Statistical Science
2020, Vol. 35, No. 3, 391–393
https://doi.org/10.1214/20-STS777
Main article: https://doi.org/10.1214/20-STS720
© Institute of Mathematical Statistics, 2020

Comment: Clarifying Endogeneous Data Structures and Consequent Modelling Choices Using Causal Graphs

Erica E. M. Moodie and David A. Stephens

We read with great interest the article by Qian, Klasnja and Murphy (2020), and commend the authors for focusing on principled estimation and providing a quantitative approach to healthcare delivery through mobile devices. The quantitative analyses studied here could have wide-ranging applications that may serve to increase patient empowerment by taking medical monitoring and even intervention out of the clinic and into the home.

Here, we wish to delve into two complementary aspects of the work: first, we attempt to give clarifications concerning the parameter(s) of interest, and second, we provide visualizations of potential scenarios that may help to clarify estimands and when biases due to endogeneity may arise.

1. TREATMENT EFFECTS: ONE, TWO OR TOO MANY?

The authors focus on the setting of (micro)randomized trials, where the treatment of interest is assigned entirely at random. In this setting, one would often expect to be able to perform causal inference, since a key barrier to doing so— confounding—is eliminated thanks to the randomized nature of the treatment assignment. In the motivating HeartSteps study, for instance, the question of interest is to determine the optimal treatment strategy to remind an individual to exercise or not based on their location and recent step activity, with the goal of maximizing steps taken over the next 30 minutes. This question suggests a causal estimand, targeting the effect of the reminder and any modification of the reminder effect by individual covariates.

We attempt here to provide a more precise focus on the estimand in plausible scenarios of interest: a single treatment effect (a 'main effect model'), an effect that is modified by covariates (an interaction model), or a truly individualized treatment effect characterized by a random slopes model. In such cases, the effect of covariates beyond their modification of treatment are not of primary interest. Looking to equation (15) and Table 2 of Qian, Klasnja and Murphy (2020), the βs are the only parameters of interest, while the random effects b_{i1} will, themselves, also be essential to tailoring treatment recommendations. As we will demonstrate in the next section, sharpening attention to the parameters of interest allows the analyst to step back from the complexities of *all* dependencies within the longitudinal data generating structure, and take note of those most relevant to the scientific question.

Suppose that there does exist heterogeneity in the treatment effect that cannot only be explained by covariates, but rather requires a random slope term. This would imply a treatment strategy that requires knowing or inferring an individual's random effect prior to being able to implement the treatment strategy. In the setting described in the study, where there are over 150 measures available on average for the participants, this is feasible; however, the strategy would not immediately generalize to future users. Rather, a potentially significant volume of data would first need to be collected to estimate each user's random slope.

2. VISUALIZATION WITH ACYCLIC GRAPHS: UNDERSTANDING ENDOGENIETY AND CONSEQUENT MODELLING CHOICES

Qian, Klasnja and Murphy (2020) raise a number of interesting scenarios where bias arises even in seemingly simple situations, such as when treatment is randomized. Some of the scenarios raised may be familiar to those with a causal inference background, whereas there are others that are less obvious and perhaps made somewhat less clear without the explicit specification of the estimand of interest. Here, we attempt to clarify, through the visual means provided by causal diagrams (Greenland, Pearl and Robins, 1999), the estimand(s) of interest and possible sources of bias. To emphasize the estimand-focused framework of a causal paradigm, the effects of interest are shown as black arrows, with other conditional dependencies in the data-generative model are shown in grey.

Consider Figure 1, panel A, which follows the notation of Qian, Klasnja and Murphy (2020): a longitudinal setting where X is endogenous. The relationship of interest

Erica E. M. Moodie is Associate Professor, Biostatistics, McGill University, 1020 Pine Ave W., Montreal, Quebec, Canada H3A 1A2 (e-mail: erica.moodie@mcgill.ca). David A. Stephens is Professor, Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Ave W., Montreal, Quebec, Canada H3A 0B9 (e-mail: david.stephens@mcgill.ca).

is the effect of X_j on Y_{j+1}. Using the usual principles of conditional independences in directed acyclic graphs (DAGs), there is the potential for bias in estimation of the effect of interest if there exists an open 'backdoor' path between X_j and Y_{j+1}, that is, a sequence of edges connecting nodes that is neither blocked by conditioning on the node variables nor by a 'collider' node that has two inbound directed edges (Greenland, Pearl and Robins, 1999). In panel A, we observe a backdoor path from Y_{j+1} to X_j through the random effect b_0 and the previous outcome Y_j. It is therefore clear that it is necessary to block this path, and one means to doing so is by using a model that conditions on the random effect. Note that we assume no correlation between the covariates and the random effect; if such correlations were suspected, these could be incorporated into the graph, for example, by including a new node in the graph that was a common cause of both X_j and b_0.

