SISCER Survival Analysis

Problem Set 2

1. Straightforward
2. Cox model versus AFT model
   (a) Write down the AFT model in hazard functions, then comparing two model would lead to the answer.
   (b) Weibull distributions
3. See Fleming & Harrington (1991, pp. 153-159)

In the early to mid 1980s, the rate of liver transplantation in patients with advanced stage PBC increased substantially, largely due to the improvement in transplantation results through the use of immunosuppressive agents such as cyclosporine and OKT3.

We first show that DPCA has a negligible effect on prognosis, then use the data from the 312 randomized cases to develop a natural history model. Such a model will be useful not only in counseling patients and in understanding the course of PBC in untreated patients, but also in providing historical control information to evaluate the efficacy of new therapeutic interventions such as liver transplantation. These evaluations of liver transplantation will be important since randomized trials comparing transplantation with non-surgical management will not be performed and since PBC is one of the most common indicators for liver transplantation in adults. In this Chapter, we present the analyses to evaluate DPCA, to develop a natural history model, and to illustrate the model's use in evaluating liver transplantation. The data set of 112 nonrandomized patients is used in model validation of the natural history model and to illustrate its use in survival prediction. Appendix D contains the data for 106 of the 112 patients, since six were lost to followup soon after their initial visit to the Mayo Clinic. Of the 106, 36 had died by July 1986, with six others having undergone transplantation.

In the database of 418 patients, the 25 transplanted patients were considered censored at the date of transplantation. This induces a small bias in a natural history model. Transplantation occurred after a median followup of 66 months for the 19 transplanted clinical trial patients and 50 months for the 6 transplanted non-trial patients.

Effect of DPCA on Patient Survival

Figure 4.4.1 presents the Kaplan-Meier estimates of survival of PBC patients following randomization to either DPCA or placebo. The curves show little separation. The median survival time of the pooled group is just under 10 years.

Under the proportional hazards assumption, the Cox regression model can be used to measure treatment effect. If treatment is coded by $Z = 0$: DPCA, $Z = 1$: Placebo, then in the model

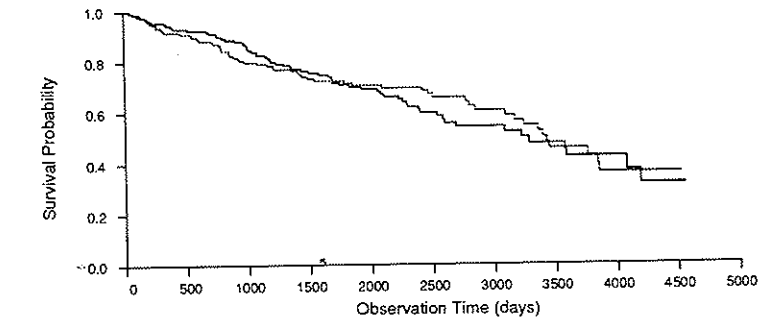
$$\lambda(t|Z) = \lambda_0(t)\exp(\beta Z), \tag{4.1}$$

$\lambda_0(\cdot)$ represents the hazard function for death while being treated with DPCA, and β is the log of the hazard ratio; i.e., if $\lambda_1(t) \equiv \lambda(t|Z = 1)$ then, for all t,

$$\lambda_1(t)/\lambda_0(t) = e^{\beta}.$$

If $L(\beta)$ is the Cox partial likelihood for β based on the censored survival data and $\mathcal{L} \equiv \ln L$, then the score statistic is given by

$$U(\beta) \equiv \frac{d}{d\beta}\mathcal{L}(\beta),$$



	Time Interval				
Group	0-1000	1000-2000	2000-3000	3000-4000	4000-5000
DPCA	23/158	22/128	13/74	5/31	2/10
Placebo	31/154	12/120	7/70	10/32	0/11

(# events/# at risk)

Figure 4.4.1 Estimated survival curves in DPCA and placebo groups, PBC data.

and Fisher's observed information is

$$\mathcal{I}(\beta) = -\frac{d^2}{d\beta^2}\mathcal{L}(\beta).$$

For these data $U(0) = -1.781115$, $\mathcal{I}(0) = 31.19845$, and $\mathcal{L}(0) = -639.9799$. Hence the standardized score statistic or Rao test statistic is

$$\{U(0)\}^2/\mathcal{I}(0) = 0.10168.$$

Later it will be established that this statistic is distributed asymptotically as a chi-square with one degree of freedom when $H : \beta = 0$. Since there are no nuisance covariates in this model, this score statistic is identical to the logrank statistic for no treatment effect.

