Confi ence Regions For The Equality Of Two Survival Curves

John P. Klein, Ph.D.
Mei-Jie Zhang, Ph.D.

Technical Report 13

July 1 6

Division of Biostatistics
Medical College of Wisconsin
8701 Watertown Plank Road
Milwaukee WI 53226
Phone: (414)456-8280

Confidence Regions For The Equ lity Of Two Surviv l Curves

John P. Klein and Mei-Jie Zhang
Division of Biostatistics
Medical College of Wisconsin
8701 Watertown Plank Road
Milwaukee, WI 53226

Abstract

Often when comparing the survival rates of individuals given either of two treatments the anal sis stops with a test of the h pothesis of no treatment difference and perhaps a plot of the two survival functions. The h pothesis test is usuall a comparison of the two survival curves over the entire observational period. n alternative approach to this problem is to provide an investigator with a confidence region for the set of times at which the survival rates of the two treatments are the same. We discuss how such confidence regions can be constructed in three situations. First, we construct confidence regions when there are no additional factors that need be adjusted for. Second, based on a proportional hazards model, we show how to construct the desired confidence regions adjusted for explanator covariates that are not confounded with the two treatments. Lastl , we extended these results to allow for explanator covariates that are confounded with treatment. These approaches are illustrated on retrospective data gathered to compare the survival rates of allogeneic and autologous bone marrow transplants for acute leukemia.

1 Intro uction

common problem arising in biomedical applications is the comparison of the survival functions or hazard rates of two treatments. Most standard statistical tests are based on comparing the survival curves or equivalently the hazard functions over a given time period. The time period considered is typically the period from initiation of the treatment to some point in time where observation of the patients ceases. This comparison may be made by the log ran test (cf. ndersen et al. 1993), for e ample, when there are no other covariates that may influence survival. When there are other covariates that may affect outcome in addition to the treatments under consideration, testing of treatment effects is carried out by some type of regression technique. Tests may be based on any number of parametric or

semi-parametric models,

off between early high mortality with allogeneic transplants and lower reoccurrence rates. To help in the decision between these two competing treatment modalities a confidence region for the times at which the survival probabilities of the two treatments are the same is of interest. lso, since autologous transplants are easier to perform as no donor is needed, a confidence region for those times where the survival probability for a autologous transplant patient is not smaller than the corresponding survival probability for an allogeneic transplant patient is also of interest.

2 Confi ence Regions When There Are No Other Explanatory Covariates

$S_1(t) = S_2(t)$ as $\{t_0 : -z_{\alpha/2} \leq \hat{\Delta}(t_0)/\sqrt{Var[\hat{\Delta}(t_0)]} \leq z_{\alpha/2}\}$. Note that the confidence region can also be written as $\{t_0 : \hat{\Delta}(t_0) - z_{\alpha/2}\sqrt{Var[\hat{\Delta}(t_0)]} \leq 0 \leq \hat{\Delta}(t_0) + z_{\alpha/2}\sqrt{Var[\hat{\Delta}(t_0)]}\}$, so that the desired region is the set of all those time points for which a $(1-\alpha)\times 100\%$o00(the)-2714.9995TJfl

3 A justment For Covariates Not Confoun e With Outcome

In many e periments there are other ris factors that need to be adjusted for prior to ma ing the main comparison between the two treatments. Let $\boldsymbol{Z} = (Z_1, \cdots, Z_p)$ be a vector of fi ed time covariates that influence survival. In this section we assume that there is no significant interaction between the comparison of interest and any of these covariates.

The confidence region, adjusted for these other covariates, is based on the proportional hazards model (Co (1972)). Here we fit a proportional hazards model for the e planatory covariates stratifying on the treatment of interest. That is we fit the model

$$\lambda(t|\boldsymbol{Z}, \text{Treatment}) = \begin{cases} \lambda_{10}(t)\, \text{e p}\{\beta^T \boldsymbol{Z}\}, & \text{for treatment 1,} \\ \lambda_{20}(t)\, \text{e p}\{\beta^T \boldsymbol{Z}\}, & \text{for treatment 2.} \end{cases} \tag{3.1}$$

Let $\hat{\beta}$ and $I(\hat{\beta})$ be the partial ma imum li elihood estimator and the observed information for this model. n estimator of the baseline cumulative hazard rate for treatment j, $j = 1, 2$ is given by Breslow's (1975) estimator

$$\hat{\Lambda}_{j0}(t) = \int_0^t \frac{dN_j(\ )}{S_j^{(0)}(\hat{\beta},\ )}, \quad \text{where} \tag{3.2}$$

$$S_j^{(0)}(\hat{\beta},\ ) = \sum_{i=1}^{n} Y_{ij}(\ )\, \text{e p}\{\beta^T \boldsymbol{Z}_i\} \tag{3.3}$$

with $Y_{ij}(\ )$ the indicator of whether the ith individual is at ris at time and is in the jth group.

