Tobit Regression for Modeling Mean Survival Time using Data Subject to Multiple Sources of Censoring

Qi Gong[1] and Douglas E. Schaubel[2]

[1]Gilead Science Inc., Foster City, CA 94404,

[2]Department of Biostatistics, University of Michigan,

Ann Arbor, MI 48109 *

Abstract

Mean survival time is often of inherent interest in medical and epidemiologic studies. In the presence of censoring and when covariate effects are of interest, Cox regression is the strong default, but mostly due to convenience and familiarity. When survival times are uncensored, covariate effects can be estimated as differences in mean survival through linear regression. Tobit regression can validly be carried out through maximum likelihood when the censoring times fixed (i.e., known for each subject, even in cases where the outcome is observed). However, tobit regression is generally inapplicable when the response is subject to random right censoring. We propose tobit regression methods based on weighted maximum likelihood which are applicable to survival times subject to both fixed and random censoring times. Under the proposed approach, known right censoring is handled naturally through the tobit model, with Inverse Probability of Censoring Weighting (IPCW) used to overcome random censoring. Essentially, the re-weighting data are intended to represent those that would have been observed in the absence of random censoring. We develop methods for estimating the tobit regression parameter, then the population mean survival time. A closed form large-sample variance estimator is proposed for the regression parameter estimator, with a semiparametric bootstrap standard error estimator derived for the population mean. The proposed methods are easily implementable using standard software. Finite-sample properties are assessed through simulation. The methods are applied to a large cohort of patients wait-listed for kidney transplantation.

KEYWORDS: Inverse Weighting, Mean Lifetime, Random Censoring, Survival, Tobit Model

*Corresponding to: Qi Gong, Gilead Science Inc., Foster City, CA 94404

1 INTRODUCTION

In biomedical studies, Cox regression[1] is the most popular modeling approach for the analysis of censored data in the presence of covariate adjustment. However, the dominance of the Cox model is mostly due to convenience and the availability of pertinent statistical software. It is possible that the investigator may prefer to describe the results of a survival analysis to a stakeholder in terms of differences in mean lifetime. For example, a transplant nephrologist may prefer to describe risk factors affecting pre-transplant survival in terms of differences in mean survival time, as opposed to ratios of hazard functions. The Cox model is inappropriate in a lot of settings due to violation of the proportional hazards assumption. The default modification of the Cox model to handle time-varying covariate effects (the Cox non-proportional hazards model) generally yields covariate effects with an undesirable interpretation [2]. In the presence of proportionality, each Cox model parameter pertains to an ordering of survival functions, which is a very useful property. However, this simplicity disappears under violations of the proportionality assumption. Note that, if one ignores departures from proportionality by fitting a time-constant covariate effects in the absence of proportional hazards (i.e., with the hope of obtaining an average effect), the resulting parameter estimates represent an average which is difficult to define explicitly and depends on the (nuisance) censoring distribution [3].

Mean survival time is generally not estimable non- or semi-parametrically in the presence of censoring since, in most applications, the survival function will not drop to zero. In particular, the maximum follow-up time may correspond to a survival probability that is well above 0, in which case the curve is not closed and mean lifetime is inestimable. In terms of nonparametric estimation, the Kaplan-Meier [4] estimator will drop to 0 if the subject with the maximum follow-up is observed to die, but this final drop in the curve can be quite unstable and lack face validity. For

example, survival probability could be estimated at 0.4 for $t = 9.9$, but 0 for $t = 10$, which is not believable and which would be subject to great imprecision.

Extrapolation can be used to complete the open survival curve to obtain the full mean lifetime estimate. Specifically, the area under the survival curve, estimated either by Kaplan-Meier estimator or Cox model, attached with an extrapolated tail serves as the mean lifetime estimate. Various methods are available for this purpose. Beyond the last observation, Efron [5] used 0, while Gill [6] used the probability estimate of the last event. Brown et al. [7] and Gelber et al. [8] used an exponential distribution to fit the survival curve while Moeschberger and Klein [9] used a Weibull distribution. Gong and Fang [10] derived the hybrid estimate, along with corresponding asymptotic results, for any parametric function fit to the curve. However, the procedures developed in [10] involve the explicit combination of two separate survival analysis techniques to estimate the survival probability of interest, which may be inconvenient and difficult to justify or trust. Moreover, in its existing implementation, the method cannot be used to estimate covariate effects.

