Institute for Health and Society, Medical College of Wisconsin

Datum

Biostatistics NEWSLETTER
Key Function of the CTSI
& MCW Cancer Center Biostatistics Unit

Comparing Survival Curves vs. Comparing Survival Probabilities at a Fixed Time Point

Jennifer Le-Rademacher, PhD, Assistant Professor, Division of Biostatistics

Time to an event of interest such as death or infection is often the response of interest in medical studies. In such studies, patients are followed from the beginning of the study until the event or until the end of the study period or until their last follow-up whichever occurs first. For patients who experienced the event during the study period, their event times are known. Whereas the event times for patients who did not have the event by the end of the study period or by their last follow-up, their event times are not known. It is only known that they occur after the patient's last follow-up time or after the end of the study. The event times for these patients are censored at the end of study or at the time of their last follow-up. Analysis of time-to-event data, also known as survival analysis, must take into account the complexities caused by censored observations.

To evaluate the effect of two treatments, say treatment A and treatment B, on a time-to-event outcome, say death, it is necessary to compare the survival experience of patients receiving these two treatments. Two common but distinct approaches to compare survival outcomes are: i) comparing the entire survival curves and ii) comparing the survival probabilities at a fixed point in time. These approaches test distinct hypotheses concerning survival experience.

The first approach compares the survival experience at all time points by comparing the mortality rates (hazard rates) between treatment A and treatment B. The null hypothesis in this approach is that the risk of mortality after treatment A is the same as the risk of mortality after treatment B at all time points. In univariate analysis, this comparison is performed using the log-rank test. When it is necessary to adjust for other covariates, the hazard rates between treatment A and treatment B can be compared in multivariate analysis using the Cox regression models. Note that the Cox models assume that the hazard rates are proportional at all time points. In other words, it assumes that the hazard ratio between the treatments is constant over time. It is important to ensure this assumption is met when fitting the Cox models.

The second approach compares the survival probabilities at a fixed time point t_0, for example at five years after treatment. The null hypothesis in this approach is that the survival probability at time t_0 after treatment A is the same as the survival probability at time t_0 after treatment B. In univariate analysis, the survival probabilities at time t_0 are estimated separately for the treatment groups using the Kaplan-Meier estimator. The estimated survival probability at time t_0 for the two groups are then compared. In multivariate analysis, the survival probabilities at time t_0 can be compared using the estimated survival probabilities after adjusting for other covariates using the method of Zhang et al.[1] or using the pseudo-value approach of Klein et al.[2] Since the survival probabilities at only one time point are considered, the proportional hazards assumption is not relevant when comparing survival probabilities at a fixed point in time.

(Continued on page 2)

Volume 18, Number 5
November/December 2012

In this issue:

Figure 1: Survival Curves for Example 1

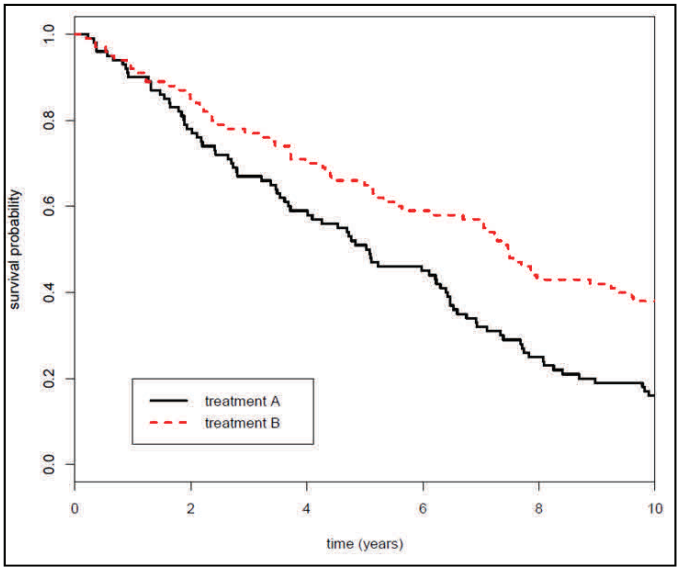
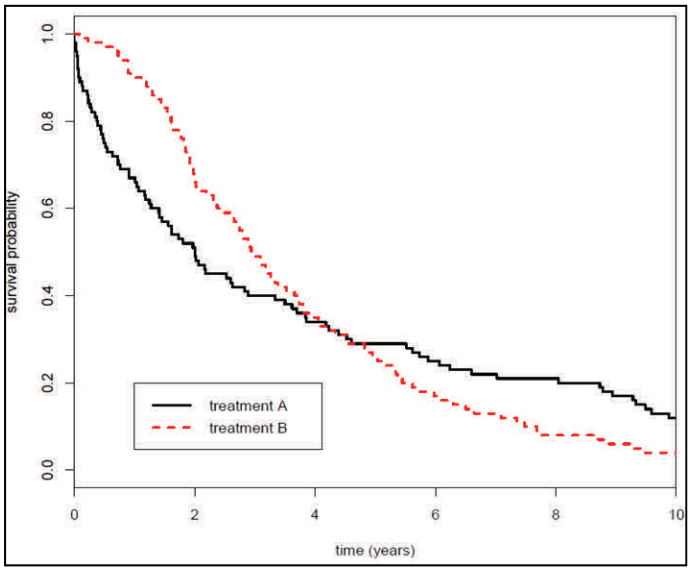


