13

Assessing the Effect on Survival of Kidney Transplantation with Higher-Risk Donor Kidneys

Douglas E. Schaubel and John D. Kalbfleisch

University of Michigan, Ann Arbor, MI

13.1 Introduction

Patients with end-stage renal disease (ESRD; also known as kidney failure) must address the deficit in renal function through dialysis or kidney transplantation. Kidney transplantation has been repeatedly demonstrated to be superior to dialysis in terms of patient survival (Schaubel et al., 1995; Wolfe et al., 1999; Rabbat et al., 2000). However, there are tens of thousands more patients in need of a kidney transplant than there are available donor kidneys. As a result, patients typically begin renal replacement therapy with dialysis, and those deemed medically suitable are placed on a wait list for transplant. Once on the wait list, a patient may later be removed if his or her health condition declines to the point where transplant surgery is considered futile. In the United States, deceased-donor kidneys are allocated on a first-come first-served basis. Patients on the wait list move toward the top of the list when patients above them on the list die, receive a transplant, or are removed.

The continuing shortage of deceased-donor kidneys has prompted the increased frequency of transplantation with expanded criteria donor (ECD) kidneys (Port et al., 2002). These are kidneys obtained from deceased donors who were age ≥ 60; or of age 50–59 with one or more of the following characteristics: death due to stroke, history of hypertension, serum creatinine $\geq 1.5\mu$mol/L. As established by Port et al. (2002), ECD kidneys are associated with somewhat poorer outcomes (e.g., a 70% relative increase in the rate of graft failure). On the other hand, ECD kidneys are more available than non-ECD kidneys since patients are generally more likely to decline ECD offers. The question we address is whether a patient should accept an ECD transplant (considered as an experimental treatment), or whether he/she should opt for "conventional therapy." The latter would entail refusing ECD transplantation, with the hope

of later receiving a non-ECD transplant; this comes with the risk that they will have to wait so long that the patient dies on the wait list or is removed.

In clinical settings, patients often have to choose between different therapies. For chronic conditions (particularly those associated with high rates of adverse events), a patient may choose the treatment course that offers the longest survival time, although other criteria such as quality of life may also enter the decision. The randomized controlled trial, in which patients are randomly assigned to experimental treatment or control groups, is widely viewed as the gold standard for the evaluation of treatments. However, the randomization of treatments is not possible in many settings due to ethical and/or logistical considerations. In such settings, so-called observational data offer the primary basis of treatment evaluations. The lack of randomization requires that the statistical analyses accurately account for imbalances between treatments with respect to measured patient characteristics, such as age, BMI or co-morbid conditions. These are often referred to as covariates, and of particular concern are covariates that are strongly related to the risk of death. Failure to adjust for such risk factors may artificially make one treatment appear to be better than another when, in fact, treatment-specific differences in outcomes rates were due only to differences in treatment-specific covariate distributions.

In clinical studies concerning adverse events, the observation period typically concludes before some subjects have experienced the event of interest. In such cases, a subject's event time (often termed a "failure" time) is said to be right censored; it is known only to be greater than their follow-up time. Survival analysis methods are well-suited to handling right censored failure times. For example, the Kaplan–Meier estimator (Kaplan and Meier, 1958) gives a simple estimate of the probability of remaining alive at various follow-up times. In the case of observational studies of patient survival, it is often desirable to assume a particular statistical model that expresses the death rate (more formally known as the hazard function) as a function of patient characteristics or covariates. Such models are referred to as regression models, and for close to forty years, the Cox regression model (Cox, 1972) has been the method of choice for regression analysis of censored survival data. This model allows for the estimation of covariate effects and, as such, yields covariate-adjusted comparisons of treatment options. Covariates in a Cox model may be fixed at baseline (i.e., time 0, the start of follow-up), or may vary during follow-up. The description of such methods is available from many sources; see, e.g., Kalbfleisch and Prentice (2002), Klein and Moeschberger (2003), and Lawless (2003).

In this chapter, we describe the application of both traditional and more recently developed methods in survival analysis for the comparison of therapies. Here, the available data are complicated by several issues: treatment is not assigned at time 0, but rather at some point after initial eligibility; although experimental and conventional forms of treatment are available, treatment availability is limited due to a perpetual excess of demand relative to supply;

the treatment received by a patient is not randomized; and patients can be declared ineligible for treatment.

In the next section, we describe the kidney transplant registry data used for our analyses and formulate the problem more precisely. In Section 13.3, we describe a traditional attempt to address the issues described in the preceding paragraph. In Section 13.4, we describe a modification of the methods used in Section 13.3. A more recently developed and generally more satisfactory approach to the problem is described in Section 13.5. Some discussion is provided in Section 13.6, including a comparison of related methods not used in this chapter.

