Paper 228-2011

Matching-Adjusted Indirect Comparison Analysis Using Common SAS® 9.2 PROCEDURES

Elisabetta Malangone, Analytica International, New York, NY, USA
Steven Sherman, Analytica International, New York, NY, USA

ABSTRACT

This paper presents a novel matching-adjusted approach to indirectly compare survival estimates for competitive treatment options.

Using patient-level data for the treatment arm and summary patient characteristics and survival outcomes for the comparator, matching variables prognostic for survival are chosen. A program involving an extension of a common SAS® 9.2 procedure, PROC SURVEYSELECT, is applied to select 1000 random repeated sub-samples from the original population with the same distribution of matched variables. Using Output Delivery System (ODS) and survival analysis procedures, the survival estimates are computed for each bootstrapped sample and a 95% confidence interval (CI) is inferred around the mean of the sampled survival estimates. These outcome measures may then be compared to those reported for the comparator treatment.

In the absence of head-to-head randomized controlled trial (RCT) data, comparative effectiveness of treatment outcomes can be evaluated using a matching-adjusted indirect comparison, which accounts for observed differences between populations and results in an effect of treatment exposure on survival outcomes less likely due to confounders.

INTRODUCTION

In the absence of head-to-head clinical trial data, examining the comparative effectiveness of competing therapies often presents a challenge to researchers. There are several methods for indirect comparisons of clinical outcomes. Most approaches, such as naïve and adjusted matching, use aggregate data from independent studies that have a common comparator. For example, researchers may compare studies of treatment X vs. placebo and treatment Y vs. placebo in order to make inferences about the relative efficacy of X vs. Y.[1] Unfortunately, such methods do not adjust for differences in baseline patient characteristics that can potentially confound the observed outcomes.

In another situation, researchers may have access to patient-level data from one study, but only aggregate data from another. This situation provides an opportunity to align the distributions of patient characteristics on key variables between studies by using patient-level data matched to available aggregate data (termed a “matching-adjusted” approach). Matching-adjusted techniques can be useful for survival estimate comparisons, with methods currently available to indirectly compare absolute survival estimates (e.g., median overall (OS) and progression-free survival(PFS)) and relative survival estimates (e.g., hazard ratios).[2]

The objective of this paper is to introduce a novel matching-adjusted indirect comparison methodology that provides adjusted survival estimate comparisons between two treatments in the absence of head-to-head clinical trial data.

METHODS

In cases where patient-level data are available for one study and only summary data are available for another, the proposed matching-adjusted technique can be performed in 6 steps:

Step 1: Selection of comparator study

Investigators should initiate a careful search for a comparator study that matches the inclusion/exclusion criteria present in the index study (i.e., the study for which patient-level data are available). It is also possible to balance two studies with differing inclusion/exclusion criteria when the equivalent patient types of the comparator study population are present as a subset of the index study. In some cases it may be possible to contact the investigators of a potential comparator study to perform subset analyses as necessary when populations do not sufficiently overlap.

Step 2: Alignment of inclusion/exclusion criteria

If the inclusion/exclusion criteria between the index and selected comparator study are not balanced, and the patient

types of the comparator study are present as a subset of the index study, additional exclusions may be made to the patient-level data to further align the two populations. Following this step, studies will include similar patient types, although some differences in the distributions of these patient characteristics should be expected.

Step 3: Selection of matching variable(s)

After a preliminary investigation of the distributions and potential prognostic effects of common data elements available for the two studies, variables that differ in their distributions and are likely to be predictive of the patient risk of outcome should be considered as potential matching variables.

Step 4: Distribution matching of patient-level data to aggregate data

Once inclusion/exclusion criteria are aligned, the resulting pre-match population of the index study serves as a pool from which repeated random samples of the patient-level data without replacement are selected. Using SAS version 9.2, the SURVEYSELECT procedure is used to take 1000 repeated subsets of the pre-match population that mimic the distribution of the comparator study for the pre-selected matching variable(s). Each of the resulting samples will have distributions for the matched variable equal to those from the comparator study.

General SAS syntax that can be used to derive repeated random samples matched to aggregate data:

```
PROC SURVEYSELECT DATA=in_dataset
SEED= {value} METHOD=SRS REP=1000 SAMPSIZE= (x-value, y-value, z-value)
OUT=out_dataset;
STRATA matching_variable;
run;
```

- `SEED option` provides the initial seed for random number generation
- `METHOD = SRS` denotes a simple random sampling methodology without replacement
- `REP = 1000` designates the number of random samples to be taken
- `SAMPSIZE` provides parameters for the number of observations for each level of the matching variable to be represented in each of the final 1000 samples that is needed to mimic the distribution observed in the comparator study
    - For example, for a matching variable categorized into three levels (low, medium, and high), the value for “low” would correspond with x-value, the value for medium would correspond with y-value, and the value for high would be entered in place of z-value

Step 5: Deriving survival estimates and 95% confidence intervals for absolute and hazard ratio measures of effect

Median Survival Estimation:

Median survival can be determined for each of the 1000 samples using the SAS LIFETEST procedure. A SAS ODS statement is used to aggregate all median survival estimates, thereby providing a distribution of all observed sample-specific outcomes. The mean of these values is calculated, which provides the value of the point estimate for the index study. The 95% CI is derived from the 2.5 and 97.5 percentiles of the survival distribution.

