Marginal Structural Models
The Way Forward for Life-course Epidemiology?

Bianca L. De Stavola and Rhian M. Daniel

The study of life-course socioeconomic disadvantage and health raises several important conceptual and methodologic problems. Nandi and coauthors[1] attempt to address some of these in this issue of EPIDEMIOLOGY. In this commentary, we review them and discuss whether their proposed solution is sufficiently broad to be invoked more generally in life-course research, and we highlight the caution required when doing so.

CONCEPTUAL MODELS

Definitions

Three main conceptual models underpin research in this field[2,3]: the cumulative exposure, the critical/sensitive periods, and the pathways model. The first model assumes that current risk is related to cumulative exposure over an extended time interval, with every period in that interval equally influential. The critical/sensitive periods model instead allows for certain sections of the time interval to be the only ("critical") or the most important ("sensitive") contributors to risk, whereas for the pathways model, sensitive periods interact in their impact on risk.

Empirical Evidence and Its Limitations

Empirical support for any of these models across a range of adult health outcomes is inconsistent[2,4] for several reasons. The analytic methods commonly used may not be adequate, as discussed later in the text, but also the data used to assess the various conceptual models may not be appropriate for the task in hand. Socioeconomic position/status/condition is somewhat loosely defined and its meaning (and impact on health) has changed considerably during the past century in high-income countries and is now changing even more rapidly in low- and middle-income countries. Its measurement is particularly prone to error, often relying on self-reported income, access to assets, attained education, occupation, etc, with several heterogeneous groups merged together into somewhat arbitrary categories. Furthermore, socioeconomic data are rarely available at >2 time points over a person's life-span, limiting the ability to identify critical periods, or indeed to examine the effect of individual socioeconomic trajectories.

Nandi and colleagues[1] acknowledge these difficulties when analyzing the US Health and Retirement Study to examine the effect of life-course exposure to adverse socioeconomic status (SES) on 3 related adult diseases: coronary heart disease (CHD), diabetes, and stroke. They address the difficulty of measuring SES by adopting a latent variable approach and assuming (implicitly) that missingness in manifest indicators occurs at random, and that data on childhood exposure are not affected by recall bias. These are often inevitable assumptions, but they should nevertheless be assessed, for example, through sensitivity analyses.[5] More importantly for this discussion, Nandi and colleagues

From the Centre for Statistical Methodology, London School of Hygiene and Tropical Medicine, London, United Kingdom.
Supported in part by the UK Economic and Social Research Council (grant ES/I025561/1). The authors reported no other financial interests related to this research.
Correspondence: Bianca L. De Stavola, Centre for Statistical Methodology, London School of Hygiene and Tropical Medicine, U.K., Keppel Street, London, UK WC1E 7HT. E-mail: bianca.destavola@lshtm.ac.uk.

ISSN: 1044-3983/12/2302-0233
DOI: 10.1097/EDE.0b013e318245847e

have access to exposure data on just 2 life periods: childhood and adulthood. Therefore, they can examine only very simplified versions of the conceptual models of interest. The authors use marginal structural models[6] to estimate the direct effect of childhood SES on each of the outcomes of interest, as we now describe. Evidence that the effect of childhood SES is not completely mediated by adult SES would support the hypothesis that "early environmental conditions have lasting effects,"[1] ie, that the critical model in which there is no effect of childhood SES is not supported by the data (Fig., model b).

From Conceptual to Marginal Structural Models

Assuming for simplicity that SES in childhood and adulthood are both binary variables (the extension to 4 levels of SES is immediate), the causal model—a marginal structural model—fitted by Nandi et al (on a log risk scale) is:

$$log[Pr\{Y(a,m) = 1\}] = \alpha + \beta a + \gamma m + \delta am \quad (1)$$

where $Y(a,m)$ is the potential value of the binary outcome (CHD, diabetes, stroke) under a hypothetical intervention that sets childhood SES—the exposure—to level a, and adult SES—the mediator—to level m.

