

Understanding Counterfactual-Based Mediation Analysis Approaches and Their Differences

Stijn Vansteelandt

Current approaches for mediation analysis based on natural direct and indirect effects differ primarily in terms of the statistical models on which they rely¹:

1. A model for the expected outcome Y , given mediator, exposure, and baseline covariates (confounders) W .
2. A model for the distribution of the mediator M , given exposure and confounders W .
3. A model for the distribution of the exposure X , given confounders W .

Available approaches require correct specification for two of these three models¹; the proposal of Albert² in this issue of *Epidemiology* avoids reliance on a model for the mediator distribution in view of the usual difficulties in specifying it.

THE MEDIATION FORMULA

The mediation formula³ is pivotal to nearly all methods for the estimation of natural direct and indirect effects. This formula prescribes estimating $E\{Y(x, M(x'))\}$ by standardizing predictions from the outcome model corresponding to exposure level x , relative to the mediator distribution corresponding to exposure level x' . This formula has inspired the development of maximum likelihood estimators, obtainable by substituting the outcome and mediator distributions by their maximum likelihood estimators under suitable parametric models⁴⁻⁶; some traditional mediation analysis approaches⁷ can be viewed as approximations^{4,5} thereof. Provided correct models for the outcome mean and mediator distribution, it follows that direct application⁴⁻⁶ of the mediation formula (based on maximum likelihood estimators) delivers natural direct and indirect effect estimators that are at least as precise as those obtained through alternative approaches that avoid specification of either the outcome mean or the mediator distribution. The simulation study in Albert² shows that the loss of precision can be sizeable.

As Albert² correctly recognizes, faultless model specification may however be a thorny issue. When the mediator is strongly associated with exposure or confounders, then misspecification of the mediator's effect on the outcome may be difficult to diagnose, and extrapolation bias is likely. The same is true for the exposure's effect on the outcome when the exposure is strongly associated with covariates. Moreover, models for the mediator distribution can be difficult to postulate. This is especially so when the mediator is continuous; its distribution is then not entirely described by its mean. However, note that it is not always necessary to correctly specify the entire mediator distribution for direct application of the mediation formula to give valid results. For instance, it suffices to correctly specify the mediator's expectation when the outcome model is linear⁴ in the mediator; for natural indirect effects, this is also the case when the outcome is binary, rare, and modeled through logistic regression.^{5,8} Bearing this and the aforementioned potential for precision loss in mind, further simulation studies seem warranted under degrees of model misspecification that are more

From the Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan, Belgium.

Correspondence: Stijn Vansteelandt, Ghent University, Department of Applied Mathematics and Computer Science, Krijgslaan 281, S9, 9000 Gent, Belgium, E-mail: Stijn.Vansteelandt@ugent.be.

Copyright © 2012 by Lippincott Williams & Wilkins
 ISSN: 1044-3983/11/2306-0889
 DOI: 10.1097/EDE.0b013e31826d0f6f

realistic (in the sense of being difficult to diagnose or amend) than those described by Albert.²

ALTERNATIVE APPROACHES FOR MEDIATION ANALYSIS

Alternative approaches for mediation analysis substitute an exposure model for either the outcome or mediator model. These approaches, including that of Albert,² are indicated primarily when the investigator has relatively greater confidence in the correctness of the exposure model, as is typically the case when the exposure is randomly assigned. Such a priori information is ignored by maximum likelihood approaches because it does not help to increase precision, even though it partially insulates the results from model misspecification bias. These alternative approaches may be less attractive than direct application of the mediation formula when the exposure is continuous because this can make the modeling of the exposure distribution relatively more cumbersome; in what follows, as in the article by Albert,² we will therefore concentrate on dichotomous exposures.