For an individual with a covariate vector $\boldsymbol{Z_0}$, the two treatments will have the same survival rate at time t_0 if $\Lambda(t|\boldsymbol{Z_0}, \text{Treatment 1}) = \Lambda(t|\boldsymbol{Z_0}, \text{Treatment 2})$, which from (3.1) is equivalent to having $\Lambda_{10}(t_{0}$g

Since at t_0 an α level test of the equality of the two survival functions is accepted when $\hat{\Delta}(t_0)/[Var(\hat{\Delta}(t_0))]^{1/2}$

To construct the confidence region where the survival rates are the same for the two treatments a stratified proportional hazards model is used. We fit the model

$$\lambda(t|\boldsymbol{Z}, \text{Treatment}) = \begin{cases} \lambda_{10}(t)\,\mathrm{e}^{\ \mathrm{p}\{\gamma_1^T \boldsymbol{Z}_1 + \theta^T \boldsymbol{Z}_2\}}, & \text{for treatment 1,} \\ \lambda_{20}(t)\,\mathrm{e}^{\ \mathrm{p}\{\gamma_2^T \boldsymbol{Z}_1 + \theta^T \boldsymbol{Z}_2\}}, & \text{for treatment 2.} \end{cases} \tag{4.1}$$

Estimates for $\beta = (\theta_1^T, \gamma_1^T, \gamma_2^T)$ are found by fitting a Co model, stratified on treatment group to the data with an augmented covariate vector $\boldsymbol{Z}^T = (\boldsymbol{Z}_2^T, \boldsymbol{Z}_1^T I[\text{Treatment} = 1], \boldsymbol{Z}_1^T I[\text{Treatment} = 2])$. For a given set of confounding factors, $\boldsymbol{Z}_{10}$, the two treatments will have the same survival rate at time t_0 if

$$\Delta(t_0|\boldsymbol{Z}_{10}) = \Lambda_{20}(t_0)\,\mathrm{e}^{\ \mathrm{p}\{\gamma_2^T \boldsymbol{Z}_{10}\}} - \Lambda_{10}(t_0)\,\mathrm{e}^{\ \mathrm{p}\{\gamma_1^T \boldsymbol{Z}_{10}\}} \tag{4.2}$$

is equal to zero. The estimator of $\Delta(t_0|\boldsymbol{Z}_{10})$ given by

$$\hat{\Delta}(t_0|\boldsymbol{Z}_{10}) = \hat{\Lambda}_{20}(t_0)\,\mathrm{e}^{\ \mathrm{p}\{\hat{\gamma}_2^T \boldsymbol{Z}_{10}\}} - \hat{\Lambda}_{10}(t_0)\,\mathrm{e}^{\ \mathrm{p}\{\hat{\gamma}_1^T \boldsymbol{Z}_{10}\}}$$

follows from the fitted Co model with $\Lambda_{j0}()$ estimated using Breslow's estimator (3.2).

n estimator of the asymptotic variance of $\hat{\Delta}(t_0|\boldsymbol{Z}_{10})$ can be shown to be

$Var(\hat{\Delta}(t_0|\boldsymbol{Z}_{10}$

(dichotomized as ≤ 1 yr or > 1 yr) and two interaction covariates. The interaction covariates are $Z_{11} = 1$ if age > 30 and allo transplant and $Z_{12} = 1$ if age > 30 and auto transplant. Figure 4a and 4b show the standardized difference between the cumulative hazard rates for patients under 30 and o

difference of the two Kaplan-Meier estimators of the survival function. We find

Survival Curve Based On Transformations. *Scandinavian ournal Of Statistics*, **17**, 35-41.

Breslow, N. E. (1975). nalysis Of Survival Data Under The Proportional Hazards Model. *Int rnational Statistics R vi w*, **43**, 45–58.

Co , D. R. (1972). Regression Models nd Life-Tables (With Discussion). *ournal Of Th Royal Statistical Soci ty B*, **34**, 187–220.