13.2 Study Population and Notation

Data were obtained from the Scientific Registry of Transplant Recipients (SRTR; `www.srtr.org`), a nationwide population-based organ failure registry. Patient-specific data are reported by the transplant centers to the Organ Procurement and Transplantation Network, which oversees solid organ transplantation in the United States.

For analyses presented in this chapter, the study population was comprised of adult (age ≥ 18) patients initially wait-listed for kidney transplantation in the US between January 1, 1998 and December 31, 2006. For each patient, follow-up began at the date of wait-listing and concluded at the earliest of death, receipt of living-donor transplant, loss to follow-up, or the end of the study's observation period: December 31, 2006.

Table 13.1 presents a summary of events for the study population. Patients entered the SRTR database at the time of being placed on the wait list, which serves as the natural time 0. Just under 30% of the study population received a deceased-donor kidney transplant, and less than one-fifth of these received an ECD kidney. More than 10% of patients were removed from the wait list prior to transplant or death. Since the death time is known for all patients

TABLE 13.1: Analysis of SRTR data: Event counts.

Event	Count	Percentage of Wait-Listed
Wait-Listed	170,415	100
Transplanted: ECD	9,423	5.5
Transplanted: non-ECD	49,382	29.0
Died	39,475	23.2

who died before December 31, 2006, it is not censored by removal. There were approximately 40,000 deaths observed either before or after transplant.

As described in Section 13.1, the objective is to compare the experimental treatment (ECD kidney transplantation) with conventional therapy. In the seminal paper by Wolfe et al. (1999), kidney transplantation was shown to be associated with a strong and significant decrease in mortality relative to dialysis among wait-listed patients. Rabbat et al. (2000) replicated this result through a registry linkage study using data from Ontario, Canada. The main contribution of these works was to restrict the study population to patients actually wait-listed for transplantation. Previous studies, e.g., Schaubel et al. (1995), had compared kidney transplant to dialysis; these studies suffer the limitation that some of the dialysis patients in the comparison group were not actually wait-listed and, hence, not really eligible to receive a kidney transplant. The survival benefit of ECD transplantation is not obvious, based on the above-listed studies. In aggregate, such studies reveal that ECD kidneys are significantly more likely to fail, and that kidney transplantation (averaging over ECD and non-ECD transplants) offers reduced mortality.

We now introduce some notation that will allow us to describe the data for specific individuals, and the analyses that we consider in subsequent sections. Let D_i denote death time for patient i, measured in days since wait-listing. Let C_i represent censoring time. Since patients either leave the study by dying or being lost to follow-up, only the minimum of D_i and C_i is observed, and we set $X_i = \min(D_i, C_i)$. The time of transplant (if it occurs) is given by T_i. As mentioned earlier, a patient may be removed from the wait list prior to transplantation and, if this occurs, we denote the removal time by R_i. We let $A_{1i} = 1$ if patient i receives an ECD kidney transplant, and 0 otherwise. Analogously, A_{2i} is a 0/1 indicator for receiving a non-ECD transplant. The covariate vector is represented by $\mathbf{Z}_i$ and, in the present context, is intended to contain information on characteristics on patient i associated with both patient survival and the probability of being transplanted. Note that, for purposes of this chapter, the covariate information is recorded at time $t = 0$ and not updated.

It is useful to think of the data structure in terms of a state diagram, as in Figure 13.1. From this perspective, all patients enter the wait list (WL) state at time $t = 0$. From there, the patient will either transit to one (and only one) of the ECD, non-ECD, or removal (R) states, or the patient may reach the death state (D) without experiencing any of these three events. The death state can be accessed from any other state. We treat transitions into each of the ECD, non-ECD and removal states as non-reversible in that a patient that receives an ECD transplant is considered to be an ECD patient thereafter; the same holds for non-ECD transplantation. Examples of elements of $\mathbf{Z}_i$ include age, gender (represented as an indicator covariate: 0 = male, 1 = female), race, blood group, and diagnosis category.

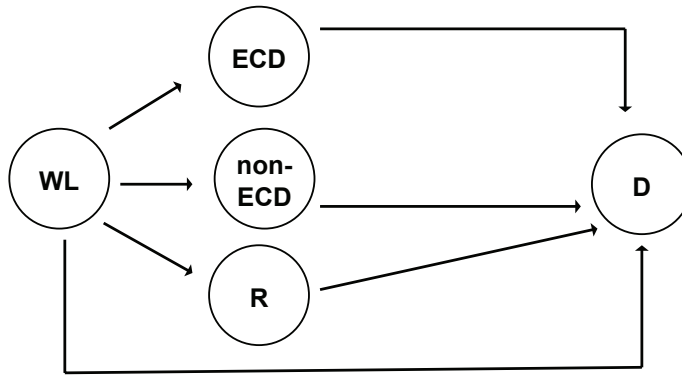

FIGURE 13.1: End-stage renal disease states. WL = wait list; ECD = expanded criteria donor kidney transplant; non-ECD= non-ECD kidney transplant; R = removed from wait list; D = death.