A first class of alternative approaches avoids reliance on a model for the mediator distribution. Here, progress is made by noting that the counterfactual $M(x')$ equals the observed value M within the subgroup of individuals with exposure level x' , so that $Y(x, M(x'))$ equals $Y(x, M)$. The expectation $E\{Y(x, M(x'))|X = x'\}$ can therefore be calculated by predicting $Y(x, M)$ on the basis of an outcome model with X set to x and with M and W set to their observed values, and then averaging this within this subgroup. Inverse probability weighting (by the reciprocal of $P(X = x'|W)$) can be used to account for the selective nature of subjects with $X = x'$, and thus to transport the results to the general population, ie, to calculate $E\{Y(x, M(x'))\}$. This forms the basis of the developments of Albert² and Vansteelandt et al.⁹ These approaches are primarily indicated when the investigator has a priori knowledge about the exposure distribution, or when direct application of the mediator formula requires not only correct specification of the mediator's expectation but also its entire distribution.

A second class of alternative approaches avoids reliance on a model for the outcome mean. These approaches are indicated primarily with concern for extrapolation bias in the outcome regression model. Here, $E\{Y(x, M(x'))\}$ is estimated as the sample average of the outcome in subjects with $X = x$, but weighting by the reciprocal of $P(X = x|W)$ to account for the selective nature of those subjects, and additionally by $P(M|X = x', W)/P(M|X = x, W)$ to standardize the results to the mediator distribution at exposure level x' (rather than the observed level x). That is, $E\{Y(x, M(x'))\}$ is estimated as:

$$\sum_i I(X_i = x) w_i Y_i / \sum_i I(X_i = x) w_i \tag{1}$$

using weights

$$w_i = P(M_i | X_i = x', W_i) / \{P(X_i = x | W_i) P(M_i | X_i = x, W_i)\}. \tag{2}$$

This forms the basis of the developments of Hong¹⁰ and Lange et al.¹¹ A limitation is that this approach can suffer more weight instability⁹ (ie, w_i can be highly variable individuals i) and that correct specification of the mediator distribution can be more demanding than correct specification of the outcome mean.

Tchetgen Tchetgen and Schpitser¹ develop a so-called triply robust approach, which essentially combines the aforementioned three approaches. It provides valid results when two of the three working models for the exposure, mediator, and outcome are correctly specified but does not require the user to specify which two. This approach is promising, in that it lessens the concern about model misspecification bias more than the foregoing approaches and does not require the user to make a specific choice of working models to rely on. However, the inferior performance of related strategies in simulation studies⁹ indicates that further work is needed on how to best fit the exposure, mediator, and outcome models before routine application can be advised.

MARGINAL VERSUS CONDITIONAL EFFECTS

Although the focus of Albert² is on marginal or population-averaged effects, one may alternatively choose to focus on conditional natural direct and indirect effects^{4,5,9,12}: $E\{Y(x, M(x')) - Y(x', M(x')) | W\}$ and $E\{Y(x, M(x)) - Y(x, M(x')) | W\}$, where W includes all confounders. When W is discrete with few levels, these effects can be estimated by applying the above approaches with each stratum separately. In general, some form of modeling is required to allow for the borrowing of information across strata. This can, for instance, be performed through the so-called natural-effect models^{9,11} of the form:

$$E\{Y(x, M(x')) | W\} = \beta_0 + \beta_1 x + \beta_2 x' + \beta_3 W, \tag{3}$$

where β_1 and β_2 capture the natural direct and indirect effect of exposure on outcome. These can be estimated using a regression imputation⁹ approach that is closely related to the approach of Albert² in this issue of Epidemiology. In particular, the counterfactuals $Y(0, M(X))$ and $Y(1, M(X))$ can be predicted on the basis of the outcome model with X set to either 0 or 1, and with the mediator and covariates set to their observed values. Let Y^* denote these predictions and let X^* be an artificial exposure variable, which correspondingly assigns 0 or 1 to these predicted values. Model (3) can then be fitted through the corresponding standard regression model:

