sequences of mediators are not identified without making strong biological and modeling assumptions, although such effects may be seemingly simple to calculate in the traditional approaches [4,13]. The counterfactual approaches have moreover emphasized the need to adjust for mediator-mediator (M-M) confounders in the analysis [14] (although it should be noted that the estimation of certain pathways in the traditional approaches and specific counterfactual extensions thereof is immune to such confounding) [15]. Regardless of the approach used, the correct temporal order between the mediators and the outcome, as well as adjustment for mediator-outcome (M-OC) confounders, is of critical importance for a valid mediation finding, even when based on data from an randomized controlled trial (RCT) [7,10,14].

In this study, we aimed to (i) describe how MA was conducted and reported in recent RCTs and to (ii) propose recommendations for the planning, conduct, and reporting of MAs in future RCTs.

2. Methods

2.1. Study design

We conducted a methodological systematic review and followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [16]. The

protocol was not registered in The International Prospective Register of Systematic Reviews as this review does not contain direct health-related outcomes [17].

2.2. Information sources and search strategy

As the primary aim of this review was to describe the current practice in conducting and reporting MA in clinical research, we restricted the search strategy to only include recent reports published on MEDLINE (accessed through PubMed) over the last 2 years (i.e., from January 1, 2017, to December 1, 2018, search date: December 11, 2018). The full search strategy is presented in Appendix 1.

2.3. Eligibility criteria

We included all reports of RCTs and the secondary analyses of previously published RCTs that conducted a MA to investigate the role of one or more mediators in explaining the pathway from treatment to a specific health outcome. We excluded non-English publications, articles for which full texts were unavailable, and protocols and RCTs in which an MA was conducted, but not to explain the effectiveness of the randomly assigned treatment. For instance, studies were excluded if their data were used to investigate the mediated effect of an exposure that was not the randomized treatment.

2.4. Study selection

We exported the references retrieved from the search into an MS Excel document (Microsoft Corp). Two reviewers (T.-T.V. and C.S.) independently screened the titles and abstracts of all retrieved references to identify the eligible reports. The full-text copies of potentially eligible reports were also obtained and independently examined for further assessment if needed. In the case of disagreement, consensus was determined by a discussion between the two reviewers. The result of this process was reported through a PRISMA flowchart [16].

2.5. Data extraction

A data-extraction form was designed, pilot-tested, and refined by a reviewer (T.-T.V.) to extract the following information from the eligible reports: (i) the number of mediators investigated; (ii) the method used to assess the indirect treatment effect through the mediators; (iii) the methodological characteristics of the reported MA; and (iv) the discussion of authors regarding the assumptions underlying the method being used (see Appendix 2 for the data-extraction form).

For item 2 in the aforementioned list, we classified the methods used to assess mediation into two groups, namely (i) traditional approaches (including the Baron and Kenny's framework, the difference- and product-of-coefficient approaches) and (ii) counterfactual-based approaches (Table 1). When multiple mediators were assessed, we further reported whether the different mediators were included in the analysis in a parallel or serial manner (Table 1). If a parallel MA was considered, we investigated whether the (parallel) mediators were assessed simultaneously in the same model or separately in different models. We made this distinction because the use of separate MAs, each involving a single mediator, is arguably invalid (see further discussion in Appendix 3 and [14]). The use of different approaches was also stratified based on the outcome type (continuous, binary, time-to-event, or others) and the outcome model used.

For item 3, we determined whether the eligible studies (i) reported a sample size calculation for the MA or commented on the impact of the sample size on the MA results, (ii) conducted a complete-case analysis or considered an alternative approach to handle the missing data, (iii) included or evaluated the potential of treatment-mediator interaction(s), (iv) considered confounding adjustment for the M-M and M-OC relationships, (v) considered the statistical significance of the ITT treatment effect as a condition to investigate mediation, (vi) conducted sensitivity analysis to assess the impact of the mediators' measurement error, and (vii) reported the goodness of fit (i.e., statistics that describe how well a model fits the set of observations) of the mediator(s) and outcome models. For item 4, when the authors reported confounding adjustment but did not clearly state whether it was for the mediator or for the outcome model, we assumed that the authors at least considered covariate adjustment in the outcome model. The latter is desirable because adjustment for covariates in the mediator model is less important when treatment is randomized.

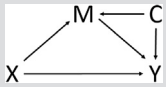
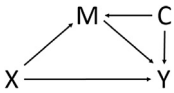
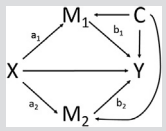
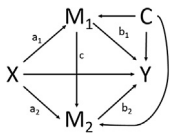
For item 4, we determined whether authors discussed the assumptions that underlined the MA and which assumptions were discussed. We extracted the exact assumptions being discussed using open-ended questions. We did not assess the interpretation of the clinical impact of the mediation findings. For instance, when a statistically significant mediation finding was found, the authors often discussed how such finding would be useful in medical practice (e.g., in patient management). We did not assess such discussion.