See SAS statements for the derivation of the median survival estimates and the corresponding bootstrapped 95% confidence intervals for matched patient-level data below:

```
ODS LISTING CLOSE;

/*Create a dataset of quartile values for the survival distributions*/
ODS OUTPUT QUARTILES=ODS_quartiles;

/*Calculate the survival outcomes for each matched sample*/
PROC SORT DATA=in_dataset; by replicate;
PROC LIFETEST DATA=in_dataset;
TIME time_to_event*event(0);
BY replicate;
RUN;
```

```
ODS OUTPUT CLOSE;
ODS LISTING;

/*Identify the median values for each matched sample*/
DATA ODS_median;
SET ODS_quartiles;
WHERE percent = 50;
RUN;

/*Derive the 2.5 and 97.5 percentiles for the outcome distribution*/
PROC SORT DATA=ODS_median; by median_survival;
PROC UNIVARIATE DATA=ODS_median;
VAR estimate;
OUTPUT OUT=bootstrapped_95_CI pctlpts= 2.5, 97.5 pctlpre=P_;
RUN;

/*Print the lower and upper bounds of the 95% confidence interval*/
PROC PRINT DATA=bootstrapped_95_CI;  run;

/*Print the mean and standard deviation of the outcome distribution*/
PROC MEANS DATA=ODS_median n mean std;
VAR estimate;
RUN;
```

An example utilizing this methodology was presented at the ESMO 35th Congress.[3] An indirect comparison of patient outcomes was performed to evaluate the efficacy of second-line everolimus vs. sorafenib treatments for metastatic renal cell carcinoma (mRCC) patients who had previously failed on sunitinib therapy. Patient level data was available for the everolimus population and aggregate data was available for the sorafenib population. In the original populations, a significant difference in the distribution of MSKCC risk score was present, which was likely to confound survival estimate comparisons. The matched-adjusted indirect comparison approach was applied to remove differences in MSKCC risk score distribution and compare the absolute survival estimates (median overall survival and median progression free survival) in the populations with similar MSKCC score. This resulted in a less biased comparison of survival estimates, not confounded by MSKCC score.

Hazard Ratio Estimation:

When a hazard ratio (HR) is reported in the comparator study, and a HR is derivable from the patient-level data with a comparator common to both studies, such as placebo, the PHREG procedure can be used to derive a distribution of HR survival outcomes from all 1000 matched samples. This procedure requires investigators to first match characteristics for not only the treatment group of interest, but also for the comparator treatment in both studies. As an example, suppose a comparator study examines drug X versus placebo, while the index study examines drug Y versus placebo. In order to have HR estimates that are comparable between the studies, not only should the patient characteristics for drug X and drug Y be similar, but characteristics for those in the respective placebo arms must be similar as well. Investigators may accomplish this by performing steps 1 through 4 above for the corresponding active treatment and placebo populations (**Figure 1**).

Using the PHREG procedure as described above, the mean of the distribution of the 1000 HR estimates serves as the point estimate for the index study versus placebo, while the 2.5 and the 97.5 percentiles provide the 95% CI. A ratio of the HRs derived from the index study and the comparator study provides the final HR estimate of treatment X versus treatment Y, whereby the placebo effects cancel each other out from each study, as shown in **Figure 1**.

See SAS statements for the derivation of hazard ratio survival estimates and the corresponding bootstrapped 95% confidence intervals for matched patient-level data below:

```
*Derive hazard ratios for each matched sample for treatment X vs. placebo;

PROC SORT DATA=All_matched_data;
BY replicate;
RUN;

PROC PHREG DATA= All_matched_data OUTEST=all_hazard_ratios;
MODEL time_to_death_event* event (0)=treatment/risklimits;
```

```
BY replicate;
RUN;

*Exponentiate hazard ratio estimates;
DATA all_hazard_ratios_exp;
SET  all_hazard_ratios;

/*Create new variables for the sample-specific exponentiated hazard ratio
estimates*/

HR=exp(treatment);
RUN;

/*Print the mean and standard deviation of the outcome distribution*/

PROC MEANS mean std DATA=all_hazard_ratios_exp;
VAR HR;
TITLE "Mean hazard ratio based on 1000 samples: treatment X vs. placebo";
RUN;

/*Derive the 2.5 and 97.5 percentiles for the outcome distribution*/
PROC SORT DATA=all_hazard_ratios_exp;
BY HR;
RUN;

DATA OS_HR_exp_95CI;
SET OS_HR_exp;
BY HR;
count+1;
IF count=25 THEN _95CI= HR;
IF count=975 THEN _95CI=HR;
RUN;

/*Print the lower and upper bounds of the 95% confidence interval*/
PROC MEANS min max data= OS_HR_exp_95CI;
VAR  _95CI;
TITLE "95% Confidence interval for hazard ratios based on 1000 samples
treatment X vs. placebo ";
RUN;
```