13.3 Analysis Based on Time-Dependent Treatment Indicator

In this section, we describe results from a traditional analysis. As a lead-in, we discuss some ideas related to the analytic approach we take in this and all subsequent sections of the chapter.

For a potentially censored failure time variate such as D_i, interest often lies in the hazard function, $\lambda_i(t)$, which represents the death rate for individual i at time t; this rate is conditional on survival up to time t. Thus,

$$\lambda_i(t) = \Pr(D_i = t | D_i \geq t).$$

Suppose that interest lies chiefly in comparing treated versus untreated subjects with respect to survival, and that treatment is randomly assigned at the beginning of follow-up ($t = 0$). We let $A_i = 1$ for treated subjects and $A_i = 0$ for untreated subjects. In this simple case, the following Cox regression model is often used:

$$\lambda_i(t) = \lambda_0(t) \exp(\beta A_i) = \begin{cases} \lambda_0(t) & \text{if } A_i = 0, \\ \lambda_0(t) \exp(\beta) & \text{if } A_i = 1. \end{cases}$$

The quantity $\lambda_0(t)$ is the baseline hazard function which, in this case, equals the hazard function for an untreated subject ($A_i = 0$). The hazard function for a treated subject is $\lambda_0(t) \exp(\beta)$; thus, the treatment multiplies the hazard function by a factor $\exp(\beta)$, which is commonly referred to as the relative risk due to treatment. If, for example, $\exp(\beta) = .6$, then treated subjects have a

death rate that is 60% that of untreated subjects (i.e., there is a 40% reduction in mortality), which this model assumes to be constant throughout the follow-up period. In this model, no functional form is assumed for $\lambda_0(t)$ and, somewhat surprisingly, β can be estimated without simultaneously estimating $\lambda_0(t)$. Various extensions of this simple model are used in the analyses presented in this chapter.

Three features of the kidney transplantation study require us to extend the model described in the preceding paragraph. First, treatment (kidney transplantation) is not randomized, so a valid comparison of transplanted and wait list patients requires that patient characteristics be accounted for in the analysis. Second, treatment is inherently time-dependent; all patients begin follow-up on the wait list (i.e., untransplanted), with some eventually receiving a transplant. Moreover, there are are two forms of treatment: ECD kidney transplant and non-ECD transplant. Each of these aspects is accommodated by the first analysis we present, which is based on the following model,

$$\lambda_i(t) \quad = \quad \lambda_0(t) \exp\{\beta_1 A_{1i}(t) + \beta_2 A_{2i}(t) + \boldsymbol{\beta}_3^\top \mathbf{Z}_i\}, \tag{13.1}$$

where $A_{1i}(t) = 1$ if patient i receives an ECD transplant before time t (and 0 otherwise); $A_{2i}(t) = 1$ if patient i receives a non-ECD transplant before time t (0 otherwise); and $\mathbf{Z}_i$ is a vector of covariates thought to be associated with both mortality and kidney transplantation.

For example, if the covariates considered were age (in years) and diabetes status (coded as 1 for diabetics and 0 for non-diabetics), then the covariate vector for a 60 year-old diabetic would be equal to $(60, 1)^\top$, with the $^\top$ denoting vector transpose. For this patient, the component in Model (13.1) corresponding to $\mathbf{Z}$ would be $\boldsymbol{\beta}_3^\top \mathbf{Z} = 60\beta_{31} + \beta_{32}$, where β_{31} and β_{32} are the parameters measuring the effect of age and diabetic status on the death rate.

The analysis focuses on estimating the parameters β_1, β_2 and β_3. The inclusion of the covariate vector in Model (13.1) affects the interpretation of β_1 and β_2. Specifically, $\exp(\beta_1)$ is the relative risk (or hazard ratio) for an ECD-transplanted patient versus a patient on the wait list at time t, assuming the patients have identical covariate vectors. Analogously, $\exp(\beta_2)$ is the relative risk for a non-ECD patient versus wait-listed patient, with all covariates equal. The essence of the covariate adjustment is that between-treatment comparisons are interpreted as being among patients with identical covariate vectors. The distribution of patient characteristics may be quite different across treatment groups, but, provided Model (13.1) is correct, this does not introduce bias, since the model has explicitly captured the effect of such factors.