An alternative is to estimate restricted mean lifetime [11]; e.g., given by the area under the survival curve up to a pre-specified time, τ. Many authors have employed Cox regression to estimate treatment effects expressed as differences in restricted mean lifetime. For example, Zucker [12] and Chen and Tsiatis [13] proposed a method which averages the fitted values from the Cox model. Zhang and Schaubel [14], based on Chen and Tsiatis [13] paper, extended the method to take into account dependent censoring. Later Zhang and Schaubel [15] further incorporated double robust estimation into their approach. Potential limitations of using restricted mean lifetime include the subjective choice of τ, as well as interpretability. Stepping back for a moment, survival times are inherently continuous, as assumed by standard nonparametric Kaplan-Meier [4] and semiparametric Cox [1] methods. Their continuous

nature makes modeling the mean function especially appealing.

At this juncture, it is useful to make the distinction between two types of right censoring: fixed (which we denote by L_i for subject i) and random (denoted by C_i). Fixed censoring times are known for a subject at the time follow-up begins. For example, in a clinical trial in which each subject is prospectively followed for 5 years, $L_i = L$=5 years for all i. As another example, consider an observational study (e.g., using registry data) spanning a 5 year period, with staggered entry and follow-up for all subjects ending at the end of the 5-year period. In this case, L_i would not be equal for all i, but would be known for each subject at the date of study entry. For instance, $L_i = 3$ years for a patient who enters the study 2 years into the observation period. In contrast to fixed censoring, a random censoring time is not known at the time of study entry. Examples include a patient's voluntary withdrawal from the study, or random loss to follow-up. While it is commonplace in survival analysis to simply consider censoring time, $\min\{L_i, C_i\}$, the distinction between fixed and random censoring is important for the methods proposed in this report.

When survival times are not subject to censoring, the sample mean is the best way to summarize average survival. To adjust for covariates in this setting, survival times could be linearly regressed on the covariates. In linear regression, lifetime as a dependent variable is assumed to follow a normal distribution. In the vast majority of practical settings, survival times are subject to (at least fixed, and perhaps random) right censoring. The Tobit model [16] was proposed as a parametric method to describe the relationship between covariates and a non-negative censored dependent variable. Inference proceeds via Maximum Likelihood Estimation, with survival times assumed to follow a Normal distribution. However, the Tobit model can generally only deal with fixed right censoring, but not random censoring. This is an important limitation on the context of survival analysis, since random censoring occurs

frequently in clinical trials and observational studies.

In this report, we propose Tobit regression methods which can accommodate both fixed and random right censoring. In a given study, if all censoring times were fixed, existing Tobit regression methods could be applied. We essentially propose a weighted version of a complete-case analysis, fitting the Tobit regression model to only subjects either observed to die or observed to live until their fixed end-of-follow-up time. We weight such subjects using a variant of Inverse Probability of Censoring Weighted (IPCW), originally proposed by Robins and Rotnitzky [17] and Robins and Finkelstein [18]. Basically, IPCW weights the censoring and events by the inverse of the probability of not being censored, such that the weighted version of the likelihood (and hence resulting estimators) reflect that which would have been observed in the absence of censoring. The probability of being censored is modeled via Cox regression, with random censoring as the event. The assumed death model is the same as in the uncensored setting, such that the estimated regression parameters have a straightforward interpretation. An estimator of marginal mean survival time is also proposed as the empirical average of fitted values. The proposed methods are a useful option for practitioners when mean survival time is of inherent interest, or when the proportional hazards assumption is violated.

The remainder of the article is organized as follows. The data set-up and proposed methods are described in Section 2. A simulation study is presented in Section 3 to evaluate the operating characteristics of the proposed procedures. The proposed methods are applied to a cohort of end-stage renal disease patients in Section 4. Finally, some discussion is provided in Section 5.

2 PROPOSED METHODS

In this section, we first formalize the data structure described in Section 1. We then describe how estimation would unfold for (i) uncensored data (ii) survival times subject to known right censoring (iii) death times subject to both fixed and random censoring.

2.1 Set-up and Model

We begin by setting up the required notation. Let i denote subject ($i = 1, \ldots, n$). We let D_i denote the death time, and consider the setting wherein D_i is subject to two types of right censoring. Specifically, let L_i represent the known right censoring time, which occurs administratively at the end of the observation period. The random right censoring time is given by C_i. Note that L_i is always known, even if $D_i < L_i$ or $C_i < L_i$. Covariate vectors pertinent to D_i and C_i are represented by Z_i and Z_i^C, respectively, each of which is assumed not to change after baseline (time $t = 0$). One observes a vector of data for each subject, $(X_i, L_i, Z_i, Z_i^C, \Delta_i^D, \Delta_i^C, \Delta_i^L)$, where follow-up time is defined as $X_i = \min(D_i, C_i, L_i)$, $\Delta_i^D = I(X_i = D_i)$, $\Delta_i^C = I(X_i = C_i)$, $\Delta_i^L = I(X_i = L_i)$ with $I(\cdot)$ being the standard 0-1 indicator function.