Figure 1: Hazard ratio derivation from indirect comparisons

An example utilizing this methodology was presented at the 13th Annual European ISPOR Congress.[5] A matching-adjusted indirect comparison analysis using individual patient data from the SATURN study and previously-published summary data from the JMEN study was conducted to simulate a head-to-head trial evaluating the relative efficacy of erlotinib versus pemetrexed as first-line maintenance therapies in non-squamous metastatic non-small cell lung cancer patients. Due to an unequal distribution of race and smoking status between the SATURN and JMEN studies, and the strong relationship these factors have with survival, race and smoking status were chosen as matching

variables. Survival comparisons after matching were less likely to be biased than survival comparisons prior to adjustment.

Step 6: Statistical inference for median survival and HR estimates

Median survival statistical inference:

The 95% CI for each study can be used to assess statistical significance of any observed difference in the point estimates from each of the studies.[4] An overlap suggests no statistically significant difference between the two estimates at the 0.05 alpha level.

HR survival statistical inference:

The standard error for the treatment X versus treatment Y HR can be calculated using the equation depicted in **Figure 2** .[6] Ninety-five percent confidence intervals can then be calculated from the standard errors for the survival HRs. As above, an overlap of the 95% CI suggests no statistically significant difference between the two treatments at the 0.05 alpha level.

$$SE(\log HR_{AB}) = \sqrt{SE(\log HR_{AC})^2 + SE(\log HR_{BC})^2}$$

Figure 2: Equation for deriving standard errors for an indirect treatment comparison of hazard ratio estimates

Note: HR_{AC} = hazard ratio of treatment X vs. placebo

and HR_{BC} = hazard ratio of treatment Y vs. placebo

LIMITATIONS

While indirect analyses are a viable means of comparing outcomes from two different treatments in the absence of a head-to-head trial, the study findings should be interpreted as those of an observational study since causation cannot be inferred from the results due to loss of randomization. Matching-adjusted indirect comparisons may also be limited by sample size and are recommended for studies which require matching on only one or two main variables.

CONCLUSION

In the absence of head-to-head RCT data, a matching-adjusted indirect comparison accounts for observed differences between populations, rendering them more comparable. This approach leads to survival estimates that are less likely to be influenced by confounding variables, thereby making the results more robust for use in an indirect comparison analysis.

REFERENCES

1. Sultana A, Ghaneh P, Cunningham D, Starling N, Neoptolemos JP, Smith CT. Gemcitabine based combination chemotherapy in advanced pancreatic cancer-indirect comparison. BMC Cancer 2008;8:192.

2. Signorovitch JE, Wu EQ, Yu AP, et al. Comparative effectiveness without head-to-head trials: a method for matching-adjusted indirect comparisons applied to psoriasis treatment with adalimumab or etanercept. Pharmacoeconomics;28(10):935-45.

3. Di Lorenzo G, Casciano C, Malangone E, Buonerba C, Sherman S, Wang X, Liu Z, De Placido S. An Adjusted Indirect Comparison of Everolimus and Sorafenib Therapy in Sunitinib-Refractory mRCC Patients Using a Robust Matching Technique. In: 35th ESMO Congress; Milan, Italy; 2010.

4. Wood M. Bootstrapped Confidence Intervals as an Approach to Statistical Inference. Organizational Research Methods 2005;vol. 8 (no. 4):454-70.

5. Casciano R, Bischoff H, Nuijten M, Malangone E, Ray J. Maintenance Erlotinib Versus Pemetrexed For the Treatment of Non-Small Cell Lung Cancer: Indirect Comparison Applying Real-Life Outcomes. In: 13th European

ISPOR Congress; Prague, Czech Republic; 2010.

6. 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:** 683-91.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the corresponding author at:

Name: Elisabetta Malangone
Enterprise: Analytica International
Address: 24 West 40th Street
City, State ZIP: New York, NY, 10018
Work Phone: 1.203.823.9286
Fax: 1.212.686.8601
E-mail: emalangone@analyticaintl.com
Web: www.analyticaintl.com