The maximum partial likelihood estimate for β is $\hat{\beta} = -0.0571242$ and $\mathcal{L}(-0.0571242) = -639.9290$. Hence, the likelihood ratio statistic for the hypothesis $H : \beta = 0$ is

$$-2\{\mathcal{L}(0) - \mathcal{L}(\hat{\beta})\} = 0.10193.$$

Since the standard error for $\hat{\beta}$ is estimated by $\{\mathcal{I}(\hat{\beta})\}^{-1/2}$, where $\mathcal{I}(\hat{\beta}) = 31.1525$, the Wald statistic for $H : \beta = 0$ is

$$\hat{\beta}^2\mathcal{I}(\hat{\beta}) = 0.10166.$$

As expected, the Rao, Wald, and likelihood ratio statistics yield nearly identical results.

Under the proportional hazards assumption, the hazard ratio

$$r \equiv \lambda_1(t)/\lambda_0(t) = e^{\beta}$$

is independent of t. In large samples,

$$\hat{\beta} \sim N(\beta, \{\mathcal{I}(\hat{\beta})\}^{-1});$$

thus $\hat{r} = e^{\hat{\beta}} = 0.94448$, and a 95% confidence interval for r is

$$\exp\{\hat{\beta} \pm 1.96\{\mathcal{I}(\hat{\beta})\}^{-1/2}\} = (0.66479, 1.34184).$$

We estimate the failure rate on placebo to be 94.4% that on DPCA, and there is evidence against it being more than 134% that on DPCA. If DPCA must improve patient survival by more than a factor of 1/2 to offset the drug's expense, toxicity and inconvenience of administration, then this trial supports not using the drug in this disease.

An analysis of subsets defined by clinical, biochemical and histological variables failed to yield evidence of important survival differences between the drug and the placebo in patient subgroups.

Natural History Model for PBC

The data in Appendix D on the 312 PBC randomized patients can be used to build a statistical model for the influence of covariates on disease outcome. Table 4.4.1 provides the distributions of 14 clinical, biochemical and histological variables. With the exception of 4 missing platelet counts and two missing urine copper values, the data are complete.

For the remainder of this section, we use the model

$$\lambda(t|\mathbf{Z}) = \lambda_0(t)\exp(\beta'\mathbf{Z}), \tag{4.2}$$

where $\mathbf{Z}' \equiv (Z_1, Z_2, \ldots, Z_K)$ is a vector of K predictors and $\beta' \equiv (\beta_1, \beta_2, \ldots, \beta_K)$ are the regression coefficients. Each predictor Z_i could be defined in a variety of ways, such as using the variables in Table 4.4.1, transformations or crossproducts of these variables, etc. In model (4.2), each individual patient is given a risk score $R \equiv \beta_1 Z_1 + \ldots + \beta_K Z_K$. Let $S(t|\mathbf{Z})$ denote the probability that patient with risk factors given by $\mathbf{Z}$ (and with risk score R) is still alive t years after time 0, and let $S_0(t)$ denote the survival function for individuals having risk score $R = 0$. Then

$$\begin{aligned} S(t|\mathbf{Z}) &= \{S_0(t)\}^{\exp(R)} \\ &= \{e^{-\Lambda_0(t)}\}^{\exp(R)}, \end{aligned}$$

where time $t = 0$ usually denotes the time the measurements in the covariate vector $\mathbf{Z}$ are obtained. In the PBC data set in Appendix D, time $t = 0$ is the date of treatment randomization. One can estimate $S(t|\mathbf{Z})$ by

$$\hat{S}(t|\mathbf{Z}) = \{e^{-\hat{\Lambda}_0(t)}\}^{\exp(\hat{R})}, \tag{4.3}$$

where the maximum partial likelihood estimate vector $\hat{\beta}$ is used to obtain $\hat{R} = \hat{\beta}_1 Z_1 + \cdots + \hat{\beta}_K Z_K$, and where $\hat{\Lambda}_0(\cdot)$ is the Breslow estimator in Eq. (3.29). In the proportional hazards model in Eq. (4.2), each coefficient β_k has the simple interpretation that every unit increase in the kth covariate, Z_k, changes the hazard function by the multiplicative factor $\exp(\beta_k)$.