Since we seek to evaluate the merits of ECD kidney transplantation specifically, the parameter of chief interest is β_1. The key feature of Model (13.1) is the use of time-dependent treatment indicators, $A_{1i}(t)$ and $A_{2i}(t)$. At $t = 0$, both will be set to 0, since by definition patients begin follow-up at the time of wait-listing. At most one of $A_{1i}(t)$ and $A_{2i}(t)$ will become 1 during follow-up. The treatment indicators are non-reversible in that, once they jump from 0 to 1, they stay at 1 for the remainder of follow-up.

TABLE 13.2: Analysis of SRTR data: Results from time-dependent models (13.1) and (13.2).

Model	Contrast	Hazard Ratio	(95% CI)
(13.1)	ECD vs. WL	.72	(.69, .75)
	non-ECD vs. WL	.49	(.48, .50)
(13.2)	ECD vs. (WL + non − ECD + R)	.98	(.94, 1.02)

Results based on fitting Model (13.1) to the study data are given in Table 13.2. We find that the estimated hazard ratio (relative risk) is $\exp(\widehat{\beta}_1) = .72$, indicating that the death rate or hazard is reduced by 28% for a patient receiving ECD transplantation, compared to a patient not transplanted (the reference group). The confidence intervals in Table 13.2 represent a margin of error for the relative risk estimates. The estimated relative risk for a non-ECD transplant is $\exp(\widehat{\beta}_2) = .49$, indicating a 51% reduction in death rate for a patient who has received a non-ECD transplant, as compared to the same reference group. Notwithstanding the importance of this result, it does not completely answer our research question, due to the manner in which removals and non-ECD transplants are accommodated; issues that we now describe.

In Model (13.1), a patient who is removed from the wait list stays in the reference group; this is to avoid what is termed dependent censoring. To clarify, although censoring is an inherent feature of survival analysis, it is typically assumed that a subject being censored at time t carries no information about the future death time (other than $D_i > t$, of course) that would have been observed in the absence of censoring. Here, patients tend to get removed from the wait list when their condition has deteriorated to the point at which transplantation is considered futile. Therefore, censoring a patient upon removal would induce dependent censoring.

Although Model (13.1) avoids dependent censoring from removals, the inclusion of removals in the comparison group is also of concern. The motivation for covariate adjustment is to compare an ECD patient to an otherwise equivalent non-transplant patient. However, a patient who is removed is no longer eligible for an ECD (or a non-ECD) kidney transplant. Therefore, the interpretation of $\exp(\beta_1)$ as reflecting "otherwise equal" patients is certainly suspect.

Moreover, Model (13.1) does not handle non-ECD transplantation in a manner consistent with the primary research goal, which was to evaluate the benefit of receiving an ECD transplant. This we consider to be the comparison of the outcomes of an ECD kidney transplantation with the outcomes that would have been observed in the absence of an ECD transplant. A patient who foregoes ECD transplantation (i.e., rules it out as a therapeutic option)

may have the opportunity to later accept a non-ECD transplant. In fact, patients who decline ECD transplantation are not doing so because they think ECD transplantation is no better than remaining perpetually on the wait list, but because they hope later to receive a non-ECD transplant. Since Model (13.1) includes the $A_{2i}(t)$ term, patients are essentially transferred out of the reference group (to which ECD is compared) upon receipt of a non-ECD transplant. One could argue that such patients should be left in the comparison group.

Due to the these limitations associated with the analysis of Model (13.1), our questions regarding the benefit of ECD kidney transplantation remain unanswered at this point. We investigate a simple alternative to this model in the next section.

13.4 Modification to Time-Dependent Analysis

In an attempt to remedy the handling of non-ECD transplants, an alternative time-dependent model is given by

$$\lambda_i(t) \quad = \quad \lambda_0(t) \exp\{\beta_E A_{1i}(t) + \boldsymbol{\beta}_0^\top \mathbf{Z}_i\}, \tag{13.2}$$

which results from deleting the $A_{2i}(t)$ component from Model (13.1). Under this approach, patients appropriately remain in the reference category both after being removed and after receiving a non-ECD transplant.

Results based on Model (13.2) are given in Table 13.2, where it now appears that there is essentially no difference between receiving ECD transplantation and not receiving an ECD transplant, with $\exp(\widehat{\beta}_1) = .98$ and $p = .25$. The reference category is fundamentally different from Model (13.1), in that survival of the ECD group is being compared to survival of all others (wait list, non-ECD, removed) combined.

However, Model (13.2) also has some important drawbacks. It treats removals exactly as in Model (13.1) and so is subject to the same criticism as above. Thus, a patient receiving an ECD transplant at time t is still compared to a reference group that includes patients who were removed prior to time t and so were not eligible for transplantation (ECD or non-ECD) at time t.