We now add a randomized treatment to the scenario in panel B of Figure 1, without any further complications: we assume no heterogeneity in the treatment effect, either through a random slope or via interactions with the covariates. It is evident that there is no path between A_j and Y_{j+1} except the one which is directly of interest, and which encapsulates the estimand of the direct effect of treatment on the proximal outcome. The endogeneity of X is irrelevant here, because X is not a confounder nor does it play any role in the estimand of interest, which is the treatment (A) effect.

Finally, consider scenarios most pertinent to behavioral treatment delivery: in the first, treatment is modified by covariates, and thus what is optimal for a given individual at one time point may not be so for someone else, or even for that individual at another time due to their evolving covariate profile. In the second, treatment effects vary across individuals, with this heterogeneity being captured in a random slope. Visualizing these scenarios is less straightforward.

There has been some discussion, and a number of possible representations, of interactions in DAGs (Weinberg, 2007, Lopez, Subramanian and Schooling, 2019) and, as recently as 2014, it has been stated that "DAGs cannot represent interactions or effect modification" (Foraita, Spallek and Zeeb, 2014) though many would disagree (VanderWeele and Robins, 2007). Regarding the DAG as a structural but not quantitative representation of the probabilistic relationship between variables, we may conclude that an interaction node is merely a node that receives inbound directed edges from the interacting variables in a deterministic relationship. Consequently, we represent an interaction via arrows from the interacting variables to both the outcome and a product node (see Figure 1, panels C and D).

In panel C, the critical pathways to determining individualized treatment recommendations are then the direct

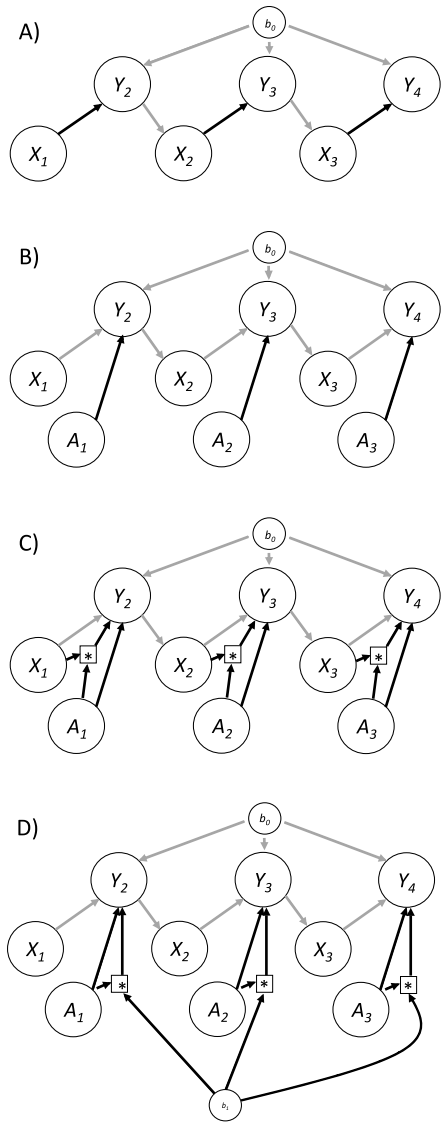

FIG. 1. *Directed acyclic graphs depicting potential scenarios of interest.* (A) *Endogenous covariate X and a random intercept b_0.* (B) *Scenario A + a sequentially randomized treatment A.* (C) *Scenario B + treatment covariates interactions ($A \cdot X$), depicted by *.* (D) *Scenario B + treatment heterogeneity in the form of a random slope, b_1, that interacts with treatment as depicted by *. Black arrows denote relationships of interest (which will ultimately be parameterized in some way to describe the estimand of interest); grey arrows denote conditional dependencies that are not of explicit interest. For simplicity, serial dependence between Xs and between Ys is omitted.*

effect of A_j on Y_{j+1} (i.e., the impact of treatment at reference levels of covariates X_j) and all effects of A_j on Y_{j+1} that are moderated by X_j. Here, as in the first scenario from panel A, we see there exists a path through the random intercept that could bias estimation. Although there is an open path between Y_{j+1} to X_j, conditioning on X_j blocks the back door path between $X_j A_j$ and Y_{j+1}, implying conditioning on the random intercept b_0 is not necessary. In contrast, in panel D, treatment heterogeneity is in the form of a random slope (but no treatment-covariate interactions). The effect of interest is thus encoded in the

pair of arrows from A_j and $A_j b_1$ into Y_{j+1}, and so it is clear that conditioning on b_1 is needed to account for the individual-specific component of the treatment effect.

