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

Jianmin Wang, Dawn Odom, Costel Chirila, Qingyao Zheng
RTI Health Solutions, Research Triangle Park, NC, United States

BACKGROUND

- Individual patient-level data (IPD) from all studies are rarely available for indirect treatment comparisons; therefore, they are typically conducted using aggregate study-level data (ASD).
- For special cases where IPD are available from one study, but only ASD are available from another study, matching-adjusted indirect comparison (MAIC) has been proposed as an alternative to the anchored indirect treatment comparison methods because of its potential to adjust for differences in baseline characteristics between trials.[1]
- MAIC can be implemented using at least two different techniques: resampling and weighting. This study is focused on the resampling method.

OBJECTIVE

- This study evaluates the performance of the MAIC method proposed by Malangone and Sherman,[2] 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[2] 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.
  - ASD available from Trial 1 were as follows:
    - Cross-tabulations (i.e., count and percentages) of age group and sex
    - Median survival time estimate using the Kaplan-Meier method
    - Hazard ratio (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[2] paper was followed with no stratification by treatment.
  - The median (50% percentile) of the 1,000 bootstrapped HR estimates was used as the bootstrapped HR. A 95% confidence interval (CI) was obtained from the 2.5% percentile and 97.5% 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 its 95% CI obtained from bootstrapped Trial 2 data. The comparison of HR (treatment A vs. placebo) and HR (treatment B vs. placebo) was conducted by examining CIs, as seen in the Malangone and Sherman[2] 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).

RESULTS

- The summary statistics of the two simulated data sets are shown in Table 1.
- Note that the simulated trial datasets were constructed such that there was a higher percentage of patients that were female and younger in Trial 1 and a higher percentage of patients that were male and older in Trial 2.
- 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 efficacious, but still better than placebo, in male patients in the old age group (Figure 1). Treatment B in Trial 2, however, interacts with age group and sex in an opposite trend. As is 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 matching were shown in Table 2. For each scenario, the matching procedure resulted in identical baseline characteristic distributions across the two studies.
- Median survival time (95% CI) from prematched trials and postmatched samples are presented in Table 3.
- Hazard ratios (95% CI) from prematched trials and postmatched samples are presented in Table 4.
- Applying MAIC to the two scenarios resulted in the following (Table 4, Figure 2):
  - SCENARIO 1: When the first data set (Trial 1) provided ASD
    - The aggregate HR (95% CI) for treatment A versus placebo was 0.283 (0.246-0.325); the bootstrapped HR (95% CI) for treatment B versus placebo was 0.586 (0.466-0.740). Because the 95% CIs do not overlap and HR for treatment A versus placebo is lower than HR for treatment B versus placebo, the conclusion is that treatment A is more efficacious than treatment B in terms of extending survival time.
  - SCENARIO 2: When the second data set (Trial 2) provided ASD
    - The bootstrap 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 efficacious than treatment A in terms of extending survival time.

Table 1. Simulated Subject Characteristics Descriptive Statistics

Baseline Characteristics	Trial 1 (N = 1,800) n (%)	Trial 2 (N = 1,800) n (%)
Age group		
Young	1,500 (83)	300 (17)
Old	300 (17)	1,500 (83)
Sex		
Female	1,200 (67)	600 (33)
Male	600 (33)	1,200 (67)
Age group stratified by sex		
Young age group, female	1,000 (56)	100 (6)
Young age group, male	500 (28)	200 (11)
Old age group, female	200 (11)	500 (28)
Old age group, male	100 (6)	1,000 (56)
Treatment		
Active[a]	915 (51)	892 (50)
Placebo	885 (49)	908 (50)
Censor indicator		
Censored	712 (40)	713 (40)

[a] Patients took treatment A as active treatment in Trial 1 and took treatment B as active treatment in Trial 2.

