Covariance Analysis of Heart Transplant Survival Data

JOHN CROWLEY and MARIE HU*

This paper presents a number of analyses to assess the effects of various covariates on the survival of patients in the Stanford Heart Transplantation Program. The data have been updated from previously published versions and include some additional covariates, such as measures of tissue typing. The methods used allow for simultaneous investigation of several covariates and provide estimates of the relative risk of transplantation as well as significance tests.

KEY WORDS: Heart transplant; Survival data; Covariance analysis.

1. INTRODUCTION

The Stanford Heart Transplantation Program, described in detail in Clark et al. [4], has received close statistical scrutiny in Turnbull, Brown, and Hu [16], Mantel and Byar [13], and Brown, Hollander, and Korwar [3]. The goal in [16] and [13] was to assess the effect on survival of transplantation, treating the patient population as homogeneous, while in [3] the influence of a number of covariates was investigated through pairwise correlations. The purpose of the present report is to explore, via the techniques in Cox [6] and Breslow [2], the simultaneous effect of several covariates and to see for what values of these covariates, if any, transplantation is likely to prolong survival. No effort is made to assess the clinical impression of the physicians in the program that the quality of life is decidedly improved by transplantation. The statistical methodology is also of interest in itself as an example of the usefulness of the general approach pioneered by Cox, in a somewhat nonstandard situation.

Briefly, patients are admitted to the Stanford program (given personal and family consent) after a medical conference at which it is decided that the patient is not likely to respond to other forms of therapy. A donor heart, matched on blood type, is then sought; this search has lasted anywhere from a few days to almost a year. Some patients die before a suitable heart is found. The total number of patients waiting when a heart is found is generally one or two, though it has been as high as seven. When more than one potential recipient is matched on blood type for a given donor heart, other criteria, such as the availability of the patient and recently, tissue typing, have been used. In a few cases a patient was passed over until a pulmonary infection, which might cause complications after transplantation, had cleared. However, in general it is thought that no serious systematic bias has been introduced by this selection of patients for the transplant operation. In fact, the belief of the physicians in the program is such that, if anything, *less* hardy patients tend to receive hearts preferentially over hardier ones.

Four patients to date have improved enough while waiting for a heart that a transplant was judged no longer necessary. These patients were then deselected and followed as long as possible (two have been lost to follow-up and two have died). For the purposes of the analysis they are treated like other patients who were not transplanted, as they offer evidence of the potential for survival without a new heart.

The data are presented and discussed in Section 2, and the method of analysis is explained in Section 3. Section 4 gives results of the analysis and checks of the fit of the model, and Section 5 contains further analyses which concentrate on investigating the effect of tissue typing.

2. DATA AND NOTATION

We modify the notation of [16] somewhat to conform closely throughout to that of [2]. Associated with each patient is a date of acceptance, T_1^*, and a date last seen, T_2, so that the survival time in days can be defined as $T_2 - T_1$. The survival time is said to be uncensored or censored depending on whether T_2 is the date of death or the closing date of the present study, April 1, 1974. (The survival time for two patients who were deselected was censored by dates prior to the closing date on which they were lost to follow-up.) For the M patients receiving a heart there is also a date of transplantation, T_3, where $T_1 \leq T_3 \leq T_2$. The waiting time in days for these patients is defined as $W = T_3 - T_1$, and their post-transplant survival is $Y = T_2 - T_3$, with $X = W + Y$. Let m denote the number of the M transplant survival times which are uncensored and $M - m$ the number

* John Crowley is Assistant Professor, Departments of Statistics and Human Oncology, University of Wisconsin, Madison, WI 53706. Marie Hu is Research Associate, Biostatistics Division, Stanford University, Stanford, CA 94305. This work was done with support from NIH Grants HL 13108-05, 2 T01 GM 00025-16, and 1 R01 GM 21215-01. The authors are grateful to Charles Bieber for supplying the data on the measures of HLA mismatch and for valuable conversations on the HLA system, and to the Stanford Heart Transplantation Program and its principal investigator, Eugene F. Dong, for encouraging us to explore and evaluate the data presented here. Thanks are also due to Byron Wm. Brown, Jr., for his help and advice, and to Rupert Miller for his careful reading of the manuscript. The valuable comments of the referees are also acknowledged.