Figure 2: Survival Curves for Example 2



The decision whether to compare the entire survival curves or to compare the survival probabilities at a specific time point depends the objective of the study as well as the anticipated effect of the treatments on the outcome. This decision should be made during the design stage and should be stated clearly in the study hypothesis. If the study objective is to show that treatment B results in superior survival compared to treatment A at all time points as in example 1 (Figure 1, Table 1) then comparing the entire curves is an appropriate hypothesis. In contrast, if the study objective is to show that treatment B results in superior survival at a specific time point, for example in the early period as in example 2 (Figure 2, Table 1), then comparing the survival probabilities at an earlier time point would be more appropriate. If the objective is to show that treatment B provides a long-term survival benefit although mortality might be higher in the early period, then comparing survival probabilities at a later time period would be an appropriate hypothesis. When the hazards rates are proportional, comparing the entire survival curves uses more information, hence is more powerful than comparing the survival probabilities at a fixed time point. On the other hand, if the hazard rates are not proportional, the power of the log-rank test is low. For illustration, let look at the following examples.

In example 1 (Figure 1), survival rate is higher after treatment B compared to treatment A at all time points and the hazards rates for these treatments are proportional. Table 1 shows the Kaplan-Meier estimates and the 95% confidence intervals for the survival probabilities at six months, one year, and five years. The log-rank test comparing the entire survival curves provides strong indication that treatment B results in superior survival compared to treatment A with a p-value of 0.002 (Table 2). However, as seen in Figure 1 and in Table 1, the survival probabilities after both treatments are similar in the first year. The estimated survival probability at six months is 96% (95% CI of 91% - 99%) after Treatment A vs. 97% (93% - 99%) after treatment B. The difference in six month survival probabilities between the two treatments is 1% with a non-significant p-value of 0.699 (Table 2). Similarly, there is no difference in one-year survival between the treatments with a p-value of 0.620(Table 2). With a higher mortality rate, survival after treatment A declines faster than treatment B resulting in a 5-year survival probability of 51% (41% - 61%) compared to 65% (CI 55% - 74%), respectively. The difference in 5-year survival probabilities is 14% and is marginally significant with a p-value of 0.044.

In example 2 (Figure 2), treatment B improves survival in the first four years post treatment. However, after the first four years, survival after treatment B is worse compared to treatment A. The log-rank test comparing the entire survival curves indicates no net survival difference between treatment A and treatment B with a p-value of 0.481 (Table 2). However, there is a significant difference in survival probabilities in the earlier period. At six months, the estimated survival proba-

Table 1: Kaplan-Meier Estimate and 95% Confidence Interval for Survival Probabilities

Time Point	Example 1		Example 2	
	Treatment A	Treatment B	Treatment A	Treatment B
6 months	0.96 (0.91, 0.99)	0.97 (0.93, 0.99)	0.75 (0.66, 0.83)	0.97 (0.93, 0.99)
1 year	0.90 (0.84, 0.95)	0.92 (0.86, 0.96)	0.67 (0.58, 0.76)	0.90 (0.83, 0.95)
5 years	0.51 (0.41, 0.61)	0.65 (0.55, 0.74)	0.29 (0.21, 0.38)	0.26(0.18, 0.35)

(continued on page 3)

bility after treatment A is 75% (66% - 83%) compared to 97% (93% - 99%) after treatment B. Treatment B results in a 22% increase in 6-months survival with a strongly significant p-value of < 0.0001. Similarly, there is a statistically significant difference in one-year survival with a p-value of < 0.0001. However, the mortality rate after the first year increases much faster after treatment B than after treatment A leading to similar survival probabilities between treatments by four years after treatment. By five years, the estimated survival probability is 29% (21% - 38%) after treatment A and 26% (18% - 35%) after treatment B. The difference in five year survival probabilities is 3% with a non significant p-value of 0.634.