Initially model (4.2) was fit to the data with $\mathbf{Z}' = (Z_1, \ldots, Z_{14})$ chosen to be the 14 variables in Table 4.4.1. If $\mathbf{U}(\beta)$ and $\mathcal{I}(\beta)$ denote the score vector and Fisher's observed information matrix, respectively, the collection of univariate Rao or logrank statistics, $\{[\{\mathbf{U}(0)\}_k]^2/\{\mathcal{I}(0)\}_{kk} : k = 1, \ldots, 14\}$, are listed in the right-hand column of Table 4.4.1. The term $\{\mathcal{I}(0)\}_{jk}$ is the estimated covariance of $U_j(0)$ and $U_k(0)$, and since $\mathcal{I}(0) \doteq \int_0^\infty \sum_{i=1}^n \mathbf{V}(0,t)dN_i(t)$ for $\mathbf{V}$ defined by (3.23), $\{\mathcal{I}(0)\}_{jk}$ is the sum (over death times) of the covariances of Z_j and Z_k among those at risk at each death time. Thus inspection of

$$\{c_{jk} \equiv \{\mathcal{I}(0)\}_{jk}/[\{(\mathcal{I}(0)\}_{jj}\{\mathcal{I}(0)\}_{kk}]^{1/2} : j \neq k\}$$

Table 4.4.1 Prognostic Factors: Summary of Univariate Statistics (312 Patients in the PBC Clinical Trial of DPCA)

Demographic	min	1st Q	med	3rd Q	max	Missing	Rao χ^2(1 d.f.)
Age (years)	26.3	42.1	49.8	56.7	78.4	0	20.86
Sex	male:	36	female:	276		0	4.27
Clinical		**Absent**		**Present**		**Missing**	**Rao χ^2(1 d.f.)**
Ascites		288		24		0	104.02
Hepatomegaly		152		160		0	40.18
Spiders		222		90		0	30.31
Edema[1]	0: 263		1/2: 29	1: 20		0	97.89
Biochemical	**min**	**1st Q**	**med**	**3rd Q**	**max**	**Missing**	**Rao χ^2(1 d.f.)**
Bilirubin	0.3	0.8	1.35	3.45	28.0	0	190.62
Albumin	1.96	3.31	3.55	3.80	4.64	0	70.83
Urine Copper	4	41	73	123	588	2	84.35
Pro Time	9.0	10.0	10.6	11.1	17.1	0	51.76
Platelet Count	62	200	257	323	563	4	12.15
Alkaline Phos	289	867	1259	1985	13862	0	2.58
SGOT	26	81	115	152	457	0	29.59
Histologic	**1**	**2**	**3**	**4**		**Missing**	**Rao χ^2(1 d.f.)**
Stage	16	67	120	109		0	46.49

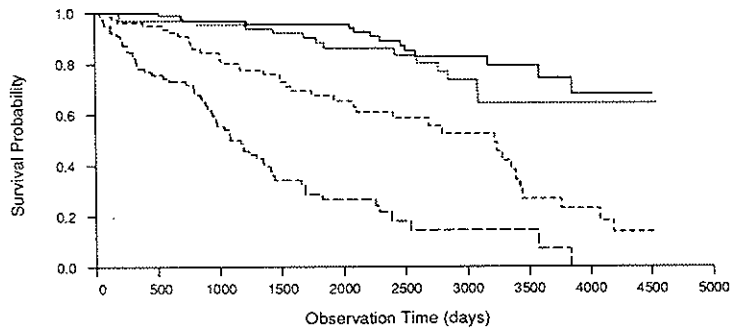
[1]Edema 0: No edema and no diuretic therapy for edema
$\frac{1}{2}$: Edema but no diuretics, or edema resolved by diuretics
1: Edema despite diuretic therapy

Rao statistics computed after six missing values were replaced by median values (i.e., 4 missing Platelet Counts, 257; 2 missing Urine Copper, 73)

is a method for finding co-linearities among the K components of $\mathbf{Z}$. The largest values of c_{jk} are 0.47 between hepatomegaly and stage, 0.37 between bilirubin and SGOT, and 0.37 between bilirubin and urine copper. Bilirubin is the strongest univariate predictor of survival. One would expect and can verify that the predictive strength of the variables SGOT and urine copper are reduced in models which adjust for bilirubin. In building a parsimonious natural history model based on easily accessible variables, there is hope that readily available measurements on hepatomegaly and bilirubin will contain much of the predictive information from the invasive variable histologic stage (which requires a liver biopsy), and in the frequently unmeasured variables, urine copper and SGOT.