Conditional on Z_i, D_i is assumed to follow a Normal distribution; i.e., $D_i|Z_i \sim N(\mu_i, \sigma^2)$, where σ is the common standard deviation. Further, we assume that $\mu_i = E[D_i|Z_i]$ is a linear combination of the covariates, such that $\mu_i = \beta_0' Z_i$, with β_0 being the unknown parameter vector of interest. We let $\mu = E[D_i]$ be the marginal expectation of the survival time at the population level; i.e., averaged over the distribution of the covariate vector, such that $\mu = E[E[D_i|Z_i]]$.

We assume that D_i is conditionally independent of C_i given the covariate Z_i. This is a somewhat stronger form of independence than would be assumed, for example, in a Cox model for $D_i|Z_i$. The standard definition of independent censoring used in

Cox regression, $\lambda^D(t|Z) = \lambda^D(t|Z, C > t)$ is very similar to (but theoretically does not imply) independence between D and C given Z.

2.2 Estimation when D is subject to known censoring L

In the absence of censoring, under the assumed model, $E[D_i] = \beta_0' Z_i$ and $\text{var}(D_i) = \sigma^2$, the response variate is naturally suited to linear regression, $D_i = \beta_0' Z_i + \varepsilon_i$ where $\varepsilon_i \sim N(0, \sigma^2)$. However, if D_i is potentially censored by known censoring time L_i, we could employ a Tobit model to obtain an unbiased estimator of β_0.

If none of the subjects are censored, the Tobit model has the likelihood function

$$\prod_{i=1}^{n} \left[\frac{1}{\sigma}\phi\left(\frac{X_i - \beta' Z_i}{\sigma}\right)\right]^{\Delta_i^D} \left[\Phi\left(\frac{\beta' Z_i - L_i}{\sigma}\right)\right]^{\Delta_i^L}, \tag{1}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the cumulative and probability density functions of standard normal distribution. The log likelihood would then be given by

$$\sum_{i=1}^{n} \Delta_i^D \log\left[\frac{1}{\sigma}\phi\left(\frac{X_i - \beta' Z_i}{\sigma}\right)\right] + \Delta_i^L \log\left[\Phi\left(\frac{\beta' Z_i - L_i}{\sigma}\right)\right]. \tag{2}$$

2.3 Estimation when D is potentially censored by L and C

It is frequently the case in practice that D_i can be randomly right censored prior to end-of-study censoring L_i. In this setting, we propose carrying out an inverse-weighted complete-case analysis of $\{i : \Delta_i^D + \Delta_i^L = 1\}$; i.e., a complete-case analysis of the data that would suffice in the absence of C, inverse weighted such that the weighted data set represents the data that would be observed in the absence of C. Analogous inverse-weighted complete case approaches have been developed for other survival analysis settings; e.g., [19]. The weight function is derived through IPCW [17, 18] and given by, $W_i = (\Delta_i^D + \Delta_i^L) \exp\left\{\Lambda_i^C(X_i)\right\}$, where $\Lambda_i^C(t) = \int_0^t \lambda_i^C(u)du$ and $\lambda_i^C(t)$ is the hazard function for C_i conditional on Z_i^C. Note that the quantity, $\exp\left\{-\Lambda_i^C(t)\right\}$, represents the probability of remaining uncensored (i.e., not randomly censored) as of time t.

A weighted Tobit model log likelihood function can be written as

$$\ell(\beta) = \sum_{i=1}^{n} \Delta_i^D W_i \log\left[\frac{1}{\sigma}\phi\left(\frac{X_i - \beta' Z_i}{\sigma}\right)\right] + \Delta_i^L W_i \log\left[\Phi\left(\frac{\beta' Z_i - L_i}{\sigma}\right)\right].$$

The regression coefficient β_0 is then estimated by $\widehat{\beta}$, the root of the score function

$$\begin{aligned} U(\beta) = & \sum_{i=1}^{n} -\Delta_i^D W_i Z_i \frac{1}{\sigma}\phi'\left(\frac{X_i - \beta' Z_i}{\sigma}\right)\phi\left(\frac{X_i - \beta' Z_i}{\sigma}\right)^{-1} \\ & + \Delta_i^L W_i Z_i \frac{1}{\sigma}\phi\left(\frac{\beta' Z_i - L_i}{\sigma}\right)\Phi\left(\frac{\beta' Z_i - L_i}{\sigma}\right)^{-1}. \end{aligned}$$

The variance of $\widehat{\beta}$ can be estimated by $\widehat{\text{var}}(\widehat{\beta}) = n[-\partial U(\beta)/\partial\beta]^{-1}$.