Table 2: Survival Comparison Results

Comparison	Example 1	Example 2
Entire curve	Chi-square statistic = 9.60 p-value = 0.002	Chi-square statistic = 0.50 p-value = 0.481
Point-wise at 6 months	Chi-square statistic = 0.15 p-value = 0.699	Chi-square statistic = 24.43 p-value < 0.0001
Point-wise at 1 year	Chi-square statistic = 0.25 p-value = 0.620	Chi-square statistic = 16.84 p-value < 0.0001
Point-wise at 5 years	Chi-square statistic = 4.05 p-value = 0.044	Chi-square statistic = 0.48 p-value = 0.634

These examples illustrate the distinction between comparison of the entire survival curves versus comparison of the survival probabilities at a fixed time point. In the first example, treatment B provides an overall survival experience that is superior to treatment A. However, the survival probabilities at six months and at one year after treatment B are not statistically better than treatment A. On the other hand, in example 2, there is no net difference in survival between treatment A and treatment B. However, treatment B shows a significant improvement in survival at six months and at one year compared to treatment A. Since these approaches test different hypotheses relating to survival, they can lead to different conclusions concerning the impact of treatments on survival. It is important to clearly specify the hypothesis in a study so that the appropriate method is used to correctly address the question of interest. ■

References:

1. Zhang X, Loberiza FR, Klein JP, and Zhang, MJ. A SAS macro for estimation of direct adjusted survival curves based on a stratified Cox regression model. *Computer Methods and Programs in Biomedicine* 2007; 88: 95 – 101.

2. Klein JP, Logan B, Harhoff M, and Andersen PK. Analyzing survival curves at a fixed point in time. *Statistics in Medicine* 2007; 26: 4505 – 4519.

To subscribe to *Datum* visit:

http://www.mcw.edu/biostatistics/CTSIKeyFunction/Newsletter/SubscriptionForm.htm

Datum

Editors:
Jennifer Le-Rademacher, PhD
Ruta Brazauskas, PhD
Brent Logan, PhD
Alexis Visotcky, MS
Haley Montsma, BBA

Datum is published by the Division of Biostatistics. It is available online at: www.mcw.edu/biostatistics/datum.htm.
We welcome your newsletter questions, comments, and suggestions. Please contact us with address updates or to be added/removed from our mailing list.

Please direct all newsletter correspondence to:
Datum Newsletter ● Institute for Health and Society ● Medical College of Wisconsin
8701 Watertown Plank Road ● Milwaukee, WI 53226—0509
Phone: 414-955-7439 ● E-mail: hmontsma@mcw.edu

Interpreting Tables

Ruta Brazauskas, PhD, Assistant Professor, Division of Biostatistics

Tables are commonly used to present data in biomedical literature. They are frequently employed to describe the study cohort in the beginning of a manuscript and summarize study findings at the end. In this article we will focus on the simplest tables presenting information in terms of counts and percentages which are encountered in many scientific papers. In addition, we will also mention some aspects which are important to keep in mind while working on the manuscript. As an example, consider a recent publication titled "Relationship between documentation status and survival for medically underserved Hispanic breast cancer patients" (Castro-Echeverry et al., 2012). This study examines survival in socioeconomically disadvantaged Hispanic documented and undocumented breast cancer patients. Table 1 presenting demographic characteristics of the study cohort is adapted from the aforementioned publication.

Table 1: Demographic characteristics of documented and undocumented patients.

Variable	Documented (N=507) (%)	Undocumented (N=244) (%)	P-value
Stage at diagnosis			<0.001
Localized	355 (70)	134 (55)	
Regional	102 (20)	62 (25)	
Metastatic	50 (10)	48 (20)	

What information is conveyed in this table? The portion of the table selected for this article focuses on describing the stage of the breast cancer at the time of diagnosis. The headings of the columns inform us that two groups of patients are being considered: documented and undocumented immigrants. The sample consists of 751 patients with 507 (68%) of them being documented immigrants while 244 (32%) of them fall into undocumented immigrant category. The table presents counts along with column percentages (calculated from the total in each of the two groups) for the six possible categories defined by documentation status and disease stage. For example, 355 out of 507 documented immigrants were diagnosed with localized breast cancer. Note that this constitutes 70% of the documented immigrants in the study. On the other hand, 134 out of 244 patients who are undocumented immigrants had localized cancer at the time of their diagnosis and that represents only 55% of patients in the latter group. Given the information in the table, we can investigate other features of the study group. For example, we calculate that about 6% (48 out of 751) of all patients are undocumented immigrants with metastatic cancer. It is easy to see that 13% of patients (98 out of 751) present with metastatic breast cancer. However, knowing that a patient is an undocumented immigrant increases her chances of having a metastatic cancer to 20% (48 our of 244).

Note that the authors chose to present column-wise percentages, i.e. those calculated using the total in each of the two groups defining the columns. Alternatively, the counts could be accompanied by row-wise percentages. In this case, the interpretation of the percentages would be slightly different. Among patients presenting with localized breast cancer, 73% were documented and 27% were undocumented immigrants while among those with metastatic cancer the percentages change to 51% and 49%, respectively. While the counts in the table will not change, the choice between which percentages to present, column or row wise, depends on the comparison (analysis) of interest.

