Scand J Work Environ Health. 2020;46(2):113–116. doi:10.5271/sjweh.3886

Advancing mediation analysis in occupational health research

In recent years, mediation analysis has become a popular means to identify and quantify pathways linking an exposure to an outcome, thereby elucidating how a particular exposure contributes to the occurrence of a specific outcome. When a mediator is a modifiable risk factor, this opens up new opportunities for interventions to block (part) of the exposure's effect on the outcome. Recent examples in *Scandinavian Journal of Work Environment Health* have addressed the mediating effect of wellbeing on the association between type of office and job satisfaction (1) and examined whether workplace social capital contributes to the association between organizational changes and employee exit from work (2).

Mediation analysis requires a specific study design ensuring that the temporal sequence of exposure, mediator, and outcome supports the argument for causation. A good illustration presented by Halonen and colleagues (3) used the Swedish Longitudinal Occupational Survey of Health (SLOSH) with biannual waves to investigate whether depressive symptoms mediated the association between effort–reward imbalance and subsequent neck-shoulder pain. The longitudinal design allowed for a mediation analysis whereby exposure was measured two years before the mediator, and the mediator was assessed two years before the outcome of interest. When exposure and mediator are self-reported, without assurance that the mediator occurred after the exposure (eg, years of education in school preceding exposure to working conditions), mediation analysis is prone to reversed causality (ie, not the exposure causing the mediator but the mediator causing the exposure) and has little to offer in terms of a better understanding of the mechanism how exposure is linked to the outcome. A linked issue is the stability of the mediator over time. Some authors have argued that mediation analysis is bound to fail in studies with a time window of measurements that does not sufficiently capture the fluctuations of the potential mediator over time (4).

Recent debates on mediation analysis have specifically addressed the methodology and assumptions that allow for a causal interpretation. The traditional approach to mediation analysis in health research has largely relied on the influential 1986 paper of Baron & Kenny (5). In this approach, mediation is assessed by estimating the effect of the exposure on the mediator and the effect of the mediator on the outcome (adjusting for the exposure), and multiplying the two estimates to derive the indirect or mediated effect (the "product-of-coefficients" approach). Alternatively, the effect of the exposure on the outcome is estimated with and without adjustment for the mediator, and the difference between the two estimates is used to quantify the mediated effect (the "difference-of-coefficients" approach). What has been largely overlooked, however, is the fact that these approaches are only valid in linear models and rely on the assumption that there is no interaction between exposure and mediator on the outcome (6). Moreover, because the importance of controlling for mediator-outcome confounders was never mentioned in Baron & Kenny, many studies have neglected to account for potential confounding factors of the mediator-outcome relationship. An eloquent illustration of the problems of the classical mediation method and the risk of severe bias in the mediated effect is presented by Pearce & Vandenbroucke (7).

In response to the limitations of traditional mediation methods, a new mediation methodology is rapidly developing and often referred to as "causal mediation analysis" (6, 8). This methodology has derived counterfactual definitions of direct and indirect effects, which are independent of any model, and – under explicitly stated assumptions – allow for a causal interpretation of the estimated effects (9). These definitions enable greater flexibility and provide researchers with the tools to scrutinize the assumptions underlying mediation analysis better. The causal inference approach to mediation has given rise to two different types of effects: (i) natural direct and indirect effects, and (ii) controlled direct effects. Natural direct and indirect effects by definition always sum up

to the total effect, even in the presence of exposure–mediator interaction. Intuitively, the natural direct effect captures the effect of the exposure on the outcome that is not due to its effect on the mediator, whereas the natural indirect effect captures the effect of the exposure on the outcome that is due to its effect on the mediator (8). Controlled direct effects, on the other hand, quantify the effect of the exposure on the outcome if the mediator was fixed at a specific value uniformly in the population. Consequently, whereas a total effect can always be decomposed into a natural direct and indirect effect, controlled direct effects are estimated for every level of the mediator (which may differ substantially depending on the magnitude of the interaction effect between exposure and mediator).

Decomposition of a total effect of an exposure on an outcome into natural direct and indirect effects requires that there is no unmeasured confounding of the (i) exposure–outcome, (ii) exposure–mediator, and (iii) mediator–outcome relationships. Furthermore, a critical assumption is (iv) that there are no measured or unmeasured mediator–outcome confounders that are themselves affected by the exposure (6). The first three assumptions are depicted in Figure 1: if C_1, C_2 and C_3 include all relevant confounders of the three pathways under study (A→Y, A→M, and M→Y), the first three assumptions are met.