This is symptomatic of a more general problem with both Models (13.1) and (13.2). Essentially, the timing of events is not accounted for properly. Consider Figure 13.2, which shows event histories for six hypothetical patients. The patient $i = 1$ receives an ECD transplant (labeled in the figure by E); at the time of that ECD transplant, patient $i = 2$ had already received a non-ECD transplant (labeled by N). Under Model (13.2), patient $i = 2$ is included in the comparison group for patient $i = 1$. But this seems inappropriate since at the time patient $i = 1$ chose to undergo ECD transplantation, patient $i = 2$ had already received a non-ECD transplant and could not have made the

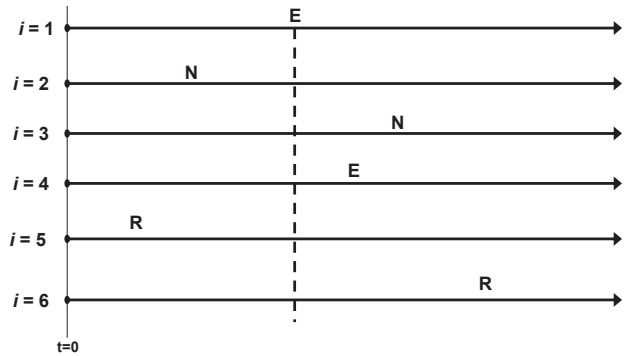

FIGURE 13.2: Event occurrences for $n = 6$ patients ($i = 1, \ldots, 6$): E = expanded criteria donor (ECD) kidney transplant; N = non-ECD kidney transplant; R = removed from wait list.

same treatment decision. However, patient $i = 3$, who received a non-ECD transplant, would be appropriate to include in the comparison group, since that patient was transplant-eligible at the time patient $i = 1$ underwent ECD transplantation. Clearly, analogous comments apply to the two patients in Figure 13.2 that were removed from the wait list ($i = 5$ and $i = 6$, whose time of removal is identified by R). It would not be appropriate to include patient $i = 5$ in the comparison group, since this patient was removed before patient $i = 1$ received an ECD transplant. However, it would be appropriate to include patient $i = 6$ since the removal time follows the time of ECD transplantation for patient $i = 1$.

In considering the limitations associated with model (13.2), and Figure 13.2, it appears that blanket rules regarding non-ECD transplantation and removal are likely to result in an interpretation for β_1 that does not address the research question. However, the preceding paragraph suggests that definition of the appropriate comparison group is straightforward for any individual patient. This is implemented in the next section where we customize the comparison groups for each ECD recipient individually.

13.5 Sequential Stratification

Ideally, ECD kidney transplantation would be evaluated through a randomized controlled trial. In this trial, all patients would be placed on the wait list at the same time, and each time an ECD kidney was procured, a patient

currently on the wait list (i.e., not already removed or transplanted) would be randomly selected to receive the donor kidney; this patient would then represent the "experimental treatment" group. A set of patients who were also eligible to (but did not) receive the ECD kidney would be used as a comparison group, perhaps through matching and/or random selection from those available. These patients represent the "conventional therapy" group, where conventional therapy allows all subsequent events except receipt of the experimental treatment. Note that not being already removed and not having already received a kidney transplant at the time of the index patient's ECD transplant are entry criteria to the comparison group. Note that, in the comparison group, a comparator patient is not censored at subsequent (post-matching) removal or non-ECD transplantation, since such events are standard sequelae of conventional therapy. Comparison patients would be censored when they receive an ECD transplant, since they have then essentially crossed over into the experimental treatment arm and, therefore, no longer contribute follow-up pertinent to the conventional therapy arm. The process of randomizing an incoming ECD kidney and selection of matched conventional therapy patients would be repeated many times, and survival outcomes could then be compared.

Such a trial can never occur, due to logistical and ethical considerations. However, since we take the time of wait list as the origin for all patients, the observed data are quite similar to this; of course the observed data are not randomized, but we could attempt to replace the randomization with careful covariate adjustment. In fact, the largest discrepancies in the setup described in the preceding paragraph and that presented in Sections 13.3–13.4 are related to the method of analysis, as opposed to the actual data structure. The implied analysis features ECD patient-specific comparison groups and, within a given matched set, the handling of removals and non-ECD transplants seems clear cut.