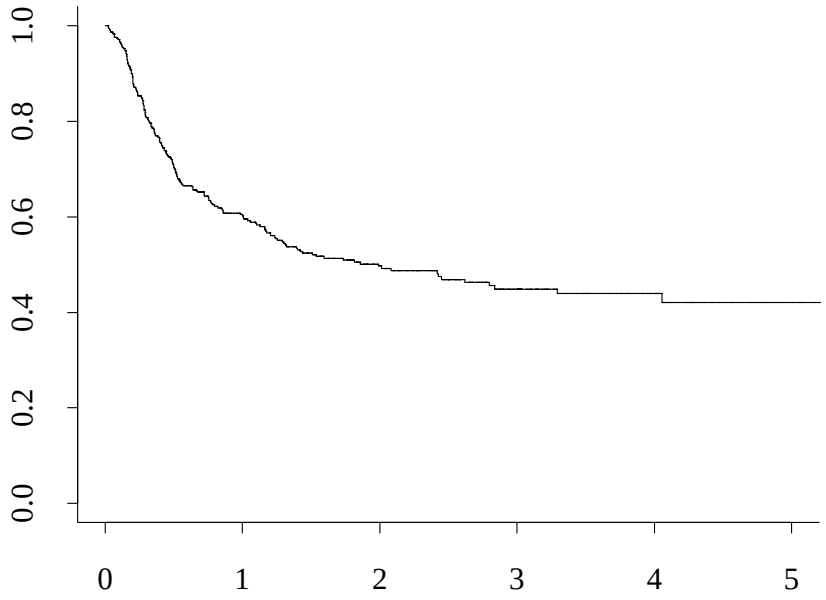
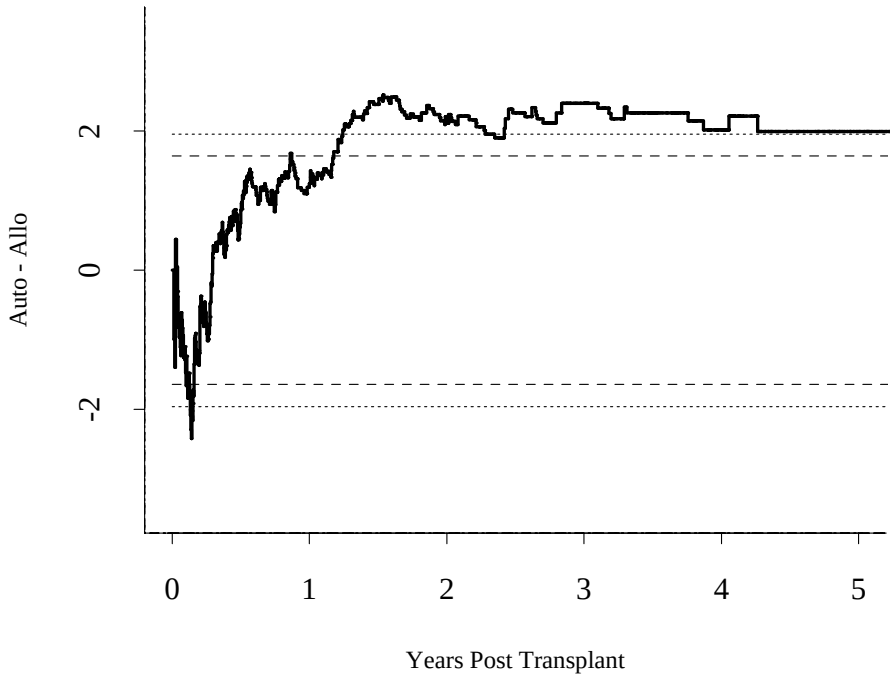
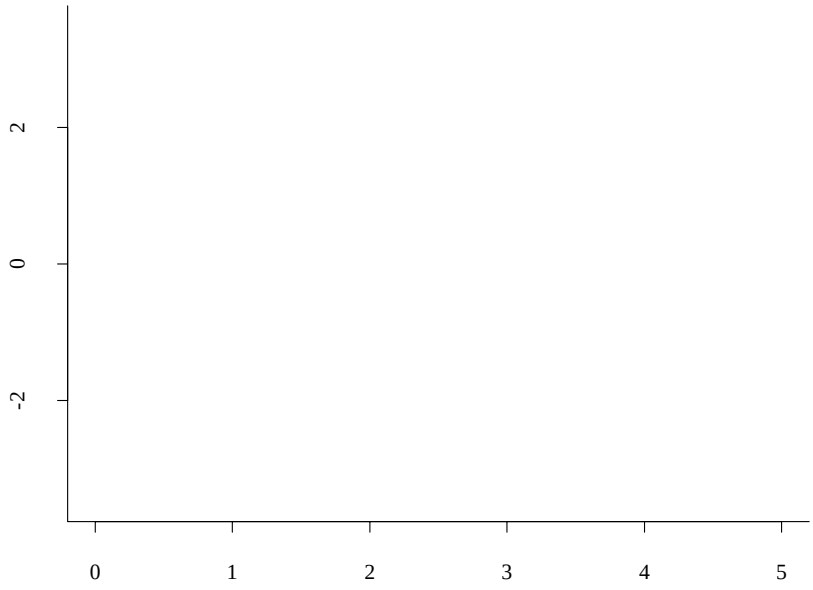

0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5

Figure 3. Standardized Difference In The Cumulative Hazards
(No Confounding)




2
0
-2
0
1
2
3
4
5