Discussion of 'Causation and Graphical Models'

Stijn Vansteelandt, Postdoctoral Fellow of the Fund for Scientific Research (Flanders, Belgium)
Ghent University, Department of Applied Mathematics and Computer Science
Krijgslaan 281, S9
9000 Gent, Belgium
Stijn.Vansteelandt@UGent.be

I wish to thank Phil Dawid for organizing this session and the discussion of these 3 very insightful and stimulating papers on causation and graphical models.

Discussion of the paper by Didelez

Vanessa Didelez (VD) addresses the difficult and important problem of adjusting for intermediate time-dependent confounders by deriving Robins' G-computation formula via augmented influence diagrams. Her incentive for using such diagrams appears to be driven by (1) her philosophy that causal effects are meaningless when logistic or other constraints preclude interventions on the exposure of interest; and (2) her desire to display with greater clarity exactly the assumptions underlying G-computation.

To identify the post-interventional law $f(Y|\overline{\sigma} = \overline{a})$, VD makes conditional independence assumptions involving intervention variables σ_t (e.g. that $Y \perp\!\!\!\perp \overline{\sigma} | \overline{A}, \overline{X}$). These variables are usually unobserved and hypothetical and do not relate to the actual study. When studying the effect of smoking on lung cancer, for instance, it is conceptually difficult to think of an actual intervention node measuring whether smoking status was assigned by intervention or whether it arose naturally, because smoking status always arises naturally in observational studies. As such, conditions involving intervention variables (e.g. $Y \perp\!\!\!\perp \overline{\sigma} | \overline{A}, \overline{X}$) have no intuitive content and thus generally obscure, rather than clarify the assumptions underlying the G-computation formula. Moreover, VD's approach appears to preclude inference on the effects of smoking (versus not smoking) because this would require one to think about interventions that force people to smoke and these are (ethically) infeasible. Such preclusions seem overly restrictive.

While VD's approach is very interesting and useful, I find the counterfactual outcomes approach more attractive here because assumptions phrased in terms of counterfactuals are *both* intuitive and verifiable by graphical techniques (Pearl and Robins, 1995). Here we consider for each subject and each sequence of interventions, $\overline{a}$, a potential outcome $Y_{\overline{a}}$. This represents the outcome that would be observed for that subject if the intervention $\overline{a}$ were applied. A typical condition underlying the G-computation formula is that $Y_{\overline{a}} \perp\!\!\!\perp A_t | \overline{X}_t, \overline{A}_{t-1}$ at each time t and for each intervention regime $\overline{a}$. In our previous example, this expresses that for people with the same smoking history and time-dependent covariates up to time $t-1$, their decision to smoke at time t is the same regardless of what their lung cancer status would be under a given smoking regime. This assumption is intuitive because it is so intimately connected to the decision process of people to smoke at each time. Such intuition helps communication with subject-matter experts, especially when graphical models are difficult to construct.

My major incentive for advocating counterfactual outcomes approaches to address this problem, is because they allow to directly parameterize intervention effects and, as such, to develop greater robustness against model misspecification. Indeed, estimation of the mean of $Y_{\overline{a}}$ requires parametric models for the distribution of high-dimensional covariates X_t conditional on past observed information. Robins (1997) argues that such models are difficult to specify and that their misspecification may introduce serious biases. Furthermore, tests of the causal null hypothesis will be infeasible via G-computation because there will in general be no parameter indexing these models that will take a fixed value, say 0, under that hypothesis. These modelling issues are so problematic that robust alternatives to G-computation are coercive. Such

alternatives have been proposed (e.g. structural nested models (Robins, 1997) and marginal structural models (Robins et al., 2000)), but - to my knowledge - have all been phrased in terms of counterfactuals. I would be very interested to see whether similar developments (based on directly parameterizing intervention effects) are possible in the framework of causal graphs.

Discussion of the paper by Lauritzen

Seeing the connection between principal and strong surrogates has been very enlightening to me. SL's notion of strong surrogates seems more useful than one based on principal strata because principal strata are unmeasurable by definition, thus creating serious conceptual and inferential problems. Strong surrogates use the notion of unmeasured confounders instead. As such, they allow the use of standard graphical model tools for inference and stimulate one to think about potential *measurable* confounders of the association between S and R.

However, a major problem with both approaches is that the conditions for a post-treatment variable to be a surrogate are difficult to verify, because of unmeasured confounders. SL suggests that these conditions are in principle falsifiable from observed data using the instrumental inequality. While nice and intriguing, this result has limited practical value because it will be impossible under certain data-generating mechanism to reject S as a strong surrogate, even with an infinite amount of data. Further, even if identification is possible, I fear that little power will be available. Finally, while the instrumental inequality also protects against measured confounders, I expect that much more power and accuracy can be gained by adjusting for those. I wonder whether the instrumental inequality can accommodate such measured confounders and continuous variables like CD4 count, bone mineral density,...

My view is that insights about the validity of surrogate endpoints must be seriously supported by a thorough understanding of the biological pathways of disease and treatment action. I think this is an attainable objective, given that surrogates are often chosen as measures of biological activity. Causal graphs offer a unique and valuable tool for incorporating such information and thus to create more accurate and powerful tests of surrogacy.

Furthermore, there will virtually always be disease pathways causally related to the true endpoint, yet unrelated to the surrogate endpoint. Such surrogates are no strong surrogates, yet may be extremely valuable. In this regard, it appears more valuable to quantify the quality of a surrogate, than to label them as strong or not. Important steps have been taken by Freedman et al. (1992), Buyse and Molenberghs (1998) and others, but all these approaches suffer from the general problem of adjustment for post-treatment variables that SL addresses. SL's paper offers valuable insights to address these problems. It would be interesting to see these approaches and SL's ideas combined.

REFERENCES

Buyse, M. and Molenberghs, G. (1998). Criteria for the validation of surrogate endpoints in randomized experiments. Biometrics, 54:1014-1029.

Freedman, L.S., Graubard, B.I., Schatzkin, A. (1992). Statistical validation of intermediate end-points for chronic diseases Statistics in Medicine, 11:167-178.

Pearl, J. and Robins, J.M. (1995). Probabilistic evaluation of sequential plans from causal models with hidden variables. Uncertainty in Artificial Intelligence, Proceedings of the 11th Conference, pp. 444-453

Robins, J.M. (1997). Causal Inference from Complex Longitudinal Data. Latent Variable Modeling and Applications to Causality. Lecture Notes in Statistics (120), M. Berkane, Editor. NY: Springer Verlag, pp. 69-117.

Robins, J.M., Hernan, M.A., Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. Epidemiology, 11:550-560.