The sequential stratification method was motivated by such considerations, in the context of kidney (Schaubel et al., 2006) and liver transplantation (Schaubel et al., 2009). In essence, this method reorganizes the observed data, then replicates the analysis described above for the ideal experiment. To formalize the ideas, suppose that subject j receives an ECD transplant at time T_j. Patient i (where $i \neq j$) is eligible to be matched to patient j if i is alive and has not been transplanted prior to time T_j (i.e., $\min(T_i, C_i, D_i, R_i) > T_j$) and $s_i = s_j$, where s denotes the matching criteria, which we assume are known at $t = 0$. For example, in our application, age and organ donation service area (the geographic region to which a wait list is intended to apply) served as matching criteria. Thus, we find a set of patients "at risk," denoted by $\mathcal{A}_j$, which comprises patient j as well as those patients who are matched to patient j at time T_j. We refer to $\mathcal{A}_j$ as a matched set of patients, or stratum. The outcomes of patient j are compared to the outcomes of patients $i \neq j$ in this set. This could be done by assuming a Cox type model

$$\lambda_{ji}(t) \quad = \quad \lambda_{0j}(t) \exp(\theta_1 \text{ECD}_{ji} + \boldsymbol{\theta}_0^\top \mathbf{Z}_i), \quad t > T_j \tag{13.3}$$

TABLE 13.3: Analysis of SRTR data: Results from sequential stratification Model (13.3).

Model	Contrast	Hazard Ratio	(95% CI)
(13.3)	ECD vs. Conventional Therapy	.78	(.75, .81)

for each patient i in $\mathcal{A}_j$, where ECD_{ji} takes the value 1 if $i = j$, the ECD recipient, and takes the value 0 otherwise. Note that the covariate vector $\mathbf{Z}_i$ would typically not contain elements used as matching criteria. The sequential stratification method handles both removal and non-ECD transplants appropriately; each is an entry criterion but not a censoring criterion for a given matched set.

A pictorial representation of sequential stratification can be obtained by referring to Figure 13.2. Here, $i = 1$ receives an ECD kidney transplant. Patient $i = 2$ is excluded as a match due to prior (non-ECD) transplantation (i.e., $T_2 < T_1$), while patient $i = 5$ was already removed by the time that $i = 1$ received an ECD ($R_5 < T_1$). Thus $\mathcal{A}_1 = \{1, 3, 4, 6\}$ is the set of patients who were alive and had not been removed or transplanted as of T_1. In terms of post-matching follow-up, patient $i = 3$ is not censored at T_3, since receipt of a non-ECD transplant is a possible sequelae of foregoing ECD transplantation. Analogously, patient $i = 6$ is not censored from the matched set at R_6. Note that patient $i = 4$ would be censored from $\mathcal{A}_1$, at the time of receiving an ECD kidney, since such receipt does not fall under the rubric of conventional therapy, and would be deemed a cross-over to the experimental treatment. In fact, T_4 would generate another stratum, $\mathcal{A}_4 = \{3, 4, 6\}$.

Results from the sequential stratification analysis are presented in Table 13.3. This more complete analysis demonstrates that ECD kidney transplantation is associated with a significant 22% reduction in mortality relative to conventional therapy, with $\exp(\widehat{\theta}_1) = .78$ and $p < .0001$.

13.6 Discussion

In this chapter, we contrast different methods of quantifying the effect of expanded criteria donor (ECD) kidney transplantation with conventional therapy, for which a person stays on the wait list until they are offered a non-ECD kidney, are de-listed, or die. There are different ways of defining a treatment effect, and the one of interest in this chapter is the effect-of-treatment-on-the-treated (Pearl, 2000), for which one aims to compare the outcomes for treated subjects against those which would have been observed (for the treated sub-

jects) in the absence of treatment. The scientific question is not well addressed by standard time-dependent survival analysis methods, which do not accommodate some important elements of the data structure. Sequential stratification, the method we advocate in this setting, involves reorganizing the observed data in a manner more in line with the research question. Aspects of the data, which are difficult to address appropriately through standard methods, can then be dealt with in a transparent and straightforward manner. The method has been used in several analyses targeting the benefit of liver or kidney transplantation; see, e.g., Miles et al. (2007), Schaubel et al. (2008), Lucey et al. (2009), Sharma et al. (2009), Englesbe et al. (2010), and Snoeijs et al. (2010).

A standard method of accommodating a time-dependent treatment in the survival analysis setting is through Cox regression with a time-dependent treatment indicator. Although this approach is suitable for some data structures, its use will generally be problematic when other time-dependent processes are at work. In the ECD setting, both the receipt of a non-ECD transplant and removal from the wait list are time-dependent processes that are important in comparing ECD transplantation to conventional treatment without the ECD option. However, such processes must be handled correctly for the model to be consistent with the particular research question. In other cases, the concomitant time-dependent process may be a covariate observed longitudinally which affects both treatment assignment and the death hazard. Sequential stratification can also be applied to this setting.