When several MAs were conducted for several outcomes of interest, we focused only on the analysis of the primary outcome. A second reviewer (C.S.) independently double-checked 25% of the extracted articles. In case of disagreement, consensus was determined by a discussion between the two reviewers (T.-T.V. and C.S.).

2.6. Data synthesis

Categorical data were summarized using frequencies and percentages. Continuous data were summarized using median and interquartile range. We reported the overall findings and stratified them based on the number of mediators investigated (i.e., single and multiple mediators). Data were analyzed using MS Excel 2010.

Table 1. Classification of mediation analysis methodology in clinical research

Methods	Basic properties
Traditional approaches[a] Single mediator  	• Different regression models are specified for the mediator and the outcome. For instance: $E(Y\|X) = c_0 + cX + c_1C$ $E(M\|X) = a_0 + aX + aC$ $E(Y\|X, M, C) = b_0 + b_1X + b_2C + bM$ • In the BK approach, a series of test of links in a causal chain is specified to assess the null hypothesis of no indirect effect (IE). No estimators of the direct effect (DE) and IE are proposed. • In other traditional approaches, the IE of X on Y through M is quantified as the product ab. The DE of X on Y is calculated as $c' = c - ab$ [3,7,8,10]. These identities hold when all models are linear without treatment-mediator interaction (see appendix 5 and [4,18,19]). • The difference-of-coefficient approach assesses the null hypothesis of no IE by using statistics linked to the difference $c - c'$. The product-of-coefficient approach assesses the null hypothesis by using statistics linked to the product ab. The underlying assumptions and the performance of these tests are discussed elsewhere [9].
Counterfactual approaches[b] Single mediator  	• Different, possibly nonlinear, regression models are specified for the mediator and the outcome. • The IE of X on Y through M is quantified by using the aforementioned models to simulate (via Monte Carlo simulation) how a change in X would affect M for each individual, how this change would in turn affect Y, and then averaging these results [1,18]. • The DE of X on Y is quantified by using the aforementioned models to simulate (via Monte Carlo simulation) how Y would change for each individual if X were changed, but M were fixed at the (simulated) level it would take if X took on some reference level, and then averaging these results [1,18].
Parallel mediators  	• Using a regression model for Y which includes X, C, and all mediators, and separate models for each mediator which include X and C. • In the traditional approaches, the individual IE for mediator i is a_ib_i and the total IE for a model including 2 mediators is $a_1b_1 + a_2b_2$, with b_i being the effect of mediator i on outcome and a_i the effect of exposure on mediator i. In the counterfactual approaches, Monte Carlo simulation is used [3,14,15]. • These approaches can correctly infer the effect along the combination of all pathways from X to M_i (along all possible paths) and further from M_i to Y (directly), even when the mediators are correlated, but not when separate analyses are considered with one mediator at a time [14].
Serial or sequential mediation  	• Assuming a causal relationship from one mediator to the other. • In the traditional approaches, IEs through different pathways are quantified as different product of coefficients, e.g., the IE through M1 then M2 is a_1cb_2. In the counterfactual approaches, Monte Carlo simulation is sequentially used with respect to different blocs of mediators [3,14]. • These approaches are sensitive to misspecification of the causal order between mediators and to the presence of unmeasured common causes of the mediators. • The traditional approaches have the drawback of enabling one to calculate IEs along all possible paths (e.g., through M_1 then M_2), even though some necessitate overly strong assumptions [13,15].

Abbreviations: X, the randomized treatment; Y, the outcome of interest; M and M_i, the potential mediator(s); C, the confounders of mediator-mediator/outcome relationship; IE, indirect effect; BK, Baron and Kenny.

[a] Including the Baron and Kenny framework and the product- and difference-of-coefficient approaches.

[b] Other causal approaches, e.g., the stochastic mediation approaches [20], have not been reviewed here.

2.7. Recommendations for future MAs

Based on the results of this systematic review, we proposed practical recommendations for the planning, conduct, and reporting of future MAs in clinical research. We consulted other textbooks and methodological reviews on MAs to assure the completeness of these recommendations [4,18,19,21–23].

3. Results

3.1. General characteristics of eligible studies

The PRISMA flow diagram summarizing the screening process is given in Fig. 1. Of 197 references identified, 99 reports were excluded for the following reasons: not an RCT ($n = 39$), no MA conducted ($n = 41$), trial protocol ($n = 8$), conducted MA not for explaining the effect of the randomized treatment ($n = 10$), non-English article ($n = 1$) (Appendix 3).