The parameters (β,γ,δ) of model (1) can be interpreted in terms of simplified two-exposure versions of the conceptual models mentioned above, as follows. All models except for the pathways model assume that $\delta = 0$. In the cumulative model, $\beta = \gamma$, whereas in the critical periods model, either $\beta = 0$ or $\gamma = 0$. Finally, in the sensitive periods model, either $\beta \gg \gamma > 0$ or $0 < \beta \ll \gamma$.

It is important to note, however, that model (1) assumes a particular parameterization of a direct effect, namely the controlled direct effect, in which the value of the mediator is fixed at m for all subjects. This raises conceptual problems ("Can we imagine a hypothetical world in which adult SES is high for everyone? And is this relevant?") as well as practical ones, mainly that—in the presence of an interaction between exposure and mediator in their effect on the outcome—there exists no corresponding definition of an indirect effect, such that the total effect of the exposure can be partitioned into its direct and indirect components. Such partitioning is possible under an alternative definition, namely that of the natural direct effect.[7]

Nandi et al avoid these difficulties since they find little evidence to reject a hypothesis of no interaction between childhood and adult SES and, hence, proceed under the assumption that $\delta = 0$, focusing on the parameter β, which could then be interpreted as both the controlled and natural direct effect of childhood SES on the outcome, unmediated by adult SES. In the presence of exposure–mediator interaction, the natural direct effect would not, in general, be identifiable from these data, due to the problem of endogenous confounding.[8] Unfortunately, however, by focusing on just the parameter β in a greatly simplified statistical model, only limited conclusions can be drawn with respect to the conceptual models.

Even within these limited confines, we nevertheless share the authors' pragmatic enthusiasm, although we find some of their results surprising. We discuss a possible explanation for this in the next section, and in doing so highlight the more general need for caution when using these and similar methods in life course research.

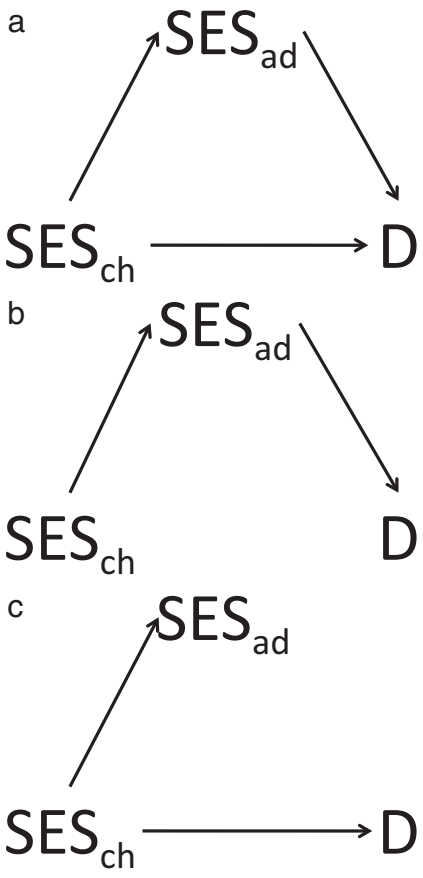

FIGURE. Diagrammatic representation of the conceptual models discussed by Nandi et al.[1] Model (a): general (cumulative, critical/sensitive periods or pathways); Model (b): critical period model where only adult SES has a direct influence on D; Model (c): critical period model where only childhood SES has a direct influence on D. SES_{ch} indicates socioeconomic status in childhood; SES_{ad}, socioeconomic status in adulthood; D, outcome.

