Thesis

Alexandros Rekkas

2022-05-15

Contents

Preface

Thesis

Introduction

In order to provide—on average—the most current medical care doctors are advised to align their clinical practice with the results of well-conducted clinical trials, or the aggregation of the results from multiple such trials [Greenfield 2007]. This approach implicitly assumes that all patients eligible for treatment similarly experience the benefits and harms of treatment of the reference trial population. Therefore, at a certain point, the accurate estimation of these average effects became crucial, transforming clinical trials from tools for assessing causality into tools for predicting patient-level treatment effects. When the strong positive overall effects derived from clinical trials could not be achieved in medical practice, the problem was attributed to the reference trial population being too narrow and not representing the "average" patient requiring treatment. Therefore the need for more pragmatic clinical trials incorporating broad patient populations was highlighted [Treweek, Trials 2009; …].

The wider clinical trial populations ensure that overall results will be generalizable to the "real-life" patients. However, generalizability comes at a cost: wider range of included patients means higher variability of measured characteristics, therefore higher variability in disease severity is observed, which, in turn, translates to higher variability of observed treatment effect sizes. In short, the estimated treatment effect derived from such clinical trials is often an average of heterogeneous treatment effects and, as such, is not applicable to most patient subgroups. This means that a positive average treatment effect estimated from a clinical trial very often is only evidence that some of the enrollees benefitted from the intervention under study. If, however, the intervention is linked to a serious adverse event, treating everyone would result in serious harms for many patients, despite the positive overall effect.

The identification and quantification of treatment effect heterogeneity (HTE), i.e. the non-random variability in the magnitude or direction of a treatment effect across the levels of a covariate (single or combination of patient attributes) [refs], is crucial for guiding medical decision-making. Despite HTE being widely anticipated, its evaluation is not straightforward. At the individual level HTE is unobservable as it requires knowledge of patient-level outcomes under all possible treatment assignments (fundamental problem of causal inference). The most common solution is, then, to assign the patient to a subgroup with similar

anticipated treatment effect, assigning the subgroup-level effect estimate to the individual. Subgroup analyses have often been considered for the evaluation of HTE, where the clinical trial population is split into sub-populations based on the levels of a certain covariate, one at a time. However, as clinical trials are most often adequately powered to detect a certain overall effect size, these subgroup analyses usually are underpowered. This can lead to falsely claiming absence of HTE or overestimating its magnitude [refs]. In addition, contrary to subgroup analyses, patients differ with regard to many baseline covariates simultaneously [Kent, BMJ 2018]. However, evaluation of two-way or higher order interactions becomes underpowered at an explonential rate.