Among the 98 eligible studies, 48.0% ($n = 47$) were RCT reports and 52.0% ($n = 51$) were secondary analyses of previously published RCTs (Appendix 4). In 68.4% of studies ($n = 67$), multiple mediators were investigated. The outcome and mediator being assessed were continuous in 84.7% ($n = 83$) and 92.9% ($n = 91$) of studies, respectively. In 53.1% of studies ($n = 52$), the mediators and outcome were measured at the same time point, which might carry an additional risk of bias because of the lack

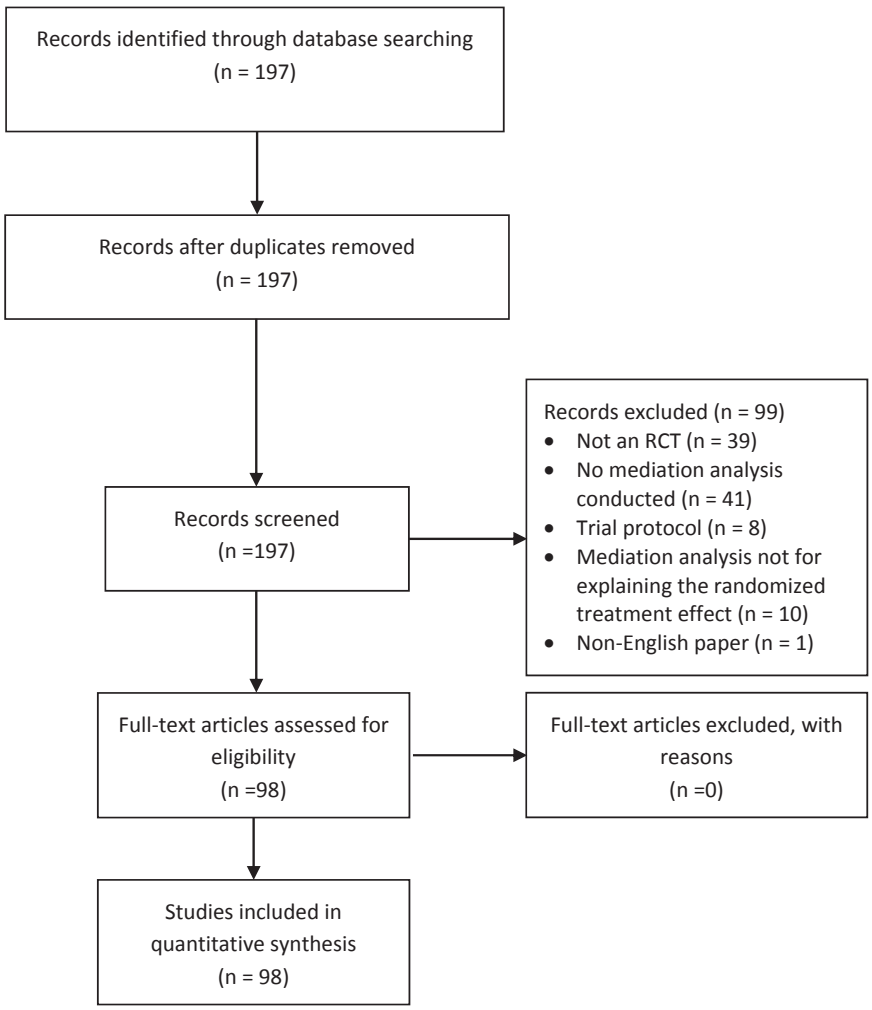


Fig. 1. – Study selection PRISMA Flowchart. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomized controlled trial.

of temporal order of the mediator(s) and the outcome [7,8,10,24,25]. Other characteristics of the eligible reports are summarized in Table 2. In data extraction, the Kappa coefficient equaled 0.86 (Appendix 2).

3.2. Methods used to investigate mediation

Among studies assessing a single mediator, the traditional approaches were applied in 90.3% of cases (n = 28). Counterfactual-based MA approaches, in contrast, were considered in two studies (6.5%). In one study (3.2%), both frameworks were used.

The traditional approaches were applied in 98.5% of studies assessing multiple mediators (n = 66). Among these, 88.1% (n = 59) considered a parallel MA, with 37.3% (n = 25) considering one common analysis for all mediators, 38.8% (n = 26) considering separate analyses for separate mediators, and 11.9% (n = 8) considering a two-step approach. In the two-step approach, the mediators were first separately analyzed. The ones showing statistical significance for the indirect effect in this step (or all mediators) were then taken forward to the second step, where one common (final) model was fitted (Table 3).

The traditional approaches were frequently used across all types of outcome. Among 11 studies with a binary or time-to-event outcome, seven studies (63.6%) applied the traditional approaches based on a logistic or Cox regression model (Table 3).

3.3. Methodological characteristics of the reported MA

Only 4.1% of studies (n = 4) reported on the evaluation of potential interaction between the treatment and the mediator(s) on the outcome (Table 4). Two studies included the interaction in the final MA and two others did not because of the lack of statistical significance.

