Assessing and Reporting Heterogeneity of Treatment Effect in Clinical Trials

Kent, David
Tufts Medical Center Inc.

Recent evidence shows that the results of randomized clinical trials might not apply to individual patients in a straightforward way, even to those within the trial. While randomization ensures the comparability of treatment groups overall, there remain important differences between individuals in each treatment group that can dramatically affect the likelihood of benefiting from or being harmed by a therapy. Averaging effects across such different patients can give misleading results to physicians who care for individual, not average, patients. The conventional way of addressing this "heterogeneity of treatment effect" (HTE) is through subgroup analysis, in which patients are serially divided into groups based on potentially influential characteristics "one variable at a time". Because patients have many attributes that might influence the risk of the outcome and the potential for benefit, subgroups are prone to spurious false positive results. Another problem with one-variable-at-a-time subgroup analysis is that they compare groups of patients that differ only on a single variable, and are thus more similar than different in terms of their outcome risk--a mathematical determinant of treatment benefit. Real patients, however, differ from one another in multiple variables simultaneously. The outcome risk within a clinical trial often varies tremendously between large subgroups of patients that are not separable with one-variable-at-a-time analyses. The investigators in this application recently proposed a framework for evaluating HTE that addresses many of these concerns. We have proposed that the distribution of risk in clinical trial populations routinely be examined through the application of a multivariable risk model, that risk-based HTE be routinely tested, and that other subgroup analyses be explicitly labeled either as primary subgroup analyses (well motivated by prior research and performed to inform clinical practice) or secondary subgroup analysis (performed to inform future research). While the theoretical reasons to prioritize the reporting of risk-based HTE are compelling, the approach has not been broadly tested, and empirical evidence from actual clinical trials remains anecdotal. This application thus proposes to refine and evaluate the approach on a set of approximately 30 large, diverse cardio- and cerebrovascular clinical trials. The goal of this study is to obtain strong preliminary empirical evidence of both the feasibility and value of this approach, and to better understand any important limitations.

RELEVANCE
This application focuses on advancing analytic methods for comparative effectiveness research. In particular, we propose an empirical evaluation of a proposed approach for assessing variation in treatment effects in randomized clinical trials, so called "heterogeneity of treatment effect." Whereas a good doctor or clinician seeks to combine their acumen, expertise and experience with their knowledge of an individual patient's circumstances and needs "evidence" for clinical practice is often derived on the average effects of treatments in large groups of patients. The research we propose is intended to address the very real incongruence between the overall effects of a treatment in a study population (the summary result of a trial) and information regarding the anticipated risks and benefits in an individual patient necessary to support patient-centered decision making. To do this, we have previously proposed a framework that suggests that risk models that can take into account multiple important patient attributes simultaneously be used routinely in the analysis of randomized clinical trials. While this approach has theoretical appeal, its value has not been systematically demonstrated. We therefore propose the first systematic evaluation of the proposed framework by analyzing approximately 30 randomized clinical trials. The results of this study will provide preliminary evidence as to whether this novel approach can and should be broadly and routinely applied.