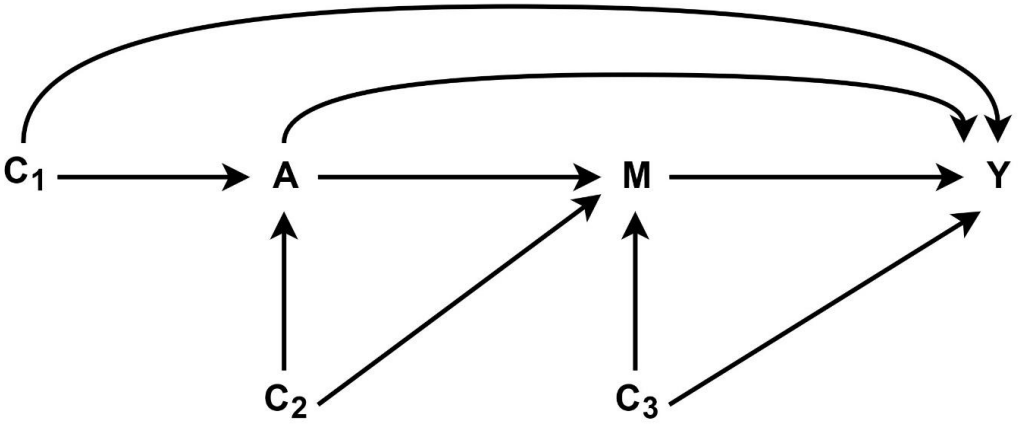

Figure 1. Generic directed acyclic graph for mediation analysis. A=exposure, M=mediator, Y=outcome, C_1=exposure–outcome confounder, C_2=exposure–mediator confounder, C3=mediator–outcome confounder.

The critical fourth assumption that the mediator–outcome association is not confounded by any variable that is itself affected by the exposure is depicted in figure 2. Here (L) is both a confounder of the M→Y path and a mediator on the A→Y path. An example would be the association between shift work (A) and diabetes mellitus (Y), as reported by Tucker and colleagues (10). Lack of leisure-time physical activity may be considered a mediator (M). Whereas controlling for confounders (C) (eg, age and sex) is straightforward, controlling for overweight (L) may be problematic. The issue here is that overweight may be a confounder of the association between leisure-time physical activity and diabetes mellitus (individuals with overweight might find it difficult to be active in their leisure), and overweight may also be on the causal pathway from shift work to diabetes mellitus (eg, by an effect of shift work on dietary patterns or physiological processes). In the presence of such exposure-induced mediator–outcome confounders, the researcher is stuck between bad choices: not adjusting for (L) will bias the estimated effect of the mediator; adjusting for (L) will bias the estimated effect of the exposure. Hence, researchers must convince themselves that such confounding is not present. When potential confounding is a serious possibility, natural direct and indirect effects cannot be identified, regardless of whether (L) is measured or not. In such a situation, new methods are required that allow for effect decomposition in the presence of exposure-induced mediator–outcome confounders by using so-called interventional (in)direct effects, which consider potential interventions on a population level (11, 12).

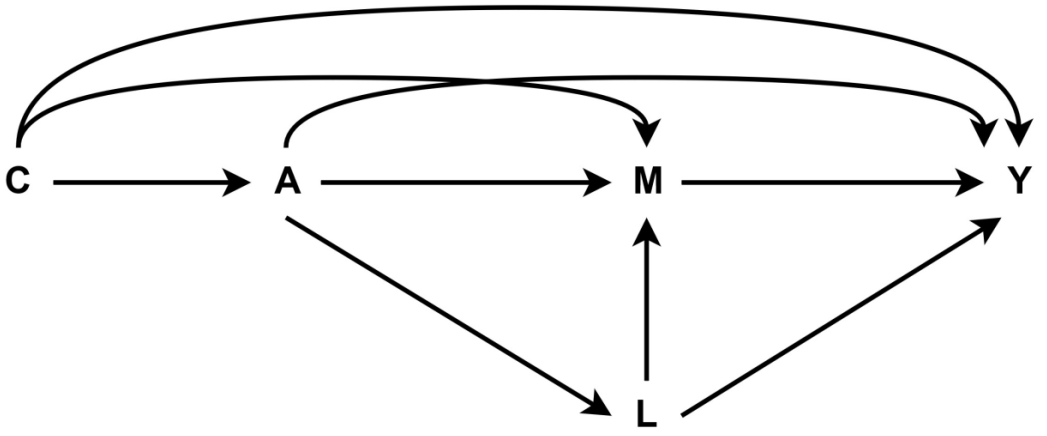

Figure 2. Generic directed acyclic graph for mediation analysis including exposure-induced mediator–outcome confounders (L).