March 1977, Volume 72, Number 357
Applications Section

1. Data March 1974

Patient	Date of birth	Date of acceptance	Date of transplant	Date last seen	Dead = 1 Alive = 0	Previous surgery	No. of mismatches	HLA-A2	Mismatch score	Reject
1	1/10/37	11/15/67		1/3/68	1	0				
2	3/2/16	1/2/68		1/7/68	1	0				
3	9/19/13	1/6/68	1/6/68	1/21/68	1	0	2	0	1.11	0
4	12/23/27	3/28/68	5/2/68	5/5/68	1	0	3	0	1.66	0
5	7/28/47	5/10/68		5/27/68	1	0				
6	11/8/13	6/13/68		6/15/68	1	0				
7	8/29/17	7/12/68	8/31/68	5/17/70	1	0	4	0	1.32	1
8	3/27/23	8/1/68		9/9/68	1	0				
9	6/11/21	8/9/68		11/1/68	1	0				
10	2/9/26	8/11/68	8/22/68	10/7/68	1	0	2	0	0.61	1
11	8/22/20	8/15/68	9/9/68	1/14/69	1	0	1	0	0.36	0
12	7/9/15	9/17/68		9/24/68	1	0				
13	2/22/14	9/19/68	10/5/68	12/8/68	1	0	3	0	1.89	1
14	9/16/14	9/20/68	10/26/68	7/7/72	1	0	1	0	0.87	1
15	12/4/14	9/27/68		9/27/68	1	1				
16	5/16/19	10/26/68	11/22/68	8/29/69	1	0	2	0	1.12	1
17	6/29/48	10/28/68		12/2/68	1	0				
18	12/27/11	11/1/68	11/20/68	12/13/68	1	0	3	0	2.05	0
19	10/4/9	11/18/68		12/24/68	1	0				
20	10/19/13	1/29/69	2/15/69	2/25/69	1	0	3	1	2.76	1
21	9/29/25	2/1/69	2/8/69	11/29/71	1	0	2	0	1.13	1
22	6/5/26	3/18/69	3/29/69	5/7/69	1	0	3	0	1.38	1
23	12/2/10	4/11/69	4/13/69	4/13/71	1	0	3	0	0.96	1
24	7/7/17	4/25/69	7/16/69	11/29/69	1	0	3	1	1.62	1
25	2/6/36	4/28/69	5/22/69	4/1/74	0	0	2	0	1.06	0
26	10/18/38	5/1/69		3/1/73[a,b]	0	0				
27	7/21/60	5/4/69		1/21/70[a]	1	0				
28	5/30/15	6/7/69	8/16/69	8/17/69	1	0	2	0	0.47	0
29	2/6/19	7/14/69		8/17/69	1	0				
30	9/20/24	8/19/69	9/3/69	12/18/71	1	0	4	0	1.58	1
31	10/4/14	8/23/69		9/7/69	1	0				
32	4/2/5	8/29/69	9/14/69	11/13/69	1	0	4	0	0.69	1
33	1/1/21	11/27/69	1/16/70	4/1/74	0	0	3	0	0.91	0
34	5/24/29	12/12/69	1/3/70	4/1/74	0	0	2	0	0.38	0
35	8/4/26	1/21/70		2/1/70	1	0				
36	5/1/21	4/4/70	5/19/70	7/12/70	1	0	2	0	2.09	1
37	10/24/8	4/25/70	5/13/70	6/29/70	1	0	3	1	0.87	1
38	11/14/28	5/5/70	5/9/70	5/9/70	1	0	3	0	0.87	0
39	11/12/19	5/20/70	5/21/70	7/11/70[c]	1	0				
40	11/30/21	5/25/70	7/4/70	4/1/74	0	1	4	0	0.75	0
41	4/30/25	8/19/70	10/15/70	4/1/74	0	1	2	0	0.98	0
42	3/13/34	8/21/70		8/23/70	1	0				
43	6/1/27	10/22/70		10/23/70	1	1				
44	5/2/28	11/30/70		1/8/71	1	1				
45	10/30/34	1/5/71	1/5/71	2/18/71	1	0	1	0	0.0	0
46	6/1/22	1/10/71	1/11/71	10/1/73	1	1	2	0	0.81	1
47	12/28/23	2/2/71	2/22/71	4/14/71	1	0	3	0	1.38	1
48	1/23/15	2/5/71		2/13/71	1	0				
49	6/21/34	2/15/71	3/22/71	4/1/74	0	1	4	0	1.35	0
50	3/28/25	2/15/71	5/8/71	10/21/73[c]	1	1				
51	6/29/22	3/24/71	4/24/71	1/2/72	1	0	4	1	1.08	1
52	1/24/30	4/25/71		8/4/71	1	0				
53	2/27/24	7/2/71	8/11/71	1/5/72[c]	1	0				
54	9/16/23	7/2/71		7/4/71	1	0				
55	2/24/19	8/9/71	8/18/71	10/8/71	1	0	2	0	1.51	1
56	12/5/32	9/3/71	11/8/71	4/1/74	0	0	4	0	0.98	0
57	6/8/30	9/13/71		2/8/72	1	0				
58	9/17/23	9/23/71	10/13/71	8/30/72	1	1	2	1	1.82	1
59	5/12/30	9/29/71	12/15/71	4/1/74	0	1	2	0	0.19	0
60	10/29/22	11/18/71	11/20/71	1/24/72[d]	1	0	3	0	0.66	1
61	5/12/19	12/4/71		12/5/71	1	0				
62	8/1/32	12/9/71		2/15/72[d]	1	0				
63	4/15/39	12/12/71	1/7/72	4/1/74	0	0	3	1	1.93	0
64	4/9/23	2/1/72	3/4/72	9/6/73	1	1	1	0	0.12	0
65	11/19/20	3/6/72	3/17/72	5/22/72	1	0	2	0	1.12	1
66	1/2/19	3/20/72		4/20/72	1	0				
67	9/3/52	3/23/72	5/18/72	1/1/73	1	0	3	0	1.02	0
68	1/10/27	4/7/72	4/9/72	6/13/72	1	0	3	1	1.68	1
69	6/5/24	6/1/72	6/10/72	4/1/74	0	0	2	0	1.20	0
70	6/17/19	6/17/72	6/21/72	7/16/72	1	0	3	1	1.68	1
71	2/22/25	7/21/72	8/20/71	4/1/74	0	0	3	0	0.97	0
72	11/22/45	8/14/72	8/17/72	4/1/74	0	0	3	1	1.46	0

Table 1. (Continued)

Patient	Date of birth	Date of acceptance	Date of transplant	Date last seen	Dead = 1 Alive = 0	Previous surgery	No. of mismatches	HLA-A2	Mismatch score	Reject
73	5/13/16	9/11/72	10/7/72	12/9/72	1	0	3	1	2.16	1
74	7/20/43	9/18/72	9/22/72	10/4/72	1	0	1	0	0.61	0
75	7/25/20	9/29/72		9/30/72	1	0				
76	9/3/20	10/4/72	11/18/72	4/1/74	0	1	3	1	1.70	0
77	8/27/31	10/6/72		10/26/72	1	0				
78	2/20/24	11/3/72	5/31/73	4/1/74	0	0	3	0	0.81	0
79	2/18/19	11/30/72	2/4/73	3/5/73	1	0	2	0	1.08	1
80	6/27/26	12/6/72	12/31/72	4/1/74	0	1	3	0	1.41	0
81	2/21/20	1/12/73	1/17/73	4/1/74	0	0	4	1	1.94	0
82	8/19/42	11/1/71		1/1/73[a,b]	0	0				
83	10/4/19	1/24/73	2/24/73	4/13/73	1	0	4	0	3.05	0
84	5/13/30	1/30/73	3/7/73	12/29/73	1	0	4	0	0.60	1
85	2/13/25	2/6/73		2/10/73	1	0				
86	3/30/24	3/1/73	3/8/73	4/1/74	0	0	3	1	1.44	0
87	12/19/26	3/21/73	5/19/73	7/8/73	1	0	2	0	2.25	1
88	11/16/18	3/28/73	4/27/73	4/1/74	0	0	3	0	0.68	0
89	3/19/22	4/5/73	8/21/73	10/28/73	1	0	4	1	1.33	1
90	3/25/21	4/6/73	9/12/73	10/8/73	1	1	3	1	0.82	0
91	9/8/25	4/13/73		3/18/74	1	0				
92	5/3/28	4/27/73	3/2/74	4/1/74	0	0	1	0	0.16	0
93	10/10/25	7/11/73	8/7/73	4/1/74	0	0	2	0	0.33	0
94	11/11/29	9/14/73	9/17/73	2/25/74	1	1	3	0	1.20	1
95	6/11/33	9/22/73	9/23/73	10/7/73[c]	1	0				
96	2/9/47	10/4/73	10/16/73	4/1/74	0	0	2	0	0.46	0
97	4/11/50	11/22/73	12/12/73	4/1/74	0	0	3	1	1.78	0
98	4/28/45	12/14/73	3/19/74	4/1/74	0	0	4	1	0.77	0
99	2/24/24	12/25/73		1/14/74	1	0				
100	1/31/39	2/22/74	3/31/74	4/1/74	0	1	3	0	0.67	0
101	8/25/24	3/2/74		4/1/74	0	0				
102	10/30/33	3/22/74		4/1/74	0	0				
103	5/20/28	9/13/67		9/18/67	1	0				