The evaluation of time-dependent treatments has received considerable attention in the survival analysis literature in the last 10–15 years. In particular, marginal structural models (MSMs) have gained much popularity in the past decade; see, e.g., Hernán et al. (2000), Robins et al. (2000), and Hernán et al. (2001). The treatment effects estimated through sequential stratification and MSMs are fundamentally different, as described by Schaubel et al. (2009) and Kennedy et al. (2010). The MSM leads to an average treatment effect, where the average is essentially taken across all patients and all possible treatment times. The MSM estimates what is known in the literature as the average causal effect, where this term has come to mean contrasting a scenario in which all patients are treated versus one in which none are treated. As noted in the previous paragraph, however, the treatment effect targeted by sequential stratification is similar to the effect-of-treatment-among-the-treated. Rather than averaging over all patients (treated or not) as would be the case in an MSM, sequential stratification implicitly averages over the patients observed to be treated with ECD and makes no effort toward inference on what would have been the treatment experience for untreated patients. The matching of similar subjects is an attempt to represent what would have been the outcomes of the treated subjects if (contrary to fact) they were untreated.

In the sequential stratification method, a patient selected as a match is censored upon receipt of an ECD transplant; that is, the time to death in the conventional therapy group is censored. Such cross-overs to the experimental

group generally constitute dependent censoring, which necessitates a variant of Inverse Probability of Censoring Weighting; see, e.g., Robins and Rotnitzky (1992). This was explored in detail and implemented by Schaubel et al. (2009) in the context of liver transplantation, where such cross-overs were relatively common. Among end-stage renal disease patients, ECD kidney transplantation is sufficiently rare that bias due to cross-overs to the ECD group among matched patients would be expected to be small.

Since the model assumed in sequential stratification is a Cox-type model, its assumptions are analogous to those in the standard use of a Cox model. Of greatest concern would be the assumption of treatment and adjustment covariate effects that are constant throughout follow-up time. The evaluation of this assumption can be accomplished by extending the model to incorporate interactions between the treatments or covariates and time and appropriate tests carried out; such extensions also can be used to describe more complicated treatment effects than a constant relative risk as discussed in Kalbfleisch and Prentice (2002). When treatment is assigned at baseline, Wei and Schaubel (2008) proposed a method that (like the baseline hazard for the Cox model) assumes no particular functional form for the treatment effect; see also Schaubel and Wei (2011). More recently, Li et al. (2013) proposed matching methods to contrast survival functions when treatment assignment is time-dependent. On a related note, methods developed by Gong (2012) measure the treatment effect in terms of either the difference in survival function, or the area between the survival curves. The work of Gong (2012) requires more modeling assumptions than Li et al. (2013), but allows for more complicated data structures.

It is likely that the survival benefit of ECD transplantation depends on several patient characteristics. For instance, older patients, or patients with diabetes, may have more to gain (relative to conventional therapy) through ECD transplantation, since their prognosis on the wait list is less favorable. In addition, ECD transplantation may be a better option for patients in regions of the country that typically have longer wait lists. Each of these directions would be a useful topic of further investigation.

Acknowledgments

This work was supported in part by National Institutes of Health Grant 5R01 DK–70869. The authors are grateful to the editor and associate editor for their many constructive suggestions. They also thank the Minneapolis Medical Research Foundation for access to the Scientific Registry of Transplant Recipients database. The Scientific Registry of Transplant Recipients is funded by the Health Resources and Services Administration, US Department of Health and

Human Services. The views expressed herein are those of the authors and not necessarily those of the US Government.

About the Authors

Douglas E. Schaubel is a professor of biostatistics at the University of Michigan. His research interests include the evaluation of time-dependent treatments, multivariate survival analysis and collaborative work on liver transplantation and end-stage kidney disease. He works closely with clinicians and surgeons at Arbor Research Collaborative for Health, and the University of Michigan Kidney Epidemiology and Cost Center. He earned a BSc from the University of Waterloo, an MSc from McGill University, and a PhD in biostatistics from the University of North Carolina at Chapel Hill. He is a fellow of the American Statistical Association.

John D. Kalbfleisch is a professor emeritus of biostatistics and statistics at the University of Michigan where he has been on faculty since 2002. Previously, he was at the University of Waterloo, where he also served as chair of the Department of Statistics and Actuarial Science and dean of the Faculty of Mathematics. His research interests include survival and life history analysis, statistical inference and applications of statistics. He is a fellow of the Royal Society of Canada, the American Statistical Association, and the Institute of Mathematical Statistics. He was the recipient of the 1994 SSC Gold Medal for research, served as SSC president in 1999–2000, and was an R. A. Fisher Lecturer for the Committee of Presidents of Statistical Societies (COPSS).

Bibliography

Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34:187–220.

Englesbe, M. J., Schaubel, D. E., Cai, S., Guidinger, M. K., and Merion, R. M. (2010). Portal vein thrombosis and liver transplant survival benefit. *Liver Transplantation*, 16:999–1005.

Gong, Q. (2012). *Semiparametric Methods for Estimating the Effect of a Longitudinal Covariate and Time-Dependent Treatment on Survival Using Observational Data with Dependent Censoring*. Doctoral dissertation, University of Michigan, Ann Arbor, MI.

