ABSTRACT

Many clinical trials are structured with multiple stages, where data analysis is conducted after each stage to inform subsequent participant recruitment and treatment allocation. This adaptive approach allows for early elimination of ineffective treatments or targeted recruitment of subpopulations showing potential benefits. Analyzing such trials presents challenges as the data is utilized twice: first for selecting the design and null hypothesis, and then for testing the chosen hypothesis using the data generated under the selected design. Classical statistical methods are inadequate as they require pre-specified data generating mechanisms and null hypotheses. Existing solutions are often limited in scope, tailored to specific designs. In this work, we propose a general framework capable of handling diverse designs and adaptive choices. Our approach leverages post-selection inference principles to develop a selective randomization p-value. Notably, it does not necessitate assumptions about the distribution of outcomes or covariates, or the dependency structure among participants. We demonstrate that our method enhances statistical power compared to other valid tests while maintaining control over the selective type-I error in simulated data and hypothetical clinical trials.