[a] Deselect.
[b] Lost to follow-up.
[c] Not typed.
[d] First heart failed, retransplanted 1/9/72.

NOTE: Blanks indicate nontransplant patients or missing data. HLA-A2 denotes mismatch on HLA-A2. Reject denotes death by rejection, yes = 1, no = 0.
Sources: Dr. Charles Bieber, Joan Miller of Stanford Heart Transplantation Program, and [3].

which are censored. For the N patients who did not receive a heart it is known only that T_3 is greater than T_2, or that W is greater than X. Let n of the N nontransplant survival times be uncensored and $N - n$ censored. Here $n = 30$, $N = 34$, $m = 45$, and $M = 69$.

Also associated with each patient is a vector of covariates. The covariates investigated in the present study include transplant status (Z_0), waiting time to transplant (Z_1), calendar time at transplant (Z_2), age at acceptance into the program (Z_3), age at transplant (Z_4), previous open-heart surgery (Z_5, Z_6), and three measures of the degree to which donor and recipient are mismatched for tissue type (Z_7, Z_8, Z_9). The relevant data are given in Table 1.

Transplant status is simply an indicator of whether or not a patient has received a heart (0 for nontransplants, 1 for transplants). Analysis of this variable alone by the methods of Section 3 is virtually equivalent to the Mantel-Cox test performed in [16] and [13] to test the efficacy of transplantation. Age at acceptance (in years) is included because of the possibility that pretransplant risks vary with age, and age at transplant (in years) because the response to the operation and subsequent therapy may also depend on age. Calendar time at transplant (in days from October 1, 1967) is used to detect possible improvement in the program itself over time, while waiting time to transplant (in days) is analyzed mainly for comparison with the techniques in [3]. (The reference [3] also includes, among other variables, an analysis of the relation of sex to posttransplant survival, which is not done in the present study both because of the negative findings in [3] and because of the small number of women accepted to date.)

Previous open-heart surgery and the three measures of tissue type mismatching deserve a brief explanation. A person's tissue type is determined by the presence of what are called HLA antigens expressed on cell membranes and genetically controlled by a pair of codominant alleles at each of two loci on the chromosomes. At the time of this study the typing laboratory at Stanford typed for 13 antigens for the A locus and 15 for the B locus. These numbers are subject to change as the classifications are further subdivided. This HLA system is thought (see [1] and [10] for opposing views) to play a part in the rejection of transplanted organs, though with the large number of possible antigens the question of matching is quantitative and not qualitative: no recipient to date has been perfectly matched.

Three measures of mismatching are studied here. The number of mismatches (Z_7) is the number of donor alleles

with no match in the recipient (1 through 4). Mismatch on HLA-A2 (Z_8) is a dichotomous variable which is 1 when the donor has the antigen HLA-A2 and the recipient has neither HLA-A2 nor the similar HLA-A28, and 0 otherwise. (HLA-A2 is an antigen to which kidney transplant patients often develop antibodies due to blood transfusions.) The mismatch score (Z_9) is a continuous score derived from antibody responses of pregnant women by Charles Bieber of Stanford University in a manner somewhat similar to that in [14] and to be described in detail in a future article. The derivation involves the estimation of many parameters from a relatively small set of data (120 women) and converting what are basically probabilities of antibody response into a reasonable measure of the strength of response. The mismatch score is studied here both to assess its present usefulness and to indicate whether the manner of measuring HLA mismatch is a fruitful field of further inquiry.

Open-heart operations are accompanied by transfusions of blood from several donors, presumably with many different HLA antigens. There is a possibility that this produces a minimal antibody response which then inhibits the antibody response when the patient receives a new heart some time later. These patients may also have a different general level of health. Two dichotomous variables (0 for no previous open-heart survey, 1 otherwise) were included to investigate these hypotheses, one (Z_5) reflecting pretransplant risk, the other (Z_6) for posttransplant effect.

A distinctive feature of the present problem is that some of the covariates are subject to change at a random point in time. The covariate for transplant status is taken as an indicator (coded 0 before the point of transplant, 1 after transplant). The pretransplant risks, age at acceptance, and previous surgery (Z_5) do not change with time. All the other covariates studied are viewed as being zero before transplant but changing from zero to the actual value of that particular covariate at the time of transplant.

3. THE METHOD OF ANALYSIS

Clark, et al. [4] included an analysis of the efficacy of transplantation in which the posttransplant survival of transplant patients was observed to compare favorably with the survival experience of nontransplant patients. Responding to the criticism by Gail [8] that the nontransplant patients are preferentially the less hardy ones, being unable to survive long enough for a suitable heart to be found, Turnbull, Brown, and Hu [16] and Mantel and Byar [13] reanalyzed the data in several ways. In both of these papers the pretransplant survival experience of all patients was used in various ways to assess whether or not posttransplant survival was longer than expected, with the general conclusion of a slight but statistically nonsignificant benefit of transplantation.

One of the analyses in [16], which did not fit the data well and was used mainly for comparison and as a step in building a better model, was based on an exponential model in which each patient is assumed to have a constant hazard λ which changes at transplant to $\lambda \exp(\beta_0)$. More formally, the hazard of the jth patient as a function of time from acceptance is $\lambda \exp(\beta_0 Z_{0j}(t))$, where

$$\begin{aligned} Z_{0j}(t) &= 0 \ , \quad \text{if} \quad t < W_j \ , \\ &= 1 \ , \quad \text{if} \quad t \geq W_j \ . \end{aligned}$$

That is, the covariate Z_{0j} depends on time.