The score statistics in Table 4.4.1 show that nearly all 14 variables are highly significantly associated with patient survival. The Kaplan-Meier plot in Figure 4.4.2 indicates that bilirubin levels distinguish patients with good and poor prognosis.

Parsimonious but accurate models based on inexpensive, non-invasive and readily available measurements are useful in clinical science, and so the variables stage, urine copper, and SGOT were eliminated temporarily from the variable selection process. The untransformed versions of the remaining 11 variables were inserted into Eq. (4.2), and a step-down procedure was employed to eliminate variables, using the Wald statistic as a criterion for deletion of the least predictive variable. Table 4.4.2 displays the first step of the procedure, which led to the elimination of



	Time Interval				
Group	0-1000	1000-2000	2000-3000	3000-4000	4000-5000
0 < bili < 0.8	3/90	1/86	7/64	3/28	0/10
0.8 < bili < 1.3	3/66	5/62	4/38	2/18	0/6
1.3 < bili < 3.4	14/78	10/59	5/30	8/15	2/5
3.4 < bili	34/78	18/41	4/12	2/2	0/0

Bilirubin measured in mg/dl (# events/# at risk)

Figure 4.4.2 Estimated survival curves for four groups determined by serum bilirubin levels, PBC data.

the variable alkaline phosphatase, and the sixth step, at which each of the remaining variables has a Wald statistic exceeding 6.0. The likelihood ratio test for the five eliminated variables has the value

$$-2(-554.237 + 550.603) = 7.268,$$

and has an approximate chi-square distribution with 5 degrees of freedom. There is little evidence to retain the variables alkaline phosphatase, ascites, platelet count, sex, or presence of spiders.

In the model in Table 4.4.2(b), all variables were entered untransformed. In such a model, an increase in the value of the ith covariate, Z_i, from x to $(x + d)$ will lead to a multiplicative increase in the hazard by a factor $\exp(d\beta_i)$, independent of the value x. However, the clinical literature suggests that changes from x to $(x + d)$ in the values of variables such as bilirubin should have a greater impact on prognosis when x is small. To evaluate the need for transformations of the four continuous variables in the six variable model in Table 4.4.2(b), the variables log(age), log(albumin), log(protime), and log(bilirubin) were added. The resulting

Table 4.4.2 Results of variable selection procedure in 312 randomized cases with PBC.

(a) First Step, log likelihood −550.603

	Coef.	Std. Err.	Z stat.
Age	2.819 e-2	9.538 e-3	2.96
Albumin	−9.713 e-1	2.681 e-1	−3.62
Alk. Phos	1.445 e-5	3.544 e-5	0.41
Ascites	2.813 e-1	3.093 e-1	0.91
Bilirubin	1.057 e-1	1.667 e-2	6.34
Edema	6.915 e-1	3.226 e-1	2.14
Hepatomegaly	4.853 e-1	2.913 e-1	2.21
Platelets	−6.063 e-4	1.025 e-3	−0.59
Prothrombin Time	2.428 e-1	8.420 e-2	2.88
Sex	−4.769 e-1	2.643 e-1	−1.80
Spiders	2.889 e-1	2.093 e-1	1.38

(b) Last Step, log likelihood −554.237

	Coef.	Std. Err.	Z stat.
Age	0.0338	0.00925	3.65
Albumin	−1.0752	0.24103	−4.46
Bilirubin	0.1070	0.01528	7.00
Edema	0.8072	0.30775	2.62
Hepatomegaly	0.5903	0.21179	2.79
Prothrombin Time	0.2603	0.07786	3.34

Results computed after the four patients with missing values for platelets were assigned the median count, 257, from Table 4.4.1.