Causal inference methods for combining randomized trials and observational studies: a review

Bénédicte Colnet, Inria Parietal

Abstract

With increasing data availability, treatment causal effects can be evaluated across different dataset, both randomized trials and observational studies. Randomized trials isolate the effect of the treatment from that of unwanted (confounding) co-occurring effects. But they may be applied to limited populations, and thus lack external validity. On the opposite large observational samples are often more representative of the target population but can conflate confounding effects with the treatment of interest. In this work, we review the growing literature on methods for causal inference on combined randomized trial and observational studies, striving for best of both worlds. We first discuss identification and estimation methods that improve generalizability of randomized clinical trials (RCTs) using the representativeness of observational data. Classical estimators include weighting, difference between conditional outcome models, and double robust estimators. We then discuss methods that combining RCTs and observational data to improve the (conditional) average treatment effect estimation, handling possible unmeasured confounding in the observational data. We compare the main methods using a simulation study and real world data to analyse the effect of Tranexamic Acid on the mortality rate in major trauma patients.