$$E(Y^* | X^*, W) = \beta_0 + \beta_1 X^* + \beta_2 X + \beta_3 W. \tag{4}$$

By controlling for confounders directly in the outcome regression model, one thus avoids the need for inverse weighting by the exposure distribution. Further advantages are that natural-effect models allow the analyst to study effect modification by covariates, and the borrowing of information across strata can deliver more powerful results. A possible limitation,

particularly germane to nonlinear models, is that the outcome model and the natural-effect model (3) may be mismatched.⁹ However, this may be less of a practical concern if one considers the natural-effect model as a convenient summary for reporting. Moreover, it can be shown that misspecification of linear natural-effect models does not bias tests of the null hypothesis of “no direct effect,” and when X is linear in W , it also does not bias tests of the null hypothesis of “no indirect effect.” Vansteelandt and Keiding¹³ provide further discussion on marginal versus conditional effects.

CONCLUSION

In summary, Albert² nicely combines regression mean imputation and inverse probability weighting ideas to infer natural direct and indirect effects. His proposal is indicated primarily when prior knowledge is available on the exposure distribution (eg, that the exposure is randomly assigned), or when direct application of the mediation formula requires a model not only for the mediator’s expectation but also for its entire distribution. His proposal might imply an important precision loss relative to direct application of the mediation formula or other more efficient estimation approaches,^{1,9} especially when the exposure is continuous or has strong predictors. In such settings, one can avoid inverse weighting by the exposure distribution and the need for modeling the mediator distribution, by focusing on conditional effects under the so-called natural-effect models.⁹

ABOUT THE AUTHOR

STIJN VANSTEELANDT is an Associate Professor of Statistics at Ghent University (Belgium) and Honorary Professor at the London School of Hygiene and Tropical

Medicine (UK). He has made contributions on causal inference methodology for mediation analysis, effect modification, instrumental variables analysis, time-varying confounding adjustment, and for the analysis of retrospective study designs.

REFERENCES

1. Tchetgen Tchetgen EJ, Schpitser I. Semiparametric Theory for Causal Mediation Analysis: efficiency bounds, multiple robustness, and sensitivity analysis. *Ann Statist*. In Press.
2. Albert JM. Mediator distribution-free mediation analysis for nonlinear models with confounding. *Epidemiology*. 2012;23:879–888.
3. Pearl J. Direct and indirect effects. In: Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence. Morgan Kaufmann, San Francisco, CA; 2001: 411–420.
4. VanderWeele TJ, Vansteelandt S. Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface*. 2009;2:457–462.
5. Vanderweele TJ, Vansteelandt S. Odds ratios for mediation analysis for a dichotomous outcome. *Am J Epidemiol*. 2010;172:1339–1348.
6. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods*. 2010;15:309–334.
7. MacKinnon D. *Introduction to Statistical Mediation Analysis*. New York: Lawrence Erlbaum Associates; 2002.
8. Tchetgen Tchetgen EJ. Formulae for causal mediation analysis in an odds ratio context without a normality assumption for the continuous mediator. *Technical Report, Harvard University*; 2012.
9. Vansteelandt S, Bekaert M, Lange T. Natural effect models for the estimation of direct and indirect effects. *Epidemiologic Methods*; 2012;1:131–158.
10. Hong G. Ratio of mediator probability weighting for estimating natural direct and indirect effects. In: *Proceedings of the American Statistical Association, Biometrics Section*. Alexandria, VA: American Statistical Association; 2010: 2401–2415.
11. Lange T, Vansteelandt S, Bekaert M. A simple unified approach for estimating natural direct and indirect effects. *Am J Epidemiol*. 2012;176:190–195.
12. Tchetgen Tchetgen EJ, Schpitser I. Semiparametric estimation of models for natural direct and indirect effects. *Technical Report, Harvard University*; 2011.
13. Vansteelandt S, Keiding N. Invited commentary: G-computation—lost in translation? *Am J Epidemiol*. 2011;173:739–742.