This exponential model can be generalized in two ways. First, since there may be many factors influencing survival, the hazard can be assumed to be a function of a (column) vector $\mathbf{Z}$ of possibly time-dependent covariates, an easy and useful way to do so being the log-linear model discussed in [2, 9, 15], in which the hazard for the jth patient is expressed as $\lambda \exp(\boldsymbol{\beta}'\mathbf{Z}_j(t))$. In addition, the hazard can be assumed to be a largely arbitrary function of time, as in Cox [6], the model for the hazard being

$$\lambda(t) \exp(\boldsymbol{\beta}'\mathbf{Z}_j(t)) \ . \tag{3.1}$$

This model has the advantage of allowing an assessment of the effect of covariates on survival which is valid over a wide class of underlying hazard functions, important in the present case because of the difficulty in finding a parametric hazard with an adequate fit to the data. The particular form of the dependence of the hazard on the covariates has proved both reasonable for survival data and easy to use computationally, though of course other models are possible and might in particular circumstances be more nearly correct.

Let $t_1 < t_2 < \ldots < t_k$ be the k distinct times to death among the $N + M$ survival times, and assume for the moment that each t_i is of multiplicity 1, so that $k = n + m$. Label the patients so that $\mathbf{Z}_i(t)$ is the vector of covariates for the patient dying at t_i, and denote by R_i the set of labels for patients whose survival times are equal to or greater than t_i. Then the argument of Cox [6] is that, given (3.1) and conditional on R_i and the fact that a death occurs at t_i, the probability that $Z_i(t)$ is as observed is the ratio of hazards

$$\frac{\lambda(t_i) \exp(\boldsymbol{\beta}'\mathbf{Z}_i(t_i))}{\sum_{j \in R_i} \lambda(t_i) \exp(\boldsymbol{\beta}'\mathbf{Z}_j(t_i))} = \frac{\exp(\boldsymbol{\beta}'\mathbf{Z}_i(t_i))}{\sum_{j \in R_i} \exp(\boldsymbol{\beta}'\mathbf{Z}_j(t_i))} . \tag{3.2}$$

Multiplying these terms for each death point gives the conditional likelihood

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{k} \frac{\exp(\boldsymbol{\beta}'\mathbf{Z}_i(t_i))}{\sum_{j \in R_i} \exp(\boldsymbol{\beta}'\mathbf{Z}_j(t_i))} . \tag{3.3}$$

Cox suggested that standard likelihood methods then be used to estimate $\boldsymbol{\beta}$ and to perform tests of significance.

Cox's arguments are more intuitive than formal, but they have been given justification when no covariates depend on time in terms of a rank-like procedure by Kalbfleisch and Prentice [11] and in terms of an approximate likelihood by Breslow [2]. The rank argument is not applicable in the case of time-dependent covariates,

2. Covariance Analysis, All 103 Patients

Line	Parameter	Z_0	Z_1	Z_2	Z_3	Z_4	Z_5	Z_6
1	$\hat{\beta}'$	.035						
	SE ($\hat{\beta}'$)	.298						
	p	—						
2	$\hat{\beta}'$	−.005	.00183					
	SE ($\hat{\beta}'$)	.322	.00522					
	p		—					
3	$\hat{\beta}'$	.438		−.000323				
	SE ($\hat{\beta}'$)	.432		.000263				
	p			.109				
4	$\hat{\beta}'$	−.103			.0316			
	SE ($\hat{\beta}'$)	.309			.0145			
	p	.371			.015			
5	$\hat{\beta}'$	−2.671				.0567		
	SE ($\hat{\beta}'$)	1.143				.0224		
	p					.006		
6	$\hat{\beta}'$	−2.196			.0117	.0452		
	SE ($\hat{\beta}'$)	1.365			.0180	.0286		
	p				.258	.057		
7	$\hat{\beta}'$	−2.339				.0512	−.900	
	SE ($\hat{\beta}'$)	1.069				.0210	.379	
	p					.008	.009	
8	$\hat{\beta}'$	−2.319				.0520		−.810
	SE ($\hat{\beta}'$)	1.084				.0211		.447
	p					.007		.035
9	$\hat{\beta}'$	−2.371				.0511	−1.111	.301
	SE ($\hat{\beta}'$)	1.078				.0210	.729	.855
	p	.014				.008	.065	.363

NOTE: $\hat{\beta}'$ = estimates of coefficients, SE ($\hat{\beta}'$) = estimated standard errors of $\hat{\beta}'$, p = one-sided p values, Z_0 = transplant status, Z_1 = waiting time to transplant, Z_2 = calendar time at transplant, Z_3 = age at acceptance, Z_4 = age at transplant, Z_5 = previous surgery (pretransplant), Z_6 = previous surgery (posttransplant).

but an adaptation of the approximate likelihood of Breslow is possible. This approach has the advantage of handling ties easily and providing a simple and reasonable estimate at the standard condition $\beta = \mathbf{0}$ of the integrated hazard function

$$\Lambda(t) = \int_0^t \lambda(s)ds$$

and the survival curve $F(t) = \exp(-\Lambda(t))$.