Sample size calculation for MA was considered in 4.1% of studies (n = 4). A complete-case analysis was conducted by 50.7% of studies (n = 34), with 19.4% (n = 13) considering some covariate adjustment to correct for imbalance between treatment groups because of missing data.

Table 2. Characteristics of the eligible studies

Characteristics	Statistics
Study design, *n* (%)	
Original research report of an RCT	47 (48.0)
Reporting information on protocol registration	19 (19.4)
Secondary analysis of a previous RCT	51 (52.0)
Number of randomized patients, median (IQR)	163 (100–362)
Type of intervention, *n* (%)	
Medication	7 (7.1)
Psychological, cognitive, behavioral intervention (including training, education, counseling with or without other components)	79 (80.6)
Simulations of different contexts	3 (3.1)
Decision aid booklets	2 (2.0)
Others	7 (7.1)
Target condition, *n* (%)	
Addiction and behavioral medicine	39 (39.8)
Mental health disorders (without other conditions)	29 (29.6)
Oncology	10 (10.2)
Diabetes and obesity	7 (7.1)
Musculoskeletal diseases	6 (6.1)
HIV and infectious diseases	3 (3.1)
Cardiovascular diseases	2 (2.0)
Others	2 (2.0)
Primary outcome, *n* (%)	
Behaviors and affects (e.g., self-esteem, alcohol consumption, decision making, social connection, pain intensity, and so forth)	57 (58.2)
Symptoms of mental disorders (stress, suicide, depression, and so forth)	23 (23.5)
Biological concentration	2 (2.0)
Cardiovascular risk	3 (3.1)
Physical functionality	2 (2.0)
Others	11 (11.2)
Number of mediators investigate, *n* (%)	2 (1–4)
Single mediator	31 (31.6)
Multiple mediator	67 (68.4)
Mediator type, *n* (%)	
Continuous variable[a]	91 (92.9)
Binary variable[a]	5 (5.1)
Ordinal variable[a]	1 (1.0)
Unclear	1 (1.0)
Outcome type, *n* (%)	
Continuous variable	83 (84.7)
Binary variable	8 (8.2)
Count variable	4 (4.1)
Survival variable	3 (3.1)
Dataset, *n* (%)	
Cross-sectional (mediator(s) and outcome measured at the same time)	52 (53.1)
Longitudinal (mediator(s) measured before outcome)	46 (46.9)

Abbreviations: RCT, randomized controlled trial; IQR, interquartile range; HIV, human immunodeficiency virus.

[a] Number of studies investigating each type of mediators (Mediators of different types, e.g., binary and continuous, were not evaluated in any studies.).

Adjusting for the M-M or M-OC confounders was considered by 57.1% of studies ($n = 56$). Such adjustment was less common among studies assessing a single mediator (48.4%, $n = 15$). Among studies considering M-M and M-OC adjustments ($n = 56$), 32.1% ($n = 18$) only adjusted for the baseline value of the outcome and/or mediator(s) of interest. The remaining 67.9% ($n = 38$) also adjusted for other baseline characteristics, with a total number of adjusted covariates (already including baseline levels of the mediator and/or outcome) ranging from 2 to 7 (IQR). No studies considered adjustment for postrandomization confounding factors of the M-OC association (Table 4).

Information on the goodness of fit of the models used in the MA was reported in 27.6% of studies ($n = 27$). In 11.2% of studies ($n = 11$), authors stated that a statistically significant ITT effect was considered as a prior condition to investigate mediation. Concerns over mediator(s)' measurement error due to the use of self-reported data were expressed in 35.7% of studies ($n = 35$), but no studies considered a sensitivity analysis to further assess these (Table 4).

3.4. Discussion on the MA

In 29.6% of the eligible studies ($n = 29$), the authors discussed the plausibility of different assumptions underlying the MA. The assumption of no unmeasured M-M and M-OC confounders was discussed in 9.2% ($n = 9$) of studies. The assumption of no treatment-induced confounders of the M-OC relationship was discussed in 2.0% of studies (i.e., two studies using counterfactual-based approaches). Concerns about the temporal order of the mediator(s) and the outcome were expressed in 19.4% ($n = 19$) of studies, with a reversed MA (i.e., switching the role of the outcome and the mediators in the analysis) conducted in 6.1% ($n = 6$) of studies. Concerns about potential misspecification of the mediator(s)/outcome models were expressed in 3.1% ($n = 3$) of studies (Table 4).

Recommendations on the planning, conduct, and reporting of future MAs are reported in Table 5.