Hernán, M. A., Brumback, B., and Robins, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, 11:561–570.

Hernán, M. A., Brumback, B., and Robins, J. M. (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *Journal of the American Statistical Association*, 96:440–448.

Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*, Second Edition. Wiley, Hoboken, NJ.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 282:457–481.

Kennedy, E. H., Taylor, J. M. G., Schaubel, D. E., and Williams, S. (2010). The effect of salvage therapy on survival in a longitudinal study with treatment by indication. *Statistics in Medicine*, 29:2569–2580.

Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*, Second Edition. Springer, New York.

Lawless, J. F. (2003). *Statistical Models Methods for Lifetime Data*, Second Edition. Wiley, Hoboken, NL.

Li, Y., Schaubel, D. E., and He, K. (2013). Matching methods for obtaining survival functions to estimate the effect of a time-dependent treatment. *Statistics in Biosciences*, in press.

Lucey, M. R., Schaubel, D. E., Guidinger, M. K., Tome, S., and Merion, R. M. (2009). Effect of alcoholic liver disease and hepatitis C infection on waiting list and posttransplant mortality and transplant survival benefit. *Hepatology*, 50:400–406.

Miles, C. D., Schaubel, D. E., Jia, X., Ojo, A. O., Port, F. K., and Rao, P. S. (2007). Mortality experience in recipients undergoing repeat transplantation with expanded criteria donor and non-ECD deceased-donor kidneys. *American Journal of Transplantation*, 7:1140–1147.

Port, F. K., Bragg-Gresham, J. L., Metzger, R. A., Dykstra, D. M., Gillespie, B. W., Young, E. W., Delmonico, F. L., Wynn, J. J., Merion, R. M., Wolfe, R. A., and Held, P. J. (2002). Donor characteristics associated with reduced graft survival: An approach to expanding the pool of donor kidneys. *Transplantation*, 74:1281–1286.

Rabbat, C. G., Thorpe, K. E., Russell, J. D., and Churchill, D. N. (2000). Comparison of mortality risk for dialysis patients and cadaveric first renal transplant recipients in Ontario, Canada. *Journal of the American Society of Nephrology*, 11:917–922.

Robins, J. M., Hernán, M. A., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11:550–560.

Robins, J. M. and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In *AIDS Epidemiology: Methodological Issues*, pp. 297–331. Birkhäuser, Boston, MA.

Schaubel, D. E., Desmeules, M., Mao, Y., Jeffery, J., and Fenton, S. (1995). Survival experience among elderly end-stage renal disease patients: A controlled comparison of transplantation and dialysis. *Transplantation*, 60:1389–1394.

Schaubel, D. E., Sima, C. S., Goodrich, N. P., Feng, S., and Merion, R. M. (2008). The survival benefit of deceased donor liver transplantation as a function of candidate disease severity and donor quality. *American Journal of Transplantation*, 8:419–425.

Schaubel, D. E. and Wei, G. (2011). Double inverse weighted estimation of cumulative treatment effects under non-proportional hazards and dependent censoring. *Biometrics*, 67:29–38.

Schaubel, D. E., Wolfe, R. A., and Port, F. K. (2006). A sequential stratification method for estimating the effect of an time-dependent experimental treatment in observational studies. *Biometrics*, 62:910–917.

Schaubel, D. E., Wolfe, R. A., Sima, C. S., and Merion, R. M. (2009). Estimating the effect of a time-dependent treatment in by levels of an internal time-dependent covariate. *Journal of the American Statistical Association*, 104:49–59.

Sharma, P., Schaubel, D. E., Guidinger, M. K., and Merion, R. M. (2009). Effect of pre-transplant serum creatinine on the survival benefit of liver transplantation. *Liver Transplantation*, 15:1808–1813.

Snoeijs, M. G. J., Schaubel, D. E., Hene, R., Hoitsma, A. J., Idu, M. M., Ijzermans, J. N., Ploeg, R. J., Ringers, J., Christiaans, M. H. L., Buurman, W. A., and van Heurn, L. W. E. (2010). Kidneys from donors after cardiac death provide survival benefit. *Journal of the American Society of Nephrology*, 21:1015–1021.

Wei, G. and Schaubel, D. E. (2008). Estimating cumulative treatment effects in the presence of nonproportional hazards. *Biometrics*, 64:724–732.

Wolfe, R. A., Ashby, V. B., Milford, E. L., Ojo, A. O., Ettenger, R. E., Agodoa, L. Y. C., Held, P. J., and Port, F. K. (1999). Comparison of mortality in all patients on dialysis, patients on dialysis awaiting transplantation, and recipients of a first cadaveric transplant. *New England Journal of Medicine*, 341:1725–1730.