Thus let $t_1 < \ldots < t_k$ be as before, and denote by d_i the number of deaths at t_i, with $\sum_{i=1}^k d_i = n + m$. Let $\mathbf{S}_i(t)$ be the sum of the covariates of those patients dying at t_i, and let R_i/R_{i+1} be the set of labels for patients whose survival times are in (t_i, t_{i+1}). In order to simplify the likelihood in such a manner as to allow time-invariant inference, Breslow [2] suggested approximating the data by moving all censored observations in (t_i, t_{i+1}) to t_i. With the additional approximation that the waiting times (W) for transplant patients which are in (t_i, t_{i+1}) are also moved to t_i, with appropriate modification of $\mathbf{Z}(t)$, the likelihood of the data is

$$\prod_{i=1}^{k} \left\{ [\lambda(t_i)]^{d_i} \exp(\beta'\mathbf{S}_i(t_i)) \cdot \prod_{j \in R_i/R_{i+1}} \exp\left[-\int_0^{t_i} \exp(\beta'\mathbf{Z}_j(s))\lambda(s)ds \right] \right\} . \quad (3.4)$$

To proceed without restriction of $\lambda(t)$ to a specific family of distributions, we again follow Breslow and assume that $\lambda(t)$ is constant between deaths,

$$\lambda(t) = \exp(\alpha_i) \ , \quad t \in (t_{i-1}, t_i] \ ,$$

where $t_0 = 0$. Denoting the log of (3.4) by $\ell(\alpha, \beta)$, we can use the methods of [2] to show that $\ell(\alpha, \beta)$ is maximized for α_i in terms of β by

$$\exp(\hat{\alpha}_i) = d_i / [(t_i - t_{i-1}) \sum_{j \in R_i} \exp(\beta'\mathbf{Z}_j(t))] \ , \quad (3.5)$$

which gives the log likelihood for β as

$$\ell(\beta) = \sum_{i=1}^{k} [\beta'\mathbf{S}_i(t_i) - d_i \ln \sum_{j \in R_i} \exp(\beta'\mathbf{Z}_j(t_i))] \ . \quad (3.6)$$

An estimate $\hat{\beta}$ of β is the solution of the system $\partial\ell(\beta)/\partial\beta = U(\beta) = 0$, with variance-covariance matrix estimated by the inverse of the matrix

$$-\partial^2(\beta)/\partial\beta^2|_{\beta=\hat{\beta}} = I(\hat{\beta}) \ .$$

Large sample tests of the composite hypothesis $\beta_2 = \mathbf{0}$, where $\beta' = (\beta_1', \beta_2')$, can be done by comparing

$$2[\ell((\hat{\beta}_1', \hat{\beta}_2'))' - \ell(\hat{\beta}_1)] \quad (3.7)$$

to a χ^2 distribution with degrees of freedom equal to the dimension of β_2. When $\beta_2 = \beta$, an appropriate large sample criterion is

$$(U(\mathbf{0}))'(I(\mathbf{0}))^{-1}(U(0)) \ , \quad (3.8)$$

an approximate χ^2 with degrees of freedom equal to the dimension of β.

Large sample tests of simple hypotheses like $\beta_i = 0$ can be done using (3.7), (3.8) where appropriate, or by comparing $\hat{\beta}_i$ to its estimated standard error as found in $(I(\hat{\beta}))^{-1}$. Empirically, these three methods are quite similar, and only the latter will be reported. Large sample theory has been carried out in detail for (3.8), with one covariate, transplant status, in [7].

An estimate of $\Lambda(t)$ is given at t_l by

$$\hat{\Lambda}(t_l) = \sum_{i=1}^{l} \{d_i/[\sum_{j \in R_i} \exp(\hat{\beta}' \mathbf{Z}_j(t_i))]\} ,$$

with linear interpolation in (t_{l-1}, t_l). Also

$$\hat{F}(t) = \exp(-\hat{\Lambda}(t)) .$$

4. EMPIRICAL RESULTS

4.1 The Model

The data in Table 1 were analyzed by the methods of Section 3 to discover for which values of the covariates transplantation is likely to be of benefit. Ties between times to transplantation and times to death were resolved conservatively by ordering deaths first, except for one patient who died on the day of his operation. The results of fitting various sets of covariates are summarized in Tables 2 and 3. All 103 patients were used in

3. Covariance Analysis, 99 Typed Patients

Line	Parameter	Z_0	Z_4	Z_5	Z_7	Z_8	Z_9
1	$\hat{\beta}'$	−2.652	.0550	−.903			
	SE ($\hat{\beta}'$)	1.119	.0219	.402			
	p		.006	.012			
2	$\hat{\beta}'$	−2.344	.0564	−.915	−.141		
	SE ($\hat{\beta}'$)	1.189	.0221	.402	.173		
	p		.005	.012	—		
3	$\hat{\beta}'$	−2.729	.0553	−.902		.239	
	SE ($\hat{\beta}'$)	1.116	.0215	.401		.374	
	p		.005	.012		.264	
4	$\hat{\beta}'$	−3.139	.0540	−.866			.432
	SE ($\hat{\beta}'$)	1.178	.0220	.402			.283
	p		.008	.016			.065
5	$\hat{\beta}'$	−3.489	.0591				.484
	SE ($\hat{\beta}'$)	1.240	.0235				.283
	p		.006				.044

NOTE: See NOTE for Table 2. Z_7 = number of mismatches, Z_8 = mismatch on HLA-A2, Z_9 = mismatch score.

the construction of Table 2, but as four patients were not tissue-typed, Table 3 is based on 99 patients. There was no trouble with convergence to a solution $\hat{\beta}$ of the system $U(\beta) = 0$ using a Newton-Raphson algorithm. Estimates of β_i and estimated standard errors are given for all covariates in each model. However, significance tests for transplant status are meaningful only in the absence of time-dependent covariates (in whose presence β_0 is analogous to an intercept of a regression equation) and are thus deleted.

When only transplant status is considered (Table 2, Line 1), transplantation appears to be associated with a slight *increase* in risk, contrasting with the reverse finding on an earlier data set in [13] and [16]. (The difference is explainable in part by several recent long waiting times, which make the estimate of pretransplant survival slightly more optimistic.) Waiting time to transplant, judged in [3] to be a significant predictor of posttransplant survival, was not significant in the present analysis. A correlation as calculated in [3] could result from a model with a decreasing hazard function before transplantation; in the present analysis it is the *change* in $\lambda(t)$ at transplant, whatever its shape before, that is being analyzed. (Further differences could be the result of the acquisition of new data. See Section 5.) Calendar time at transplant is also less striking here than in [3], and any effect present is explained by the slight trend in time toward accepting younger patients (Table 2, Line 3). (When the analysis represented by Table 2, Line 3 was performed on the data in [3], a one-sided p value of .07 was obtained, in reasonable agreement with their values of about .02.)