A look at the scientific literature reveals that some tables contain counts and percentages, other present only percentages and a few of them contain only counts. Presenting both –counts along with percentages – is the most common way of arranging a table. However, tables with multiple columns which become lengthy and cumbersome to present and read can be reduced by displaying only percentages. This is also common practice in displaying results from surveys where weighting is needed to adjust for complex sampling procedures. While displaying only counts only is possible, it should be avoided as it makes reading a table and drawing conclusions difficult. The counts can vary greatly due to unequal number of individuals in the groups considered and thus a reader should always be mindful of the total number of subjects in each group. Presenting percentages allows the audience to comprehend the information without having to convert counts into proportions.

The last number which is often presented along with the counts is the p-value. Behind every p-value there is a hypothesis being tested. In the study of documentation status of Hispanic breast cancer patients summarized in Table 1, we observed that a higher proportion of documented immigrants are diagnosed with less advanced disease as compared to the undocumented immigrants. The opposite is true for more advanced stages of the disease. Is this difference statistically significant? Statistically, the hypothesis is that the proportions of patients with localized, regional, and metastatic cancer are the same among documented and undocumented breast cancer patients. This question is answered by using a chi-squared test which compares the number of patients in each of the six categories to what would be expected if indeed there was no difference in the distribu-

(continued on page 5)

tion of disease stage in the two groups. The p-value <0.001 suggests that there is sufficient evidence to conclude that there is an association between disease stage and documentation status. Note that while interpreting the p-value a reader always sets a certain threshold to judge the significance of the findings. Usually it is a significance level alpha of 0.05 or 0.01. If the p-value is smaller that the assumed level of significance, we conclude that there is a significant difference in proportions of patients in various disease stages between the two groups. In this example, given the small p-value the same conclusion would be reached regardless if alpha of 0.05 or 0.01 is being used. ■

References:

Castro-Echeverry E, Kao LS, Robinson EK, Silberfein EJ, Ko TC, and Wray CJ. Relationship between documentation status and survival for medically underserved Hispanic breast cancer patients. *Journal of Surgical Research*, 2012 May 16. [Epub ahead of print]

Biostatistics Drop-In Consulting Schedule

1. Medical College of Wisconsin
Tuesdays & Thursdays
1:00-3:00 PM
Building: Health Research Center
Room: H2400

2. Froedtert Hospital
Mondays & Wednesdays
1:00-3:00 PM
Building: Froedtert Pavilion
Room: #L772A- TRU Offices

3. MCW Cancer Center
Wednesdays 10:00 AM -12:00 PM
Fridays 1:00-3:00 PM
Building: MCW Clinical Cancer Center
Room: Clinical Trials Support Room CLCC:3236
(Enter through C3233)

4. Zablocki VA Medical Center
1st & 3rd Monday of the month
9:00-11:00 AM
Building: Building 111, 5th Floor B-wing
Room: 5423

5. Marquette University
Every Tuesday
8:30-10:30 AM
Building: School of Nursing-Clark Hall
Room: Office of Research &
Scholarship: 112D
*Please note: Priority given to MU
Nursing and Dental School

Questions:
Haley Montsma
414.955.7439
hmontsma@mcw.edu

Bridging the sciences BIOSTATISTICS MEDICAL COLLEGE OF WISCONSIN CANCER CENTER CTSI

Biostatistics

Lecture Series 2012• 2013

The faculty of the MCW Division of Biostatistics present a yearly statistical lecture series targeted at basic and clinical investigators. The series provides an introduction to statistical techniques and concepts used in medical research; a wide range of topics are included.

For more information about the Biostatistics Lecture Series please visit our website:
www.mcw.edu/biostatistics/LectureSeries.htm

Schedule:

Basic Concepts of Bayesian Statistics
Prakash Laud, PhD
Friday, February 15, 2013
12:00-1:00 pm
Medical Education Building– M2050

Propensity Scores
Mei-Jie Zhang, PhD
Friday, March 29, 2013
12:00-1:00 pm
Clinical Cancer Center– Room K

Common Errors in Survival Analysis
John Klein, PhD
Friday, April 19, 2013
12:00-1:00 pm
Clinical Cancer Center– Room K

Bridging the sciences BIOSTATISTICS

All previous biostatistical lectures are available for viewing on our BiostatisticsMCW YouTube page.
Check it out today! www.youtube.com/user/BiostatisticsMCW

Medical College of Wisconsin
Division of Biostatistics
Suite H2100/H2400
Phone: 414.955.8280

Sponsored by: