Evaluation of Matching-Adjusted Indirect Comparison: Applied Treatment Comparison Implemented by a Resampling Method

Jianmin Wang, Dawn Odom, Costel Chirila, Qingyao Zheng

RTI Health Solutions, Research Triangle Park, NC, United States

BACKGROUND

- In practice, patient-level data are rarely available for all studies and therefore it is unlikely that the interaction can be found when conducting MAICs. However, the presence of interactions does not necessarily lead to self-contradicting conclusions—this result depends on the direction and magnitude of these interactions across the trials.

- In these simulated datasets, the effects of the interaction between baseline characteristics and treatment on the efficacy endpoint was designed to portray the nuances that can be found when conducting MAICs. However, the presence of interactions does not necessary lead to self-contradicting conclusions—this result depends on the direction and magnitude of these interactions across the trials.

- In this paper, patient-level data are rarely available for all studies and therefore it’s unlikely that the interaction can be formally tested.

OBJECTIVE

- This study evaluates the performance of the MAIC method proposed by Malangone and Sherman, which is implemented by a resampling/bootstrapping technique. We focused on the special case, where there is an interaction between baseline characteristics and treatment.

METHODS

Data

- Patient-level data sets were generated to simulate two clinical trials, with 1,800 patients each. The first data set (Trial 1) comprised treatment A and placebo, and the second data set (Trial 2) comprised treatment B and placebo.

- Other variables included in both data sets were survival time, its corresponding censoring indicator, and two baseline characteristic variables (categorical age group: young/old; sex: female/male). Survival time follows the exponential distribution in both trials.

Statistical Methods

- The SAS programs detailed in Malangone and Sherman were implemented for the MAIC analysis. First, MAIC was applied to a scenario in which only ASD were available from Trial 1 and only IPD were available from Trial 2.

- Hazard rate (HR) of treatment A versus placebo using the Cox proportional hazard model

- Using PROC SURVEYSELECT, 1,000 simple random samples (selection with equal probability and without replacement) of size 180 were drawn from Trial 2 patient data such that each sample had the same distribution of the baseline variables observed in Trial 1. Using the first dataset’s cross-tab frequency percentages, the resampling scheme used in the Malangone and Sherman paper was followed with no stratification by treatment.

RESULTS

- The median (percentile) of the 1,000 bootstrapped HR estimates was used as the bootstrapped HR. A 95% confidence interval (CI) was obtained from the 2.5th percentile and 97.5th percentile of the 1,000 HR estimates. Bootstrapped median survival along with its 95% CI was determined similarly to HR.

- The aggregate HR and its 95% CI of Trial 1 were compared with the bootstrapped median HR and 95% CI obtained from bootstrapped Trial 2 data. The comparison of HR (treatment A vs. placebo, and treatment B vs. placebo) was conducted by examining CI, as well as the Malangone and Sherman paper.

- Subsequently, the roles of two data sets were switched, and the MAIC analysis was applied once again (i.e., IPD for Trial 1 and ASD for Trial 2).

- The summary statistics of the two simulated data sets are shown in Table 1.

- Note that the two simulated trials were constructed such that there was a higher percentage of patients that were female and younger in Trial 2 and a higher percentage of patients that were male and older in Trial 1.

- In both trials, interactions between baseline characteristics and treatments were incorporated such that differential treatment effects across baseline strata were present.

- Specifically in Trial 1, the active drug treatment A works the best among female patients in the young age group and it is least effective, either through placebo, in males in the old age group (Figure 1). Treatment B in Trial 2, however, interacts with age group and sex in an opposite trend. As shown in Figure 1, Treatment B works the best among male patients in the old age group, and it is least effective among female patients in the young age group.

- The hazard is constant over time due to the exponential distribution; therefore, baseline hazard was shown in Figure 1 for illustrative purposes.

- The distribution of baseline characteristic for both trials before and after IPD and ASD preprocessing for each scenario, the matching procedure resulted in identical baseline characteristic distributions across the two trials.

- Median survival time (95% CI) from prematched trials and postmatched samples are presented in Table 3.

- Hazard ratios (95% CI) from matched trials and postmatched samples are presented in Table 4.

- Applying MAIC to the two scenarios resulted in the following (Table 4, Figure 2).

- In Scenario 1: When the first data set (Trial 1) provided ASD

- The aggregated HR (95% CI) for treatment B versus placebo was 0.237 (0.205-0.273). The bootstrapped HR (95% CI) for treatment B versus placebo was 0.226 (0.204-0.252). Because the 95% CI’s do not overlap and HR for treatment B versus placebo is lower than HR for treatment B versus placebo in prematched data, conclusion is treatment A is more effective than treatment B in terms of extending survival time.

- In Scenario 2: When the second data set (Trial 2) provided ASD

- The bootstrapped HR (95% CI) for treatment A versus placebo was 0.489 (0.390-0.612) and the ASD HR (95% CI) for treatment B versus placebo was 0.237 (0.205-0.273), using the same reasoning, the conclusion is that treatment B is more effective than treatment A in terms of extending survival time.

DISCUSSION

- In these simulated datasets, the effects of the interaction between baseline characteristics and treatment on the efficacy endpoint was designed to portray the nuances that can be found when conducting MAICs. However, the presence of interactions does not necessarily lead to self-contradicting conclusions—this result depends on the direction and magnitude of these interactions across the trials.

- Note that the two applications of MAIC produce different conclusions because in each case, the bootstrapped sample is drawn to match the characteristics of a specific trial and because there is an interaction effect between treatment and baseline characteristics that is in opposite direction for the two trials.

- In practice, patient-level data are rarely available for all studies and therefore it’s unlikely that the interaction can be formally tested.

CONCLUSIONS

- The method proposed by Malangone and Sherman is an interesting addition to the MAIC field, but results will reflect the underlying characteristics of the trial chosen for comparison.

- In the presence of interactions in the opposite direction between baseline characteristics and treatment across trials, MAIC may produce conflicting results if the same trials are included, but their roles (i.e., provider of ASD) are changed within the analysis. Therefore, it is critical to identify the specific populations for which inferences apply.

- Further research should be conducted to better understand the conditions under which this method is best suited.

REFERENCES


CONTACT INFORMATION

Jianmin Wang, PhD

Jianmin.Wang1@rti.org

Director, Biostatistics

RTI Health Solutions

200 Park Office Drive

Research Triangle Park, NC 27709

Phone: +1.919.541.7222

Fax: +1.919.541.7221

Presented at: ISPOR 20th Annual International Meeting

Fax: +1.919.541.7222

Phone: +1.919.541.6986

Research Triangle Park, NC 27709

BioStat Solutions

Director, Biostatistics

Table 1. Simulated Subject Characteristics Descriptive Statistics

Table 2. Bootstrap Matching of Age Group and Sex Across Trials

Table 3. Median Survival Time in Months (95% CI) From Prematching and Postmatching

Table 4. Hazard Ratio (95% CI) From Prematching and Postmatching

Figure 1. Baseline-Hazard Stratified by Treatment and Baseline Characteristics in Simulated Trials

Figure 2. Hazard Ratios (95% CI) for Each Scenario

Figure 3. Survival time (95% CI) for treatment A in Trial 1 and treatment B in Trial 2