Age at acceptance, which is not time-dependent, and age at transplant, which is, are nonetheless highly correlated; their effects are thus difficult to separate. However, both must be analyzed to assess the possibility that patients enter the study with risks which depend on age, and also respond to the transplant operation in a *different* age-dependent manner. The analyses of Table 2, Lines 4–6 indicate that age at acceptance as a pretransplant risk factor adds significantly to the model (and changes the direction indicated for the effect of transplant), but in a way which is largely explained by age at transplant as a differential posttransplant risk. The conclusion is that age has relatively little effect on survival of the patients before transplantation, but that older patients are less able to cope with the trauma of surgery and the ensuing difficulties with rejection of the donor heart than are younger ones. Previous surgery also seemed to be an important factor (Table 2, Lines 7–9), but as an indicator of pretransplant health, not as a posttransplant factor.

The effect of adding measures of HLA mismatch to the model is shown in Table 3. The mismatch score devised by Bieber appears to be the most promising predictor of survival of the measures studied here (Table 3, Lines 2–5). The number of mismatches influences survival in a way opposite from that anticipated, and the harmful effect of a mismatch on HLA-A2 appears modest (Table 3, Lines 2, 3).

Table 4 gives estimates of the relative risk at transplant (the ratio of posttransplant and pretransplant hazard) for several values of age at transplant using the models of Table 2, Lines 5, 6, and 7. Comparison of Parts a and b of Table 4 reinforces the conclusion that age at acceptance has relatively little effect over the range of the data on posttransplant survival. In addition, Parts d and e of Table 4 show cut-off points for the mismatch score, below

4. Estimation of Relative Risk and Cut-Off Points for Mismatch Score

Source	Age at transplant	Estimation
Relative risk at transplant exp $(\hat{\beta}_0 + \hat{\beta}_4 Z_4)$		
a. Table 2, Line 5	10	.22
	30	.38
	40	.67
	47.1	1.00
	50	1.18
	60	2.08
b. Table 2, Line 6	20	.27
	30	.43
	40	.68
	48.6	1.00
	50	1.07
	60	1.68
c. Table 2, Line 7	20	.27
	30	.45
	40	.74
	45.9	1.00
	50	1.23
	60	2.06
Cut-off for mismatch score		
d. Table 3, Line 4	35	2.89
	40	2.27
	45	1.64
	50	1.02
	55	.39
	58.1	0
e. Table 3, Line 5	40	2.32
	45	1.71
	50	1.10
	55	0.49
	59.1	0

which transplantation is advisable, estimated from the models of Table 3, Lines 4 and 5.

4.2 Checks of the Model

There are several possible ways to check the adequacy of fit of the model, that is, the way in which adjustments are made for covariates.

An often useful check on a model, as well as a valuable preliminary tool, is stratification. With the number of variables in the present study and the moderate sample size, stratification cannot be carried very far. Grouping into the age groups [8, 40), [40, 50), and [50, 65) to obtain groups of size 21, 48, and 34, the model of Table 2, Line 1, gave estimates of β_0 of $-.43$, $-.22$, and .60. While none of these estimates are significantly different from 0, the trend with age is confirmed. In fact, a quadratic effect of age is suggested. Addition of a term for age at transplant squared to the models of Table 2, Line 5 did not add significantly to the likelihood, however, and did not change the estimates of exp $(\beta' \mathbf{Z})$ appreciably.

Use can also be made of the concept of residuals based on the work of Cox and Snell [5]. As elucidated by Kalbfleisch [12], in a case where no covariates depend on time, if $\lambda(t)$ exp $(\beta' \mathbf{Z}_j)$ is the true hazard for the jth patient, then the integrated hazard forms a random sample $\Lambda(t)$ exp $(\beta' \mathbf{Z}_j)$ evaluated at the survival times from a unit exponential random variable. After fitting a model and estimating $\Lambda(t)$ and β, $\hat{\Lambda}(X_j)$ exp $(\hat{\beta}' \mathbf{Z}_j)$ should behave approximately as censored unit exponentials. Generalizing to the present case, if X_j are survival times, W_j and Y_j waiting times and posttransplant survival times, and $\hat{\Lambda}$ and $\hat{\beta}$ the estimates of Section 3, then the residuals

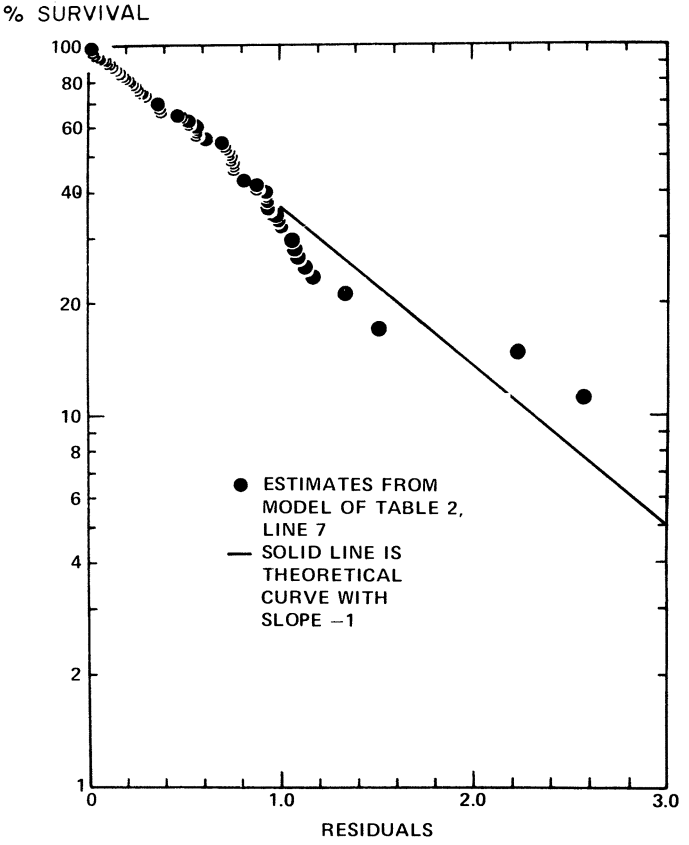
$$\hat{\Lambda}(X_j) \exp (\hat{\beta}' \mathbf{Z}_j(X_j))$$

for nontransplant patients, and

$$\hat{\Lambda}(W_j) \exp (\hat{\beta}' Z_j(W_j^-)) + [\hat{\Lambda}(W_j + Y_j) - \hat{\Lambda}(W_j)] \exp (\hat{\beta}' \mathbf{Z}_j(W_j))$$

for transplant patients, should behave approximately as censored unit exponentials. Estimating a survival curve by the method of Section 3 (no covariates) from the residuals resulting from the model of Table 2, Line 7, results in a curve which is quite close to e^{-t} (Figure A),

