for complex elderly dialysis patients. *Med Care* 2014;**52**(Suppl 3):S132–39.

15. Robins JM, Hernán MA. Estimation of the causal effects of time-varying exposures. In: Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G (eds). *Longitudinal Data Analysis*. New York, NY: Chapman and Hall/CRC Press, 2008.

16. Robins JM. A new approach to causal inference in mortality studies with a sustained exposure period: application to the healthy worker survivor effect. *Mathematical Modelling* 1986;**7**:1393–512.

17. van der Laan MJ, Petersen ML. Causal Effect Models for Realistic Individualized Treatment and Intention to Treat Rules. *Int J Biostat* 2007;**3**:Article3.

18. Centers for Disease Control and Prevention. 1993 Revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Morb Mortal Wkly Rep* 1992;**41**:1–19.

19. Hernán MA. The hazards of hazard ratios. *Epidemiology* 2010;**21**:13–15.

20. Thompson WA Jr. On the treatment of grouped observations in life studies. *Biometrics* 1977;**33**:463–70.

21. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;**11**:561–70.

22. Hernán MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol* 2006;**98**:237–42.

23. Orellana L, Rotnitzky A, Robins JM. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: Main Content. *Int J Biostat* 2010;**6**: Article 7.

24. Orellana L, Rotnitzky A, Robins JM. Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part II: Proofs and Additional Results. *Int J Biostat* 2010;**6**:Article 8.

25. Little RJ, D'Agostino R, Cohen ML *et al*. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012;**367**:1355–60.

26. May MT, Ingle SM, Costagliola D *et al*. Cohort profile: Antiretroviral Therapy Cohort Collaboration (ART-CC). *Int J Epidemiol* 2014;**43**:691–702.

27. Egger M, May M, Chene G *et al*. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. *Lancet* 2002;**360**:119–29.

28. The HIV-CAUSAL Collaboration, Ray M, Logan R *et al*. The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals. *AIDS* 2010;**24**:123–37.

29. Kitahata MM, Rodriguez B, Haubrich R *et al*. Cohort profile: the Centers for AIDS Research Network of Integrated Clinical Systems. *Int J Epidemiol* 2008;**37**:948–55.

30. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;**168**:656–64.

31. Garcia-Albeniz X, Chan JM, Paciorek A *et al*. Immediate versus deferred initiation of androgen deprivation therapy in prostate cancer patients with PSA-only relapse. An observational follow-up study. *Eur J Cancer* 2015;**51**:817–24.

32. Robins JM, Hernán MA, Siebert U. Effects of multiple interventions. In: Ezzati M, Lopez AD, Rodgers A, Murry CJL (eds). *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Geneva: World Health Organization, 2004.

33. Taubman SL, Robins JM, Mittleman MA, Hernán MA. *Alternative Approaches to Estimating the Effects of Hypothetical Interventions*. Alexandria, VA: American Statistical Association, 2008.

34. Young JG, Cain LE, Robins JM, O'Reilly EJ, Hernán MA. Comparative effectiveness of dynamic treatment regimes: an application of the parametric g-formula. *Stat Biosci* 2011;**3**:119–43.

Commentary: Should the analysis of observational data always be preceded by specifying a target experimental trial?

Vanessa Didelez

School of Mathematics, University of Bristol, Bristol, UK. E-mail: Vanessa.Didelez@bristol.ac.uk

The recent paper by Cain *et al*.[1] deals with the important practical question of when to switch antiretroviral therapy for HIV-infected individuals after virological failure.

However, the real significance of the paper is in advocating, and demonstrating the feasibility of, what we may want to adopt as a general principle: analysing

observational data by first specifying, as precisely as possible, and then emulating the ideal experimental trial for the question at hand, so as to guide the analysis of the observational data. This is not the first instance of such an analysis,[2] but they are still very (and too) rare. As demonstrated comprehensively by the authors,[1] clarifying the ideal experimental trial steers the practical decisions that need to be made for the statistical analysis and its interpretation.

Why is it important to specify and emulate the ideal experimental trial for the question at hand even if we 'only' have observational data?