Table 3. Framework used for the mediation analysis reported in the eligible studies

Item	Single mediator (*N* = 31)	Multiple mediators (*N* = 67)	Overall (*N* = 98)
Method used, *n* (%)			
Traditional approaches	28 (90.3)	-	-
Counterfactual-based approaches	2 (6.5)	-	-
Both traditional and causal approaches	1 (3.2)	-	-
Traditional approaches: parallel mediators	-	59 (88.1)[a]	-
One separate model for each mediator	-	26 (38.8)	-
One common model for all mediators	-	25 (37.3)	-
Two-step analysis[b]	-	8 (11.9)	-
Traditional approaches: serial mediators	-	9 (13.4)[a]	-
Counterfactual-based approaches, one model for each mediator	-	1 (1.5)	-
Method used, stratified by outcome, *n* (%)			
Continuous outcome	26 (83.9)	57 (85.1)	83 (84.7)
Traditional approaches	26 (83.9)	56 (83.6)	82 (83.7)
Counterfactual-based approaches	0 (0.0)	1 (1.5)	1 (1.2)
Binary outcome	3 (9.7)	5 (7.5)	8 (8.2)
Logistic regression, traditional approaches	0 (9.7)	4 (6.0)	4 (4.1)
Log-linear regression, traditional approaches	0 (9.7)	1 (1.5)	1 (1.2)
Logistic regression, counterfactual-based approaches	3 (9.7)	0 (0.0)	3 (3.1)
Time-to-event outcome	1 (3.2)[c]	2 (3.0)	3 (3.1)
Cox regression, traditional approaches	1 (3.2)[c]	2 (3.0)	3 (3.1)
Cox regression, counterfactual-based approaches	1 (3.2)[c]	0 (0.0)	1 (1.2)
Count outcome, log-linear regression, traditional approaches	1 (3.2)	3 (4.5)	4 (4.1)

Abbreviations: MA, mediation analysis; IE, indirect effect.

[a] Two studies considered both parallel and serial MA.

[b] One separate analysis for each mediator (step 1), then a common analysis for the mediators with significant IE in the previous step or all mediators (step 2).

[c] Same study with both approaches used.

4. Discussion

MA is a statistical technique that has become increasingly popular over the past decade [1–3]. In this methodological systematic review, we evaluated how MA was conducted and reported in recently published RCTs. Several remarks can be highlighted. First, our review confirmed that M-OC confounding adjustment was being considered in recent MAs of RCTs. However, the confounders taken into account were often only the baseline value of the outcome and of the mediator(s). Adjusting for a wider collection of confounders of the M-M and M-OC relationship, which may in particular include confounders that are themselves mediators, should be more encouraged in practice to strengthen the validity of the findings [19]. Besides, baseline covariate adjustment is additionally important whenever a complete-case analysis is adopted, as it may help to adjust for selection bias originating from imbalances between the treatment groups due to the exclusion of some patients from the analysis.

Regardless of the methods being used, MA is often based on quite strong assumptions [3,19]. These assumptions should be discussed to avoid any misinterpretation of the findings. As the page restriction policies of most journals may discourage researchers to do this, supplemental discussions should be encouraged in online Web appendices. Alternatively, journal editors should also consider providing more space for consideration of the assumptions described in the article.

Most MAs assessed in this review were only conducted secondarily after the end of the main trial. Planning MAs in advance (e.g., in the trial protocol) can help to handle some methodological challenges that remained unaddressed in most current studies, for example, the sample size calculation, the measurement of key confounders, the quantification of measurement error in the mediators, and the prevention of missing data in key (mediator) variables.

Finally, counterfactual mediation approaches were rarely used in recent clinical research practice. This is probably because these approaches have only become available within the last 20 years, with the literature often being highly technical. The relatively small number of articles that reported on MAs of dichotomous and survival endpoints, in spite of these endpoints being common in RCTs, may also be due to a lack of familiarity with (counterfactual) mediation approaches for such endpoints. More detailed instructions (e.g., via hand-on tutorials) are hence needed to further promote the use of these approaches as they

Table 4. Characteristics of the mediation analyses reported in the eligible studies

