Published in final edited form as:

Multivariate Behav Res. 2015 ; 50(6): 732–733. doi:10.1080/00273171.2015.1120183.

Mediational Analysis in a Planned Missingness Data Design: Alternative Model Specifications and Power of the Mediated Effect

Amanda N. Baraldi

Oklahoma State University

Using a planned missing design, this Monte Carlo study considered a scenario in which the high cost of an ideal mediator, M_1 , limited sample size to the point of inadequate power to detect the mediated effect, but in which less expensive mediators, M_2 and M_3 , could be gathered on a larger sample. Simulated multivariate normal data were generated from the latent variable mediation model shown in Figure 1, Method 1. The following factors were manipulated: (a) missingness on M_1 (20, 50, 80%) under a missing completely at random mechanism, (b) latent mediator loadings λ_2 and λ_3 (.40, .60, .80), and (c) sample size ($N=100, 200, 500, 1000$). Mediated path parameters α and β were fully crossed with all levels of zero, small, medium, and large effect sizes. For all non-zero mediated effects, direct effect τ' equaled zero. Five designs (Figure 1) were employed to examine approaches to incorporating mediators into mediational analysis: Method 1 was a latent variable model where all three mediators were indicators of a latent mediation construct; Method 2 was an auxiliary variable model with M_1 as the mediator, and M_2 and M_3 as auxiliary variables; Method 3 was an auxiliary variable model with M_1 as the mediator and M_2 as a single auxiliary variable; Methods 4 and 5 were analyses incorporating only one mediator, M_1 , with Method 4 incorporating all available data via maximum likelihood estimation and Method 5 using listwise deletion. With the exception of Method 5, all methods used maximum likelihood estimation to address missing data.

Given negligible bias, the main outcome of interest was power to detect the mediated effect. Statistical power increased with increasing effect sizes, increasing sample sizes, and decreasing missing rates. Power was the greatest for analysis methods that included all three mediators, and power decreased with methods that included fewer variables. Across all design cells relative to the complete data condition, the latent variable model (Method 1) with 20% missingness on M_1 produced only 2.06% loss in power for the mediated effect; with 50% missingness, 6.02% loss; and 80% missingness, only 11.86% loss. Method 2 exhibited 20.72% power loss at 80% missingness, even though the total amount of data utilized in Methods 1 and 2 was the same. Methods 3 – 5, based on fewer variables, exhibited greater power loss. Compared to an average power loss of 11.55% across all levels of missingness for Method 1, average power losses for Methods 3, 4, and 5, were 23.87%,

29.35%, and 32.40%, respectively. The results suggest that planned missingness may permit incorporation of an expensive mediator at a feasible cost.

Acknowledgments

I thank my SMEP sponsor, Craig Enders, as well as Leona Aiken and David MacKinnon.

Method 1. Three indicator latent variable mediation model.

Method 3. Auxiliary variable mediation model with one auxiliary variable.

Method 2. Auxiliary variable mediation model with two auxiliary variables.

Methods 4 and 5. Inclusion of M_1 only using ML missing data methods (Method 4) or ML with listwise deletion (Method 5).

Figure 1.
Analysis methods evaluated in simulation study.