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

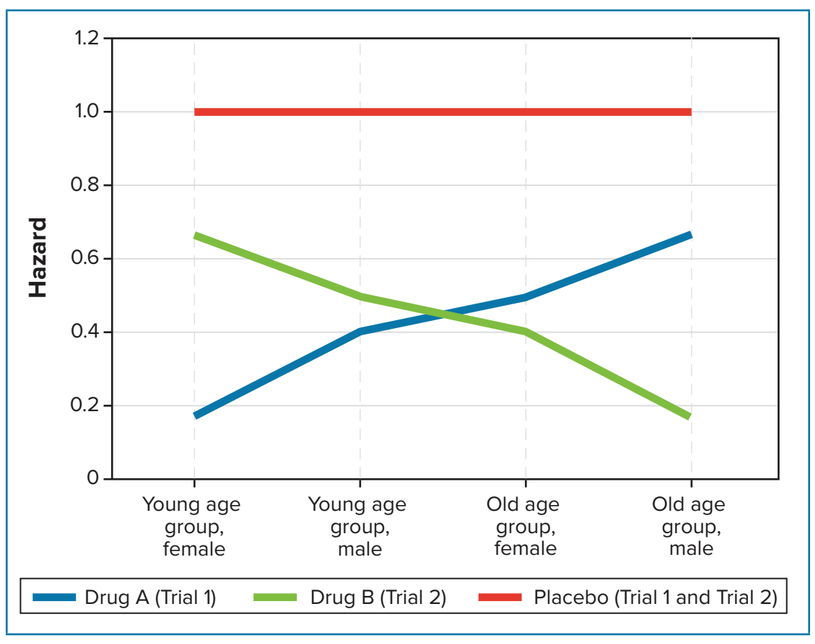


Table 2. Bootstrap Matching of Age Group and Sex Across Trials

Baseline Characteristics	Trial 1 Prematched (N = 1,800) (%)	Trial 2 Prematched (N = 1,800) (%)	Trial 1 Postmatched (N = 180 per sample, 1,000 samples) (%)	Trial 2 Postmatched (N = 180 per sample, 1,000 samples) (%)
Age group				
Young	(83)	(17)	(17)	(83)
Old	(17)	(83)	(83)	(17)
Sex				
Female	(67)	(33)	(33)	(67)
Male	(33)	(67)	(67)	(33)
Age group stratified by sex				
Young age group, female	(56)	(6)	(6)	(56)
Young age group, male	(28)	(11)	(11)	(28)
Old age group, female	(11)	(28)	(28)	(11)
Old age group, male	(6)	(56)	(56)	(6)

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

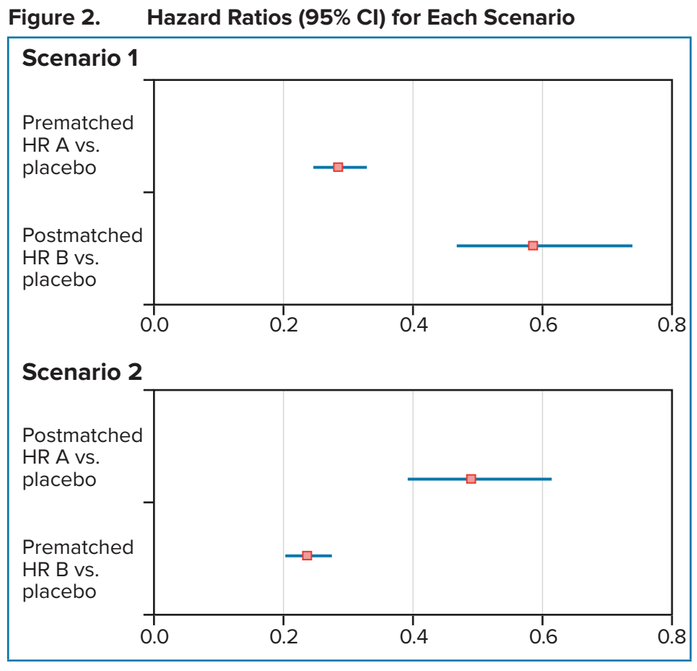
	Active (Treatment A)	Placebo	Overall
Scenario 1 (ASD from Trial 1, IPD from Trial 2)			
Trial 1 prematched	25.06 (22.16-28.49)	6.86 (5.84-7.46)	12.95 (11.64-14.22)
Trial 2 postmatched	10.24 (8.26-12.28)	6.92 (5.01-8.22)	8.24 (7.24-9.31)
Scenario 2 (ASD from Trial 2, IPD from Trial 1)			
	Active (Treatment B)	**Placebo**	**Overall**
Trial 2 prematched	28.85 (24.71-33.61)	6.57 (5.59-7.21)	13.85 (12.64-16.00)
Trial 1 postmatched	12.85 (10.76-15.16)	5.89 (4.96-7.74)	8.64 (7.73-10.76)

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

	HR (95% CI) of Active Treatment Versus Placebo	Conclusion[c]
Scenario 1 (ASD from Trial 1, IPD from Trial 2)		
Trial 1 prematched	0.283 (0.246-0.325)[a]	Treatment A is better than treatment B
Trial 2 postmatched	0.586 (0.466-0.740)[b]	
Scenario 2 (ASD from Trial 2, IPD from Trial 1)		
	HR (95% CI) of Active Treatment Versus Placebo	**Conclusion[c]**
Trial 1 postmatched	0.489 (0.390-0.612)[a]	Treatment B is better than treatment A
Trial 2 prematched	0.237 (0.205-0.273)[b]	

[a] Treatment A versus placebo.
[b] Treatment B versus placebo.
[c] The conclusion is based on nonoverlapping CIs.

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



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 necessary 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 directions for the two trials.
- In practice, patient-level data are rarely available for all studies and therefore it is unlikely that the interaction can be formally tested.

CONCLUSIONS

- The method proposed by Malangone and Sherman[2] 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 population to which inferences apply.
- Further research should be conducted to better understand the conditions under which this method is best suited.

REFERENCES

1. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol. 1997;50(6):683-91.
2. Malangone E, Sherman S. Matching-adjusted indirect comparison analysis using common SAS 9.2 procedures. SAS Global Forum 2011. Available at: http://support.sas.com/resources/papers/proceedings11/228-2011.pdf. Accessed December 1, 2014.

CONTACT INFORMATION

Jianmin Wang, PhD
Director, Biostatistics

RTI Health Solutions
200 Park Offices Drive
Research Triangle Park, NC 27709

Phone: +1.919.541.6986
Fax: +1.919.541.7222
E-mail: jmwang@rti.org