CAUSAL MEDIATION ANALYSIS

It is now broadly recognized[9] that standard regression models (and therefore traditional mediation analysis[10]) cannot deal with endogenous confounding (confounders of the mediator–outcome relationship affected by the exposure),

even if measured. This is likely to arise, however, when the exposure occurs much earlier than the mediator, as is the case in the present context. In addition, when both exposure and mediator have multiple pathways of influence on the outcome, then it is likely that variables on one pathway influence variables on another, thus also leading to endogenous confounding. More specifically in this context, childhood SES is likely to influence health behavior in early adulthood, which in turn might affect adult SES, with the same pattern repeated with adult SES and later health behaviors. Thus, smoking and diet in adolescence, for example, which are important risk factors for CHD, act both as confounders of the effect of later SES and intermediate variables from earlier SES, ie, they are potentially endogenous confounders.

The Problem With Standard Methods

Either conditioning or not conditioning on endogenous confounders leads to biased estimates of the direct effect of the exposure on the outcome. By not conditioning on the endogenous confounders while instead conditioning on the mediator, the estimate of the direct effect suffers from so-called collider-stratification bias. Conditioning on the mediator (a collider because it is a common effect of exposure and endogenous confounders) alters the association between exposure and endogenous confounders. Conditioning on the endogenous confounders is not satisfactory either, because this blocks the part of the direct effect transmitted through them; furthermore, if there are unmeasured common causes of the endogenous confounders and the outcome, the endogenous confounders are themselves colliders, and further bias is induced by conditioning on them. Note that for the effect of the mediator on the outcome, all confounders are upstream of the mediator, and thus, standard adjustment can be used (with the usual caveats of requiring a sufficient set of measured confounders).

Marginal Structural Models and Their Application to the US Health and Retirement Study

These considerations lead Nandi and colleagues[1] to adopt a modeling approach that avoids conditioning on endogenous confounders. They fit the marginal structural model of equation (1) (extended to 4 categories of SES) via inverse probability of treatment-weighted (IPTW) estimation[11] to obtain the controlled direct effect of childhood SES on the 3 binary outcomes of interest, ie, the effect that is unmediated by adult SES. Contrary to the results obtained when conditioning on the endogenous confounders, the authors find evidence in support of a direct effect of childhood SES, and therefore, against one of the 2 critical-periods models (the one in which $\beta = 0$) for each of the outcomes considered. Surprisingly, they neither discuss whether there is evidence of an effect of the mediator nor the relative magnitudes of these effects, which would potentially allow discrimination between the alternatives (cumulative, sensitive, or critical [with $\gamma = 0$]] models; they had already ruled out the pathways model). The estimated effect of adult SES, particularly on coronary heart disease (Table 2; Nandi et al), is not in line with what we would expect (eg, Kaplan and Keil[12]). The harmful effect of low SES in adulthood seen in the unadjusted analysis disappears, and even looks somewhat protective, in the adjusted analysis and in the IPTW estimation of the marginal structural model. (Note that for the effect of the mediator, as noted earlier in the text, we expect agreement between the latter 2 if there is no unmeasured confounding.) In the light of this surprising finding, the other results should be interpreted with caution, and the assumptions made in reaching them carefully questioned.

Assumptions

By using causal mediation analysis in the form of IPTW estimation of marginal structural models, the authors avoid a major potential bias: that due to endogenous confounding. This, of course, is to be commended. Another advantage of using a more rigorous approach to causal inference in this way is that the assumptions required (sometimes obscured in standard approaches) are made transparent. These assumptions fall into 3 categories: (1) those shared between the "standard" approach (ie, regression analysis where the mediator is or is not included) and the "alternative" approach, IPTW estimation of marginal structural models; (2) those additionally required by the alternative approach; and (3) those required only by the traditional approach. Although Nandi et al briefly acknowledge these, they do not discuss the potential implications of their violation.

Putting aside the assumptions concerning missing data and measurement error, the remaining assumptions are:

i. Consistency—essentially that hypothetical interventions on SES be well-defined;
ii. that the correct structural assumptions are (implicitly or explicitly) made, ie, that the analysis is based on an appropriate causal diagram;
iii. no unmeasured exposure–outcome confounding;
iv. no unmeasured mediator–outcome confounding;
v. positivity—essentially that there is sufficient movement of persons across categories of SES (from childhood to adulthood) and that this movement cannot be perfectly predicted from covariates—necessary to calculate the weights in IPTW;
vi. correct parametric specification of the models for the exposure given the baseline confounders and of the model for the mediator given exposure, baseline and endogenous confounders—necessary to calculate the weights in IPTW;
vii. correct parametric specification of the MSM.