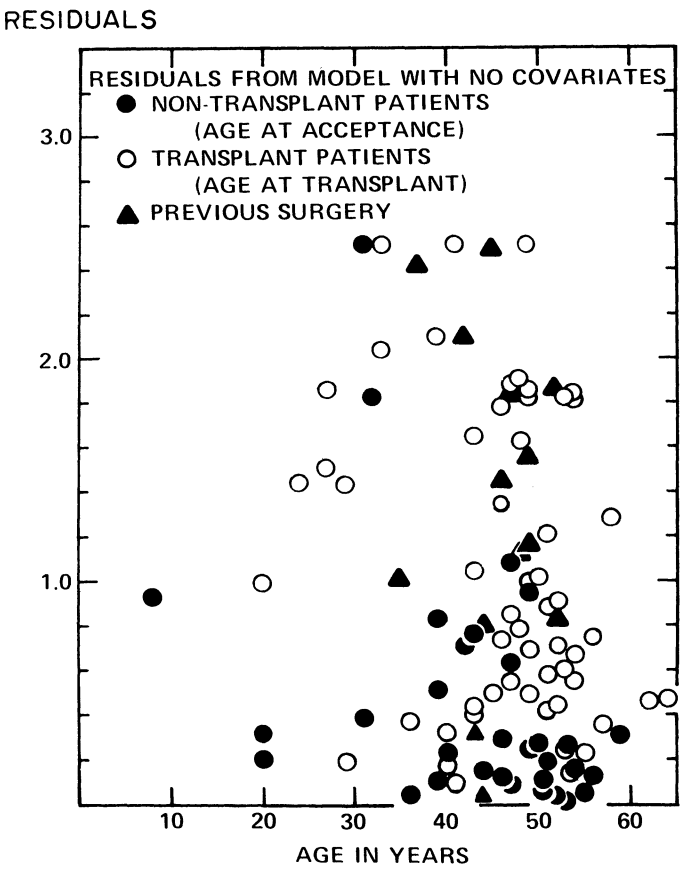
A. Estimate of Survival Curve from Residuals



but it should be noted that the analogous curve from a model with no covariates (and without regard to transplant status) is *exactly* e^{-t}.

Another use of residuals is in plotting, though here a decision must be made as to how censored residuals should be handled. Preliminary plots suggested that adding 1 (the conditional expected additional survival for unit exponentials) was too severe, so ln (2) (conditional median additional survival) was added to each censored residual.

B. Plot of Residuals Vs. Age



The resulting plot (Figure B) from the model with no covariates shows clearly that previous survey needs to be adjusted for and suggests that older patients have too many small residuals. A plot derived from fitting the model of Table 2, Line 7 shows no obvious trends. Similar conclusions emerge from plots of the residuals of Table 3, Lines 1 and 5, against the mismatch score.

5. ADDITIONAL ANALYSES

To this point we have used all the available data to assess the efficacy of transplantation and the effect of covariates by means of modeling the change in hazard at transplantation. Another approach to testing the significance of covariates (other than transplant status) which is independent of how the risk changes at transplantation, and more in the spirit of [3], is to use only transplant patients. The methods of Section 3 are applicable, though in this case none of the covariates depend on time. The model states that the hazard, *as a function of time from transplantation,* is

$$\lambda(t) \exp(\beta'\mathbf{Z}) .$$

Several models were fitted, with results which were quite similar to corresponding results in Tables 2 and 3. The effect of waiting time to transplant, while in a direction consistent with that found in [3], was not judged to be an important covariate; this is probably due to the fact that, with the accumulation of new data, the pretransplant hazard seems nearly constant over a range which contains most of the waiting times. Calendar time at transplant is more striking than in the analysis of Section 4 (and more consistent with [3]), but its effect is again attenuated by other variables.

The potential importance of the HLA typing justifies subjecting the previous surgery variable and the three measures of mismatching to an additional analysis. The antibody response is thought to be linked to only one of the possible causes of death, rejection of the donor heart (other observed causes to date being surgical, a lymphoma, a kidney failure, hepatitis, stroke, and preexisting infection). Thus the significance of HLA-related covariates in prolonging the time from transplantation to rejection might be tested more sensitively by treating other causes of death as censoring mechanisms. Categorizing deaths by independent causes as required by this analysis is always difficult, and it is perhaps even more so in this case because all transplanted patients exhibit an immune response and are maintained after the operation on immunosuppressants (at considerable attendant risk of infection). However, many patients do die of a sudden worsening of the donor organ attributable to rejection, while others die of seemingly unrelated causes while on standard immunotherapy. The really questionable cases are few and are unlikely to alter the results greatly. Thus we can proceed as above with the 65 transplanted and typed patients, but redefining the endpoint as death by rejection. (This necessitates changing the survival time of one patient from 65 to 50 days; he rejected the donor heart at 50 days but was immediately retransplanted, to survive another 15 days, and is for this analysis considered to have died by rejection at 50 days.)

The results are shown in Table 5. Waiting time is again not a key variable (Table 5, Lines 1 and 8), and the apparent importance of calendar time is again in part explained by other variables (Table 5, Lines 2, 4, and 9). The effect of previous surgery does not appear to be strengthened, corroborating its interpretation as a measure of health and not of the immune response (Table 5, Lines 5 and 7). The HLA typing, in particular the mismatch score, appears more important than before, indicating that typing may indeed play a critical role in preventing acute rejection and thus prolonging survival.

A fundamentally different type of analysis was suggested by a referee to investigate the effect of transplantation with a low mismatch score (Z_9). In this analysis a patient is censored if he has a high mismatch score at transplantation, as well as by the end of the study. Defining low mismatch score as less than or equal to one, and considering only the younger patients (less than 50 at acceptance), the methods of Section 3 (with transplant status the only covariate) give an estimate of relative risk $\exp(\hat{\beta}_0)$ of .76. The corresponding result for those 50 and over is $\exp(\hat{\beta}_0) = 1.30$. While neither is significantly different from the null value of 1, the direction of the effect is consistent with the analyses of Table 3, Lines 4 and 5, and Table 5, Lines 6 and 7.

