Discussion of "Experimental Designs for Identifying Causal Mechanisms"
by Imai, Tingley, Yamamoto — read to the RSS on 14 March 2012

Dr Vanessa Didelez, School of Mathematics, University of Bristol

vanessa.didelez@bristol.ac.uk

The importance of experimental design for causal inference is twofold. It can guarantee crucial assumptions, e.g. actual randomisation allows identification of average causal effects. A careful design also clarifies, almost defines, the target of inference — this is especially relevant in the context of sometimes woolly notions of "causal mechanisms".

Imai et al. (2012) consider in/direct effects involving $Y(t, M(t'))$. Setting treatment to t and t' for the same unit is genuinely counterfactual. Consequently, while their designs improve on the single-experiment, they cannot avoid untestable assumptions. Furthermore, do the proposed designs *define* the target of inference? The additional experiment in the parallel designs really targets the controlled direct effect, and the two require linking by untestable Assumption 5. However, the proposed cross-over designs clearly target $Y(t, M(t'))$, and under untestable Assumption 7 this comes close to observing $Y(t, M(t'))$ itself.

A different type of design is sometimes possible and clarifies the causal parameter in a decision theoretic context (Didelez et al., 2006), namely when we can manipulate the mediator, without controlling it, (almost) as if treatment were at two different values for the same unit. For example double-blind placebo-controlled studies: these target the direct effect of active ingredient not mediated by patient's / doctor's expectation. Crucially the mediator (the expectations) is not (and cannot) itself be controlled, but the design guarantees that it arises as under "drug taken". One can easily think of variations addressing the indirect effect, here the placebo effect. This type of design seems feasible whenever "treatment" comprises different aspects that could — with a bit of imagination — be separated out. Robins and Richardson (2010) use similar examples and (possibly hypothetical) interventions in augmented DAGs to discuss when $Y(t, M(t'))$ can be regarded as manipulable quantity. Does this mean that we observe $Y(t, M(t'))$ itself? Not necessarily, the design fails when there are post-treatment confounders (even if observed) of M and Y; in the placebo-controlled trial this is known as "unblinding" e.g. by side-effects of the active ingredient.

Looking at typical applications, it will be rare that cross-over or placebo-type designs can be used. The interest in causal parameters based on $Y(t, M(t'))$ therefore remains a mystery to me — what *practical* questions does it help answer that simpler approaches (causal chain, controlled effects) do not? If effect modification is the main problem, we should maybe direct more attention to investigating effect modification and design experiments accordingly.

References

Didelez, V., Dawid, A.P., Geneletti, S. (2006). Direct and indirect effects of sequential decisions. In: Proceedings of the 22nd UAI Conference (eds. R. Dechter and T. Richardson), 138-146. UAI Press, Corvallis, Oregon.

Robins, J.M., Richardson, T.S. (2010). Alternative graphical causal models and the identification of direct effects. In: Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures (eds. P. Shrout, K.M. Keyes, K. Ornstein), 103-158. Oxford University Press, New York.