Assumptions (i) to (iv) belong to category (1) above, and the remainder to category (2), although an assumption

very similar to (vii)—namely the correct specification of an outcome model—would be needed in any approach. The assumption of no endogenous confounding belongs to category (3) and therefore is not needed by Nandi et al.

The veracity of assumptions (i) to (vii) should all be carefully considered, but due to space constraints, we focus here on 2 of them.

A major concern with regard to assumption (ii) is that some or all of the presumed endogenous confounders are not confounders at all, but lie on the causal pathway from adult SES to the outcomes (ie, that the direction of arrow 6 in Fig. 1 of Nandi et al should be reversed). High blood pressure, smoking, alcohol consumption, etc, all measured in adulthood, could be caused by low adult SES, rather than vice versa. If so, what would be the implications for their analyses? In the pseudo-population created by the reweighting, the risk factors would be made independent of adult SES, eliminating any effect of adult SES transmitted through them. In addition, if there were an unmeasured common cause U (eg, underlying health status) of the risk factors and the outcome, U would be associated with childhood and adult SES in the weighted pseudo-population, even if no such association originally existed in the data. Briefly, this is due to the conditional association induced between the mediator and U given the risk factors. Furthermore, it can be shown that these conditional associations would act in opposite directions, with U being positively associated with low childhood SES and negatively associated with low adult SES in the pseudo-population, if it were positively associated with the risk factors. This is a possible explanation for the unexpected apparent protective effect of low adult SES on CHD, and the same phenomenon would inflate the estimates of the direct effect of low childhood SES. Under this alternative diagram, and in the absence of baseline confounding, the unadjusted analysis would be the correct one. The truth probably lies somewhere between these 2 extremes, with some feedback between adult SES and risk factors that can not be disentangled using the available contemporaneous measurements. Nevertheless, the fact that the risk factors are more plausibly on the causal pathway from adult SES to CHD than to the other outcomes strengthens our concern.

A further concern is the form of the chosen marginal structural model (assumption (vii)). The study is a prospective cohort study covering 14 years of follow-up (from 1992 to 2006), and the outcomes are not rare (especially CHD, 24%, but also diabetes, 18%). Despite this, Nandi et al use log-linear marginal structural models to estimate (controlled direct) risk ratios, ignoring the data on time to these events. Their results are weakened by this limitation. The reason for this modeling choice, we presume, is that causal-mediation analysis for time-to-event outcomes is currently in its infancy,[13–15] and this application highlights the need for further development.

DISCUSSION

What lessons can be learned from examining this application of mediation analysis to life-course data?

First, there is an urgent need to bridge the gap between conceptual models and their statistical counterparts. Conceptual models are complex and their comparison far from straightforward. The framework of causal mediation analysis is useful in attempting to link these concepts with relevant causal parameter(s), but also in recognizing the limitations of this endeavor.

Second, appreciation and critical understanding of all the assumptions required to obtain valid estimates of the mediation parameters are paramount. These go beyond the acknowledgment of the no-unmeasured-confounding assumptions and include drawing an appropriate causal diagram, and correctly selecting and specifying all the models involved (eg, the marginal structural model and the models to estimate the relevant weights), if the perils of a "black-box" approach are to be avoided. Approaches that allow the relaxation of certain assumptions (such as "no endogenous confounding") are greatly welcomed, but current limitations (as in the case of time-to-event outcomes) sometimes necessitate compromises on other fronts, underlining the need for the continued development of causal-mediation methods.

