Literature Review of Statistical Methods Comparisons Through Simulations When Using External Control Arm for Regulatory or HTA Submissions

Authors: Binglin Yue[1]; Kaifeng Lu[1]; Sheng Xu[1]; Shanmei Liao[1]

Affiliations: [1]BeiGene (Beijing) Co, Ltd., Beijing, China and BeiGene USA, Inc, San Mateo, CA, USA.

ABSTRACT

OBJECTIVES: Using a historical control or external control arm (ECA) to augment or replace a concurrent control arm in a randomized trial is a hot topic given the challenge of patient recruitment in rare diseases or during COVID-19 pandemic. The FDA released draft guidance in 2021 on effectiveness and safety submissions using real-world evidence. While the guidance focuses mainly on elements of study design and data source selection, there is a lack of consensus in the selection of appropriate statistical methods when constructing an ECA. This study discusses rigorous statistical methodology for ECA-supported trials in regulatory or HTA submissions.

METHODS: Targeted literature reviews of statistical simulations comparing methods for ECA in statistical journals were performed. The articles compared commonly used ECA-construction and analysis methods were selected and summarized, including but not limited to propensity score (PS)-based matching, weighting, and stratification, and PS plus Bayesian integrated approaches.

RESULTS: Type I error, power, bias, and coverage probability are common criteria used to compare different methods. When imbalances only exist in known baseline covariates and the outcome distributions are the same between the trial concurrent control and ECA, the PS method alone or paired with commensurate prior yield almost unbiased estimates, good Type I errors, and coverage probability. PS plus Bayesian approaches have wider interval width and lower power compared with PS-only methods. When there is a change in the outcome distribution over time, the PS (matching or IPTW) and commensurate prior integrated methods yield the smallest biases among all methods.

CONCLUSIONS: PS and Bayesian integrated methods outperformed the PS-only methods in terms of bias and Type I error when outcome distribution changed with current trial control. A “sweet spot” that balances all criteria through trial-specific simulations could provide the ideal setting of trial analyses plan based on specific trial design and scenarios.