These plots can of course be elaborated upon, based on assumptions concerning the data generation process—e.g. heterogeneity due to both treatment-covariate interactions and a random slope on the treatment effect could be considered. While the graphs become more complex, deliberate focus on the estimand(s) of interest may help to elucidate when a mixed model is an appropriate choice, and whether a conditional-on-random effect model is required.

Using simple visualizations requires careful consideration and a precise elaboration of the assumptions in the domain science context, as was noted by Qian, Klasnja and Murphy (2020) regarding the conditional independence assumption. Many of these assumptions will not be empirically testable; however, an explicit description can help to clarify ideas, and open dialogue between researchers.

3. TIME-VARYING CONFOUNDING AND MEDIATION

The issue of time-varying confounding is discussed in an early section of the article. We contend that the definition given warrants elaboration. It is well established in the literature that a time-varying confounder *need not* be a mediator, and different inferential methods may be required depending on whether or not mediation is present. For example, in a marginal model, inverse probability of treatment weighting (IPTW) will provide valid inference whether or not mediation is present. However, a model that conditions on the time-dependent variables may only give valid inference if there is no mediation. Graphical representations such as those in Figure 1 are again useful in elucidating these issues (see, e.g., Moodie and Stephens, 2010, Moodie and Stephens, 2011). In particular, it can easily be seen from a causal diagram that a time-varying confounder that is not also a mediator is *not*, in general, endogenous.

4. FINAL REMARKS

Mobile health is a growing area that promises to revolutionize health management. It is encouraging to see attention being paid to the statistical approaches that can make use of data from mobile phones and wearable devices, and in particular, interesting to see that familiar models such as those including random effects remain useful. Although the sample size was quite modest in the HeartSteps analysis of Qian, Klasnja and Murphy (2020), data volume will be the next challenge to consider with mobile health data. However, whatever the length of the time series or the number of individuals under study, careful consideration of the underlying data generating process should form the basis of all analyses.

ACKNOWLEDGEMENTS

EEMM acknowledges a chercheur boursier senior career award from the Fonds de recherche du Québec—Santé. The research of both authors is supported by separate Discovery Grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES

FORAITA, R., SPALLEK, J. and ZEEB, H. (2014). Directed acyclic graphs. In *Handbook of Epidemiology*, 2nd ed. (W. Ahrens and I. Pigeot, eds.) 1481–1517. Springer, New York.

GREENLAND, S., PEARL, J. and ROBINS, J. M. (1999). Causal diagrams for epidemiologic research. *Epidemiology* **10** 37–48.

LOPEZ, P. M., SUBRAMANIAN, S. V. and SCHOOLING, C. M. (2019). Effect measure modification conceptualized using selection diagrams as mediation by mechanisms of varying population-level relevance. *J. Clin. Epidemiol.* **113** 123–128. https://doi.org/10.1016/j.jclinepi.2019.05.005

MOODIE, E. E. M. and STEPHENS, D. A. (2010). Using Directed Acyclic Graphs to detect limitations of traditional regression in longitudinal studies. *Int. J. Public Health* **55** 701–703.

MOODIE, E. E. M. and STEPHENS, D. A. (2011). Marginal Structural Models: Unbiased estimation for longitudinal studies. *Int. J. Public Health* **56** 117–119.

QIAN, T., KLASNJA, P. and MURPHY, S. A. (2020). Linear mixed models with endogenous covariates: Modeling sequential treatment effects with application to a mobile health study. *Statist. Sci.* **35** 375–390.

VANDERWEELE, T. J. and ROBINS, J. M. (2007). Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am. J. Epidemiol.* **166** 1096–1104.

WEINBERG, C. R. (2007). Can DAGs clarify effect modification? *Epidemiology* **18** 569–572.