Third, the questions commonly addressed in life-course investigations are often more general than just whether a particular direct effect is supported by the data. Lack of appreciation of this may lead to ignoring signs that the results from a particular analysis may be inconsistent. This is particularly evident in the lack of discussion by Nandi et al of the apparent protective causal effect on CHD of low adult SES.

Finally, the exposures of greatest interest in life-course epidemiology (eg, not only SES but also life-time changes in body size, in systolic blood pressure, etc) vary over time, have time-varying effects, and mutually affect each other. Hence, the timing of measurement itself is crucial. However, availability of data over the life course is often dictated by opportunities and costs, thereby restricting our ability to pursue many of the investigations implied by the conceptual models.[16] This is a general issue for research in this area, and should lead to greater (but also principled) exploitation and linkage of existing data sources, such as medical records, school files, employment registries, census databases, and purposely defined studies.

Despite all these problems, we still believe that causal mediation analysis in general—and marginal structural models specifically—offer new and valuable tools to researchers involved in life-course epidemiology. We should not be afraid to use the language and methods of causal mediation analysis, and we should embrace the opportunity afforded by their formalism to be explicit about the assumptions required when using them. If we do so, sensible applications are possible and should be pursued.

ABOUT THE AUTHORS

BIANCA DE STAVOLA is Professor of Biostatistics at the London School of Hygiene and Tropical Medicine and Director of its Centre for Statistical Methodology. Her current research involves the development and implementation of statistical methods for long-term longitudinal studies with specific applications to life course epidemiology. Directly related to this work is her commitment to improving the application of causal inference methods. RHIAN DANIEL is a Research Fellow in Medical Statistics at the London School of Hygiene and Tropical Medicine. Her focus is on methods for dealing with time-varying confounding in longitudinal observational studies and also the related problem of estimating direct and indirect causal effects.

ACKNOWLEDGMENTS

We gratefully acknowledge the insights offered by Dave Leon, George Ploubidis, Costanza Pizzi, and Isabel dos Santos Silva.

REFERENCES

1. Nandi A, Glymour MM, Kawachi I, VanderWeele TJ. Using marginal structural models to estimate the direct effect of adverse childhood social conditions on onset of heart disease, diabetes, and stroke. *Epidemiology*. 2012;32:223–232.
2. Lynch J, Davey Smith G. A life course approach to chronic disease epidemiology. *Annu Rev Public Health.* 2005;26:1–35.
3. Pickles A, De Stavola BL. New methods in life course analysis. *Methods in Life Course Research*. In: Wadsworth M, Maughan B, Pickles A, eds. Oxford: Oxford University Press; 2007.
4. Galobardes B, Lynch JW, Davey Smith G. Childhood socioeconomic circumstances and cause-specific mortality in adulthood: systematic review and interpretation. *Epidemiol Rev.* 2004;26:7–21.
5. Verbeke G, et al. Sensitivity analysis for nonrandom dropout. A local influence approach. *Biometrics*. 2001;57:7–14.
6. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550–560.
7. Pearl J. Direct and indirect effects. In: Proceedings of the 17th Annual Conference on Uncertainty in Artificial Intelligence; 2001; (San Francisco):
8. Avin C, Shpitser I, Pearl J. Identifiability of path-specific effects. In: Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence; 2005; (Edinburgh):
9. Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol*. 2002;31:163–165.
10. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51:1173–1182.
11. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*. 2008;168:656–664.
12. Kaplan GA, Keil JE. Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation*. 1993;88:1973–1998.
13. Lange T, Hansen JV. Direct and indirect effects in a survival context. *Epidemiology*. 2011;22:575–581.
14. Martinussen T, et al. Estimation of direct effects for survival data by using the Aalen additive hazards model. *J R Stat Soc Series B Stat Methodol*. 2011;73:773–788.
15. VanderWeele TJ. Causal mediation analysis with survival data. *Epidemiology*. 2011;22:582–585.
16. Davey Smith G. Lifecourse epidemiology of disease: a tractable problem? *Int J Epidemiol.* 2007;36:479–480.