Facilitating communication

Methods that deal with observed (and sometimes unobserved) confounders—such as propensity scores / inverse probability weighting, g-computation, instrumental variables, Monte-Carlo simulation providing marginal standardized effects etc.—may in themselves be quite complicated and daunting, the target of inference and hence the interpretation of the outcome is made much clearer and simpler when formulated in terms of an ideal randomized trial. For instance, in Cain *et al.*[1] we see that the results can be presented in a few plots with two survival curves each, even though a complex statistical analysis—combing 'cloning', artificial (as well as other types of) censoring, inverse probability weighting, time-exposure interactions, standardization etc.—has been used. In fact, causal frameworks, such as counterfactuals, structural equations or do-calculus, can be regarded as languages for formulating an ideal interventional trial; however, they sometimes develop a life of their own that leads away from practically relevant questions.

Asking meaningful questions

It forces us to ask meaningful questions and choose meaningful targets of inference that lead to practically useful results. A large aspect of this is that it forces us to consider feasible interventions. Many observational studies, for instance, consider the "causal effect" of body mass index (BMI) on health outcomes—how would this be reformulated in terms of an ideal experimental trial? We cannot 'assign' participants to different BMI values, so what is the meaning of such an analysis? Instead we may want to look at the effect of specific exercise regimens vs dietary changes vs gastric stapling. Similarly, many studies consider the 'causal effect' of birthweight on long-term outcomes; however, we cannot change the birthweight of a newborn at will; different realistic interventions that may change birthweight may do so in extremely different ways with very different effects. It may therefore be more relevant for practice to look at what behaviours and exposures affect birthweight as well as subsequent outcomes.

Understanding assumptions

It helps with assessing, and justifying or refuting the assumptions required, to obtain causal conclusions from observational data, such as 'no unmeasured confounding'. It allows a concrete comparison between: (i) what randomization, data and decisions we would have in the experimental setting; with (ii) the (possible lack of) randomization, data and decisions we actually have in the observational setting. In Cain *et al.*,[1] the concrete example of the importance, but lack of, data on adherence and resistance illustrates this point.

Avoid pitfalls with time dependent data

It is even more important to take this approach when faced with time-dependent data and especially time-dependent exposures or treatments. As becomes clear in the example of Cain *et al.*[1] even if we were to randomize subjects to one or the other switching strategy, there can be periods of time during follow-up where a given participant's history conforms with both strategies. In an actual randomized trial, randomization will ensure that these types of subjects are at the outset comparable in both groups; in an observational trial, Cain *et al.* suggest counting them into both groups by 'cloning' them (alternatively one could randomly assign them to just one group, but that would lead to an inefficient use of the already limited data). Mistakes like assigning to a control group all subjects who dropped out or died before actually being exposed or treated will then automatically be avoided.

Guiding the statistical analysis

It also forces us to use a statistical analysis that produces quantities that are analogous to those obtained from an experimental trial. In a randomized controlled trial with two groups, it would be obvious to just plot the two survival curves where covariates would not normally be included. Hence, when analysing observational data we should target the same marginal effects, and not condition on covariates just because we need to adjust for them as potential confounders—this naturally leads to the use of marginal structural models[3] (note that Cain *et al.*[1] plot the survival curves standardized by baseline covariates). An alternative would be to weight the Kaplan-Meier curves themselves,[4] though in the context of dynamic treatment strategies this approach still needs further exploration.

Some may argue that demanding, as a rule, that any analysis of observational data must be preceded by the specification of and emulation of an experimental trial, could be very limiting; for instance, one may want to look at 'effects' of BMI or birthweight on various outcomes in order to 'generate hypotheses'. My suggestion is that if this is the case, then it should be clearly stated and the discussion should explicitly address how to translate the results into more practical questions for future research. On the other hand, we may find that specifying a target experimental trial could be very fruitful and enrich analyses of observational data as it may lead, for example, to new approaches to defining sensible exposures or treatments or better presentations of results. Clearly, observational and other non-experimental studies have many advantages over controlled randomized trials,[5] but as an experiment, let's try to complement their analysis by specifying an ideal target experimental trial for the research question at hand and see to what new ideas and findings this leads!

References

1. Cain LE, Saag MS, Petersen M *et al*. Using observational data to emulate a randomized trial of dynamic treatment switching strategies: an application to antiretroviral therapy. *Int J Epidemiol* 2016;**45**:2038–49.
2. Hernán MA, Alonso A, Logan R *et al*. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008;**19**:766–79.
3. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;**11**:561–70.
4. Westreich D, Cole AR, Tien PC *et al*. Time scale and adjusted survival curves for marginal structural Cox models. *Am J Epidemiol* 2010;**171**:691–700.
5. West SG. Alternatives to randomized experiments. *Curr Dir Psychol Sci* 2009;**18**:299–304.