Item	Single mediator (*N* = 31)	Multiple mediators (*N* = 67)	Overall (*N* = 98)
Methodological characteristics			
SS calculation, *n* (%)			
Calculating an SS for the MA	0 (0.0)	4 (6.0)	4 (4.1)
Only discussing the impact of SS on the findings	11 (35.5)	26 (38.8)	37 (37.8)
No information	20 (64.5)	37 (55.2)	57 (58.2)
Handling missing data			
ITT analysis, *n* (%)	9 (29.0)	26 (38.8)	35 (35.7)
Multiple imputation	2 (6.5)	4 (6.0)	6 (6.1)
Single imputation	2 (6.5)	3 (4.5)	5 (5.1)
Last observation carried forward	0 (0.0)	3 (4.5)	3 (3.1)
Full information maximum likelihood	5 (16.1)	13 (19.4)	18 (18.4)
Others	0 (0.0)	3 (4.5)	3 (3.1)
Complete-case analysis, *n* (%)	20 (64.5)	34 (50.7)	54 (55.1)
Percentage of data excluded, median (IQR)	15 (11–37)	12 (21–36)	19 (11–36)
Not adjusting for T-M confounders, *n* (%)	13 (41.9)	21 (31.3)	34 (34.7)
Adjusting for T-M confounders, *n* (%)	7 (22.6)	13 (19.4)	20 (20.4)
Number of adjusted covariates, median (IQR)	4 (3.5–5.5)	4 (2–5)	4 (2–5)
No missing data or nor reported, *n* (%)	2 (6.5)	7 (10.4)	9 (9.2)
Considering statistically significant ITT treatment effect as a condition to investigate mediation, *n* (%)	3 (9.7)	8 (11.9)	11 (11.2)
Evaluating treatment-mediator(s) interaction(s), *n* (%)	1 (3.2)	3 (4.5)	4 (4.1)
M-M and M-OC confounding adjustment			
No adjustment or not reported, *n* (%)	16 (51.6)	26 (38.8)	42 (42.9)
Only for the baseline value of the outcome and/or mediator(s), *n* (%)	2 (6.5)	16 (23.9)	18 (18.4)
Also for other baseline covariates, *n* (%)	13 (41.9)	25 (37.3)	38 (38.8)
Number of covariates adjusted, *n* (%)[a]	4 (3–7)	4 (2–7)	4 (2–7)
Reporting the GOF of the models, *n* (%)	8 (25.8)	19 (28.4)	27 (27.6)
Discussing the risk of the mediator(s)' measurement error	8 (25.8)	27 (40.3)	35 (35.7)
Discussion of the underlying assumptions, *n* (%)			
No discussion	19 (61.3)	50 (74.6)	69 (70.4)
Temporal order of mediator(s) and outcome[b]	8 (25.8)	11 (16.4)	19 (19.4)
Unmeasured mediator(s)-outcome confounders	4 (12.9)	5 (7.5)	9 (9.2)
Treatment-induced mediator(s)-outcome confounders	2 (6.5)	0 (0.0)	2 (2.0)
Model misspecification	2 (6.5)	1 (1.5)	3 (3.1)

Abbreviations: SS, sample size; ITT, intent-to-treat; T-M, treatment-mediator; M-M, mediator-mediator; M-OC, mediator-outcome; GOF, goodness of fit; IQR, interquartile range; MA, mediation analysis.

[a] The total number of variables adjusted for in the analysis, already including baseline values of the outcome/mediator(s).

[b] Including the studies considering a reversed mediation analysis.

generalize the traditional framework to nonlinear models and linear models that include treatment-mediator interactions and moreover enable a more refined confounding control.

Given the heterogeneity in conduct and reporting MA in clinical research, there is a strong need for a valid, consensus-based reporting guideline for MA as is commonly done in other research fields. We hope that the practical recommendations in Table 5 will serve as a first step toward such endeavor.

Our findings were in line with results of other MA reviews, which also outlined the heterogeneity in conducting and reporting MAs in practice [21–23]. However, while the previous reviews either assessed systematic reviews of MAs [21] or focused on studies using a particular approach [22] or a particular type of outcome [23], we here attempted to give a broader overview on the use of different approaches in different contexts.

Our study suffered from some limitations. First, we only considered MEDLINE for the search of eligible trials. This review is hence not extensive and representative for all trials conducted in the searching period. However, our primary aim is to provide a description of recent practice and, hence, to provide insights into important conduct

Table 5. Recommendations on the conduct and reporting of mediation analysis in clinical research