Typically, the weights $\{W_1, \ldots, W_n\}$ will not be known and will need to be estimated. Consistent with the vast majority of IPCW-based methods, we assume that C_i follows a proportional hazards model, $\lambda_i^C(t) = \lambda_0^C(t)\exp\{\beta_C' Z_i^C\}$, where $\lambda_0^C(t)$ is an unspecified baseline hazard function. Parameter estimation for this model is carried out using standard partial likelihood, based on data $\{X_i, \Delta_i^C, Z_i^C\}$ for $i = 1, \ldots, n$. The regression coefficient β_C is estimated by $\widehat{\beta}_C$, the root of the score function

$$U_C(\beta) = \sum_{i=1}^{n} \int_0^\tau \{Z_i^C - \overline{Z}_C(t;\beta)\} dN_i^C(t),$$

where $N_i^C(t) = \Delta_i^C I(C_i \leq t)$ and $\overline{Z}_C(t;\beta) = R_C^{(1)}(t;\beta)/R_C^{(0)}(t;\beta)$, $R_C^{(p)}(t;\beta) = n^{-1}\sum_{i=1}^n Y_i(t) Z_i^{C\otimes p}\exp\{\beta' Z_i^C\}$ for $p = 0, 1, 2$, and with the following definitions: $Z_i^{\otimes 0} = 1$, $Z_i^{\otimes 1} = Z_i$ and $Z_i^{\otimes 2} = Z_i Z_i'$. The Breslow estimator of $\Lambda_0^C(t)$ is given by $\widehat{\Lambda}_0^C(t) = n^{-1}\sum_{i=1}^n \int_0^t R_C^{(0)}(u;\widehat{\beta}^C)^{-1} dN_i^C(u)$.

Note that the assumption required for C_i to be independently censored by D_i (required for consistent estimation of the IPCW parameters through standard partial likelihood) is implied by the afore-listed conditional independence assumption regarding C_i and D_i.

Being estimated quantities, the weights are actually random variables, such that $\widehat{\text{var}}(\widehat{\beta})$ cannot be directly obtained from available statistical software without a considerable amount of explicit programming. However, for simplicity here we consider

the estimated weights as fixed. Note that many existing IPCW-based methods have also treated the $\widehat{W}_i(t)$ as fixed for the purposes of variance computation. We evaluate the accuracy of the proposed variance estimator in Section 3.

2.4 Estimating the marginal mean survival time

We propose to estimate $E[D]$ through $\widehat{\mu} = n^{-1}\sum_{i=1}^{n}\widehat{\mu}_i$, where $\widehat{\mu}_i = \widehat{\beta}'Z_i$. With respect to variance estimation, one option would be to use the Delta Method, but this would not account for the variability in the Z distribution. We therefore propose to estimate $\text{var}(\widehat{\mu})$ through a semi-parametric bootstrap method. Specifically, we sample B replicates of $\{Z_1, \ldots, Z_n\}$, where B is a large number (e.g., we chose $B = 1000$). We then draw B replicates of $\widehat{\beta}$ from its asymptotic distribution, $N(\widehat{\beta}, \widehat{var}(\widehat{\beta}))$. This provides B replicates, $\widehat{\mu}_b = n^{-1}\sum_{i=1}^{n}\widehat{\beta}_b'Z_{ib}$, for $b = 1, \ldots, B$. The bootstrap variance is then provided by the empirical variance of the B replicates.

3 SIMULATION

We set $Z_i = (1, Z_{1i}, Z_{2i})'$ and, correspondingly, let $\beta = (\beta_0, \beta_1, \beta_2)'$. We let Z_{1i} be a Bernoulli(0.5) covariate, while Z_{2i} is distributed as a standard Normal variate truncated by $[-2, 2]$. We also assume $Z_i^C = (Z_{1i}, Z_{2i}, Z_{3i})'$ where Z_{3i} is a count variable which evenly distributed on integers $(0, 1, \ldots, 9)$. The baseline hazard function for C_i is set to $\lambda_0^C(t) = 1/c$, with parameter vector $\beta^C = (\beta_1^C, \beta_2^C, \beta_3^C)'$.