An alternative approach in mediation analysis is to quantify controlled direct effects that only rely on assumptions (i) and (iii): no uncontrolled exposure–outcome confounding and no uncontrolled mediator–outcome confounding. Moreover, some have argued that controlled direct effects are much more policy-relevant because they estimate the proportion of the total effect of the exposure on the outcome that could be eliminated by a specific intervention on the mediator (13, 14). Although controlled direct effects can still be estimated in the presence of exposure-induced mediator–outcome confounders, this does require more sophisticated methods, such as marginal structural models (15, 16) or the g-formula (17). Albeit still infrequently used, these approaches offer an important new toolbox for health research and are often easily implemented in standard statistical software (18).

To conclude, mediation analysis has gained a lot of attention in recent years and is rapidly developing with regard to approach and methods. We invite researchers to move away from traditional mediation analysis and apply the newer mediation methods in their studies. First, a critical stance is required to evaluate whether the study design and data collected permit a meaningful mediation analysis. Second, studies should explicitly address the crucial assumptions that are fundamental for a causal interpretation of mediation. Third, sensitivity analysis is helpful to assess the robustness of the results to potential violations of the underlying assumptions.

References

1. Otterbring T, Pareigis J, Wästlund E, Makrygiannis A, Lindström A. The relationship between office type and job satisfaction: Testing a multiple mediation model through ease of interaction and well-being. Scand J Work Environ Health. 2018;44(3):330-34. https://doi.org/10.5271/sjweh.3707
2. Jensen JH, Flachs EM, Skakon J, Rod NH, Bonde JP. Longitudinal associations between organizational change, work-unit social capital, and employee exit from the work unit among public healthcare workers: a mediation analysis. Scand J Work Environ Health. 2019;45:53-62. https://doi.org/10.5271/sjweh.3766
3. Halonen JI, Lallukka T, Virtanen M, Rod NH, Magnusson Hanson LL. Bi-directional relation between effort-reward imbalance and risk of neck-shoulder pain: assessment of mediation through depressive symptoms using occupational longitudinal data. Scand J Work Environ Health. 2019;45:126-33. https://doi.org/10.5271/sjweh.3768
4. Rod NH, Lange T. Re: Current limits for understanding the underlying causal structures in observational epidemiology: reflections on refined mediation models and crude data. Epidemiology 2017;28:e60-61. https://doi.org/10.1097/EDE.0000000000000708
5. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51:1173-82. https://doi.org/10.1037/0022-3514.51.6.1173
6. 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;18:137-50. https://doi.org/10.1037/a0031034

7. Pearce N, Vandenbroucke JP. Causation, mediation and explanation. Int J Epidemiol 2016;45:1915-22. https://doi.org/10.1093/ije/dyw281

8. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol 2013;42:1511-9. https://doi.org/10.1093/ije/dyt127

9. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. Epidemiology 1992;3:143-55. https://doi.org/10.1097/00001648-199203000-00013

10. Tucker P, Härmä M, Ojajärvi A, Kivimäki M, Leineweber C, Oksanen T, et al. Associations between shift work and use of prescribed medications for the treatment of hypertension, diabetes, and dyslipidemia: a prospective cohort study. Scand J Work Environ Health. 2019;45:465-74. https://doi.org/10.5271/sjweh.3813

11. Moreno-Betancur M, Carlin JB. Understanding Interventional Effects: A More Natural Approach to Mediation Analysis? Epidemiology 2018;29:614-7. https://doi.org/10.1097/EDE.0000000000000866

12. Vanderweele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. Epidemiology 2014;25:300-6. https://doi.org/10.1097/EDE.0000000000000034

13. VanderWeele TJ. Policy-relevant proportions for direct effects. Epidemiology 2013;24:175-6. https://doi.org/10.1097/EDE.0b013e3182781410

14. Naimi AI, Kaufman JS, MacLehose RF. Mediation misgivings: ambiguous clinical and public health interpretations of natural direct and indirect effects. Int J Epidemiol 2014;43:1656-61. https://doi.org/10.1093/ije/dyu107

15. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11:550-60. https://doi.org/10.1097/00001648-200009000-00011

16. Nandi A, Glymour MM, Kawachi I, VanderWeele T. Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. Epidemiology 2012;23:223-32. https://doi.org/10.1097/EDE.0b013e31824570bd

17. Vangen-Lonne AM, Ueda P, Gulayin P, Wilsgaard T, Mathiesen EB, Danaei G. Hypothetical interventions to prevent stroke: an application of the parametric g-formula to a healthy middle-aged population. Eur J Epidemiol 2018;33:557-66. https://doi.org/10.1007/s10654-017-0344-x

18. Hernán MA, Robins JM. Causal Inference: What If. Boca Raton: Chapman & Hall/CRC, 2020.

Joost Oude Groeniger and Alex Burdorf
Department of Public Health
Erasmus Medical Center Rotterdam
Rotterdam, The Netherlands
[Email: j.oudegroeniger@erasmusmc.nl]