1. Planning
1.1 Whenever possible, plan mediation analyses a priori in the trial protocol to strengthen the validity of the findings.
1.2 Decide on the choice of mediators based on the clinical rationale underlying the mechanisms through which the treatment affects the outcome, or based on independent data.
1.3 Plan the collection of prerandomization and postrandomization confounders of the M-M and M-OC relationships. Foresee if any of these confounders is treatment-induced (e.g., collected after the onset of treatment and therefore possibly affected by treatment).
1.4 Measure the mediators before the outcome, and preferably repeatedly, to assure the causal interpretation of the findings [7,8,10,24,25].
1.5 Develop insight by constructing the causal diagram underlying the causal relationship of the treatment, mediator(s), and outcome. For a practical example, see [4,19,26].
1.6 Estimate the sample size for the MA. For detailed instructions, see [10,27].
1.7 Do not make the conduct of a mediation analysis dependent on whether a statistically significant ITT treatment effect is found. The ITT effect may be null, even when there is an important indirect effect that is of opposite sign to the remaining direct effect [10,28].
2. Conduct
2.1 Use the multiple imputation approach (or other valid approaches) to handle missing data. When a complete-case analysis is used, adjust for baseline covariates that are differentially distributed between the treatment groups, as well as between the compleated cases and the non-completed cases in the study. A sensitivity analysis can also be carried out to assess the impact of different approaches on the findings [10].
2.2 Do not consider separate analyses for separate mediators, even in a parallel MA. Evidence supporting the so-called two-step approach (which only includes in the final analysis the mediators that are statistically significant when being assessed individually) is also lacking.
2.3 Choose an appropriate framework for the analysis – The Baron and Kenny approach is conservative, has low statistical power than other approaches, and does not propose a proper measure of the direct and indirect effect [9,11]. – The product- and difference-of-coefficient approaches and the counterfactual mediation approaches based on natural direct and indirect effect coincide in the case of linear models without interactions for mediators and outcome. – Counterfactual-based mediation approaches are more encouraged when treatment-mediator interactions present or when the analysis involves binary or time-to-event endpoints that necessitate nonlinear models (e.g., logistic and Cox models). [4,5,29]
2.4 Assess potential interaction(s) between treatment and confounding factors, treatment and mediator, mediator and mediator in the mediator and outcome models. Evaluate the goodness-of-fit of each model.
2.5 Adjust for M-M and M-OC confounders. This is encouraged in both approaches, even in randomized trials and in the absence of missing data. Indeed, the mediator(s) is/are not manipulated so that M-M and M-OC confounders may present [7,10,14,30].
2.6 If possible, perform sensitivity analysis to assess sensitivity of the results to: – The assumption of no unmeasured M-M/OC confounders [31]. – Potential measurement errors of the mediators [19].
2.7 Use apt strategies when some of the M-M/OC confounders are potentially affected by the treatment (e.g., by considering these confounders as mediators themselves [18]).
3. Reporting
3.1 Report the approaches used for mediation and provide a causal diagram that underlies the analysis.
3.2 Report the sample size calculation, the actual sample size of the MA, and how the missing data are handled.
3.3 Report all confounders considered and adjusted for in the analysis.
3.4 Report the model-building procedure and the final form of all models used in the analysis. Report the goodness-of-fit of these models.
3.5 Report the point estimates and the confidence intervals of the different direct, indirect, and total treatment effects.
3.6 Report the methods and results of all sensitivity and other additional analyses (in the main article or appendices).
3.7 Discuss the validity of all causal assumptions underlying the analysis (in the main article or appendices) [3,19].

Abbreviations: ITT, intent-to-treat; M-M, mediator-mediator; M-OC, mediator-outcome; MA, mediation analysis.

and reporting issues that need to be addressed in (near) future practice. Second, the data extraction in this review was only partially double-checked (25%) by a second reviewer, which might result in potential mistakes. As the level of agreement among the two reviewers was relatively high (86%), the risk of incorrectly extracted data was relatively small. Third, we only considered MAs conducted in RCTs but not in observational studies, where the concerns about potential confounding or the temporal order of mediators and outcome are even greater. Fourth, some other aspects of the MA were not evaluated in this review, including the theoretical justification for the mediators selected and assessed in the included studies, the implementation of interventions, and the measurement of outcome. In every article, the authors provided explanation on how and why they focused on the assessed mediators. However, assessing

the appropriateness of this rationale was challenging as it required knowledge on the topic of interest. Also, we did not evaluate the implementation of interventions or the outcome assessment as these are important issues in trials, but they are not specific to MA. Finally, as most eligible studies in this review considered linear models and did not assess or report assessments regarding treatment-mediator interaction, we could not explore whether the alternative use of the counterfactual-based approaches would have been more appropriate.

5. Conclusion

The conduct and reporting of MA in recent trials remain quite heterogeneous. A valid, consensus-based methodological guideline is needed to enhance the planning, conduct, and reporting quality of MA in clinical research.

CRediT authorship contribution statement

Tat-Thang Vo: Conceptualization, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Cecilia Superchi:** Conceptualization, Data curation, Writing - original draft, Writing - review & editing. **Isabelle Boutron:** Writing - review & editing. **Stijn Vansteelandt:** Conceptualization, Writing - review & editing.

Acknowledgments

The authors would like to thank the members of the Methods in Research on Research (MiRoR) Project (http://miror-ejd.eu/) and Marie Sklodowska-Curie Actions for their support. They also convey their sincere thanks to Alice Biggane, Camila Olarte Parra, David Blanco, Efstathia Gkioni, Linda Nyanchoka, Melissa Sharp, Thu Van Nguyen, and the four ''anonymous'' reviewers of the *Journal of Clinical Epidemiology* for their valuable comments on a previous version of the manuscript.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2019.10.001.