5. Covariance Analysis, Transplant Patients, Death by Rejection

Line	Parameter	Z_1	Z_2	Z_4	Z_6	Z_7	Z_8	Z_9
1	$\hat{\beta}'$	−.00905						
	SE ($\hat{\beta}'$)	.00789						
	p	.125						
2	$\hat{\beta}'$		−.000374					
	SE ($\hat{\beta}'$)		.000327					
	p		.127					
3	$\hat{\beta}'$			.107				
	SE ($\hat{\beta}'$)			.031				
	p			.000				
4	$\hat{\beta}'$		−.000061	.105				
	SE ($\hat{\beta}'$)		.000331	.032				
	p		.425	.001				
5	$\hat{\beta}'$			.0947	−1.048			
	SE ($\hat{\beta}'$)			.0297	.622			
	p			.001	.046			
6	$\hat{\beta}'$			.109				1.043
	SE ($\hat{\beta}'$)			.033				.357
	p			.000				.002
7	$\hat{\beta}'$			.0986	−.974			1.015
	SE ($\hat{\beta}'$)			.0318	.623			.358
	p			.001	.059			.002
8	$\hat{\beta}'$	−.00616		.107				1.015
	SE ($\hat{\beta}'$)	.00759		.033				.359
	p	.212		.001				.002
9	$\hat{\beta}'$		−.000266	.104				1.129
	SE ($\hat{\beta}'$)		.000337	.034				.381
	p		.215	.001				.002
10	$\hat{\beta}'$			.106		.046		
	SE ($\hat{\beta}'$)			.031		.199		
	p			.000		.409		
11	$\hat{\beta}'$			.106			.577	
	SE ($\hat{\beta}'$)			.032			.427	
	p			.000			.088	
12	$\hat{\beta}'$			.110			−.056	1.063
	SE ($\hat{\beta}'$)			.034			.471	.393
	p			.000			—	.004

NOTE: See NOTE's for Tables 2 and 3.

6. CONCLUDING REMARKS

The general conclusion of these analyses is that transplantation can prolong survival for certain younger patients if a suitably matched heart can be found. This conclusion must of course be treated with caution. Possible bias in the selection from the queue of patients for transplantation, thought to be minimal, is still a point of potential criticism, since randomization is not employed. It is also conceivable that hardy patients tend to have other characteristics which make donor hearts more easily found.

The program has now been subjected to several statistical analyses, albeit different ones and on more data each time. Though the mismatch score seems important enough to warrant its use, no claim is made that an optimal measure of the degree of mismatching has been found: the data from which it is derived are relatively few and are based on the probability of an antibody response, while the rejection phenomenon is more related to the degree of response. Further study is indicated to attempt to improve on the way HLA mismatch is measured.

Another conclusion from this analysis is that the techniques of Cox [6] and Breslow [2] are promising, easily adaptable tools for the covariance analysis of transplant data.

[*Received November 1974. Revised August 1976.*]

REFERENCES

[1] Belzer, F.O., Perkins, H.A., Fortmann, J.C., Kountz, S.L., Salvatierra, O., Cochrum, K.C., and Payne, R., "Is HL-A Typing of Clinical Significance in Cadaver Renal Transplantation," *Lancet*, 1 (April 1974), 774–7.

[2] Breslow, Norman, "Covariance Analysis of Censored Survival Data," *Biometrics*, 30 (March 1974), 89–99.

[3] Brown, Byron Wm., Jr., Hollander, Myles, and Korwar, R.M., "Nonparametric Tests of Independence for Censored Data with Applications to Heart Transplant Studies," in Frank Proschan and R.J. Serfling, eds., *Reliability and Biometry:*

Statistical Analysis of Lifelength, Society for Industrial and Applied Mathematics, Philadelphia, Penn., 1974, 327–54.

[4] Clark, D.A., Stinson, E.B., Griepp, R.B., Schroeder, J.S., Shumway, N.E., and Harrison, D.C., "Cardiac Transplantation in Man, VI. Prognosis of Patients Selected for Cardiac Transplantation," *Annals of Internal Medicine*, 75 (July 1971), 15–21.

[5] Cox, D.R., and Snell, E.J., "A General Definition of Residuals (with discussion)," *Journal of the Royal Statistical Society, Ser. B*, 30 (June 1968), 248–75.

[6] ———, "Regression Models and Life Tables (with discussion)," *Journal of the Royal Statistical Society, Ser. B*, 34 (June 1972), 187–220.

[7] Crowley, John, "Asymptotic Normality of a New Nonparametric Statistic for Use in Organ Transplant Studies," *Journal of the American Statistical Association*, 69 (December 1974), 1006–11.

[8] Gail, Mitchell H., "Does Cardiac Transplantation Prolong Life? A Reassessment," *Annals of Internal Medicine*, 76 (May 1972), 815–7.

[9] Glasser, M., "Exponential Survival with Covariance," *Journal of the American Statistical Association*, 62 (June 1967), 651–8.

[10] Hors, J., Feingold, O., Fradelizi, D., and Dausset, J., "Critical Evaluation of Histocompatibility in 179 Renal Transplants," *Lancet*, 1 (March 1971), 609–12.

[11] Kalbfleisch, John D., and Prentice, Ross L., "Marginal Likelihoods Based on Cox's Regression and Life Model," *Biometrika*, 60 (August 1973), 267–78.

[12] ———, "Some Extensions and Applications of Cox's Regression and Life Model," paper presented at Joint Statistical Meetings, Tallahassee, Fla., March 1974.

[13] Mantel, Nathan, and Byar, David P., "Evaluation of Response-Time Data Involving Transient States: An Illustration Using Heart-Transplant Data," *Journal of the American Statistical Association*, 69 (March 1974), 81–6.

[14] Mickey, M.R., Mittal, K.K., Opelz, G., and Terasaki, P.I., "Transplant Compatibility Predictions Based on Antibody-Antigen Profile Relationships," in Paul Terasaki, ed., *Histocompatibility Testing 1972*, Baltimore: Williams and Wilkins, 1972, 485–91.

[15] Prentice, Ross L., "Exponential Survivals with Censoring and Explanatory Variables," *Biometrika*, 60 (August 1973), 279–88.

[16] Turnbull, Bruce W., Brown, Byron Wm., Jr., and Hu, Marie, "Survivorship Analysis of Heart Transplant Data," *Journal of the American Statistical Association*, 69 (March 1974), 74–80.