References

[1] Lange T, Hansen KW, Sørensen R, Galatius S. Applied mediation analyses: a review and tutorial. Epidemiol Health 2017;39:e2017035.
[2] Zhang Z, Zheng C, Kim C, Van Poucke S, Lin S, Lan P. Causal mediation analysis in the context of clinical research. Ann Transl Med 2016;4(21):425.
[3] MacKinnon DPIn: Introduction to statistical mediation analysis, x. New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates; 2008:477.
[4] VanderWeele TJ. Mediation analysis: a practitioner's guide. Annu Rev Public Health 2016;37:17−32.
[5] VanderWeele T, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. Stat Interface 2009;2:457−68.
[6] Stensrud MJ, Strohmaier S. Diastolic hypotension due to intensive blood pressure therapy: is it harmful? Atherosclerosis 2017;265:29−34.
[7] Judd CM, Kenny DA. Process analysis: estimating mediation in treatment evaluations. Eval Rev 1981;5(5):602−19.
[8] 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):1173−82.
[9] MacKinnon DP, Lockwood CM, Hoffman JM, West SG, Sheets V. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods 2002;7(1):83−104.
[10] Fairchild AJ, McDaniel HL. Best (but oft-forgotten) practices: mediation analysis12. Am J Clin Nutr 2017;105:1259−71.
[11] MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. Annu Rev Psychol 2007;58(1):593−614.
[12] Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. Am J Epidemiol 2012;176:190−5.
[13] Daniel RM, De Stavola BL, Cousens SN, Vansteelandt S. Causal mediation analysis with multiple mediators. Biometrics 2015;71:1−14.
[14] VanderWeele TJ, Vansteelandt S. Mediation analysis with multiple mediators. Epidemiol Methods 2014;2(1):95−115.
[15] Vansteelandt S, Daniel RM. Interventional effects for mediation analysis with multiple mediators. Epidemiology 2017;28(2):258−65.
[16] Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097.
[17] mh329. PROSPERO: international prospective register of systematic reviews — university of Leicester [Internet]. Available at: https://www2.le.ac.uk/library/find/databases/p/Prospero. Accessed April 4, 2019.
[18] VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. Epidemiol Camb Mass 2014;25(2):300−6.
[19] VanderWeele T. Explanation in Causal Inference: Methods for Mediation and Interaction. Oxford, New York: Oxford University Press; 2015:728.
[20] Naimi AI, Moodie EEM, Auger N, Kaufman JS. Stochastic mediation contrasts in epidemiologic research: interpregnancy interval and the educational disparity in preterm delivery. Am J Epidemiol 2014;180:436−45.
[21] Cashin AG, Lee H, Lamb SE, Hopewell S, Mansell G, Williams CM, et al. An overview of systematic reviews found suboptimal reporting and methodological limitations of mediation studies investigating causal mechanisms. J Clin Epidemiol 2019;111:60−68.e1.
[22] Liu S-H, Ulbricht CM, Chrysanthopoulou SA, Lapane KL. Implementation and reporting of causal mediation analysis in 2015: a systematic review in epidemiological studies. BMC Res Notes 2016;9:354.
[23] Lapointe-Shaw L, Bouck Z, Howell NA, Lange T, Orchanian-Cheff A, Austin PC, et al. Mediation analysis with a time-to-event outcome: a review of use and reporting in healthcare research. BMC Med Res Methodol 2018;18:118.
[24] Lemmens LHJM, Müller VNLS, Arntz A, Huibers MJH. Mechanisms of change in psychotherapy for depression: an empirical update and evaluation of research aimed at identifying psychological mediators. Clin Psychol Rev 2016;50:95−107.
[25] Hollon SD, DeRubeis RJ. Mediating the effects of cognitive therapy for depression. Cogn Behav Ther 2009;38(Suppl 1):43−7.
[26] Staplin N, Herrington WG, Judge PK, Reith CA, Haynes R, Landray MJ, et al. Use of causal diagrams to inform the design and interpretation of observational studies: an example from the study of heart and renal protection (SHARP). Clin J Am Soc Nephrol 2017;12(3):546−52.

[27] Liu X, Wang L. Sample size planning for detecting mediation effects: a power analysis procedure considering uncertainty in effect size estimates. Multivariate Behav Res 20191−18.

[28] Rucker DD, Preacher KJ, Tormala ZL, Petty RE. Mediation analysis in social psychology: current practices and new recommendations. Soc Personal Psychol Compass 2011;5(6):359−71.

[29] 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(2):137−50.

[30] Cole SR, Hernán MA. Fallibility in estimating direct effects. Int J Epidemiol 2002;31:163−5.

[31] VanderWeele TJ, Yasutaka C. Sensitivity analysis for direct and indirect effects in the presence of exposure-induced mediator-outcome confounders. Epidemiol Biostat Public Health 2014; 11(2):e9027.