We set the parameter of interest to $(\beta_0, \beta_1, \beta_2)' = (3, 0.25, 0.25)'$, and let $\sigma = 1$. We let $(\beta_1^C, \beta_2^C, \beta_3^C)' = (0.15, 0.15, 0.15)'$. We then examine 6 different scenarios, differentiated by $c = 100, 50, 30$ and $L_i = L$ set to 4 and 3.5, with each scenario having different percentages of both random censoring and known (end-of-follow-up) censoring. In scenarios 1-3, the percentage of subjects censored by L is approximately 16%, while the percentage censored by C equalled approximately 7%, 13% and 20%

respectively. Similarly, in scenarios 4-6, the percentage censored by L remained at 30%, while the percentage censored by C increased from 6%, to 12%, and to 20%. Even heavier censoring is generated in scenarios 7 and 8, in which only 40% and 20% of subjects are observed to die, respectively. Further detail is provided in refer Table 1. Sample sizes were set to $n = 100$ and $n = 250$. We generated 1000 replicates for each scenario. The weighted Tobit model was implemented by using SAS® 9.3 PROC QLIM.

Simulation results are presented in Table 1. For each parameter, bias was generally small, while the Empirical Standard Error (ESE) and Asymptotic Standard Error (ASE) were generally very close. Correspondingly, coverage probability (CP) was approximately at the nominal 95% level, even in the cases where only 53% of subjects were observed to die. Note that 47% is a fairly high percentage of censoring, relative to the standard implied based on published reports. The CP is below the nominal level for scenarios 7 and 8, which feature much higher censoring rates. It is possible that the large-sample results are more accurate in much larger sample sizes, particularly for settings where a low percentage of the sample is observed to die. With respect to bias and CP, results were generally better for $n = 250$ than $n = 100$, even for the high censoring scenarios.

Table 1: Simulation results of the mean of survival time and coefficient.

n	#	D%[1]	L%[2]	C%[3]	μ (True value=3.125)				β_0 (True value=3.00)				β_1 (True value=0.25)				β_2 (True value=0.25)			
					Bias	ESE	ASE	CP	Bias	ESE	ASE	CP	Bias	ESE	ASE	CP	Bias	ESE	ASE	CP
100	1	75	18	7	0.001	0.103	0.107	97	0.003	0.142	0.144	95	-0.004	0.205	0.203	95	-0.009	0.109	0.106	93
	2	71	16	13	0.007	0.113	0.108	94	0.012	0.158	0.143	92	-0.009	0.222	0.204	93	0.001	0.118	0.107	92
	3	66	14	20	<0.001	0.116	0.108	94	-0.002	0.164	0.144	91	0.004	0.229	0.204	90	-0.002	0.123	0.107	90
	4	61	33	6	0.008	0.116	0.116	94	0.009	0.157	0.151	94	-0.003	0.218	0.212	95	0.007	0.121	0.111	92
	5	58	30	12	-0.003	0.123	0.115	93	-0.005	0.165	0.150	93	0.003	0.225	0.211	93	-0.004	0.123	0.111	92
	6	53	27	20	-0.026	0.105	0.122	96	0.003	0.163	0.167	96	-0.001	0.232	0.236	95	0.007	0.132	0.125	94
	7	40	35	25	0.004	0.154	0.131	91	0.011	0.194	0.164	90	-0.011	0.269	0.224	90	<0.001	0.136	0.117	91
	8	20	45	35	0.027	0.265	0.198	88	0.015	0.283	0.219	88	0.024	0.340	0.262	87	0.016	0.177	0.137	88
250	1	75	18	7	0.003	0.071	0.068	93	0.003	0.093	0.091	93	-0.002	0.131	0.129	95	0.002	0.071	0.067	93
	2	71	16	13	0.002	0.069	0.068	94	0.004	0.097	0.091	93	-0.004	0.139	0.129	94	-0.001	0.074	0.068	93
	3	66	14	20	0.001	0.074	0.068	93	0.007	0.098	0.092	93	-0.011	0.144	0.130	93	0.001	0.074	0.068	92
	4	61	33	6	0.003	0.074	0.072	95	0.002	0.097	0.095	96	0.002	0.136	0.134	94	-0.001	0.073	0.070	93
	5	58	30	12	0.005	0.075	0.072	94	0.004	0.101	0.093	94	0.001	0.142	0.134	94	-0.001	0.075	0.070	93
	6	53	27	20	-0.003	0.080	0.072	93	-0.007	0.105	0.094	92	0.008	0.147	0.134	92	0.002	0.076	0.070	93
	7	40	35	25	<0.001	0.093	0.082	92	-0.002	0.114	0.103	92	0.004	0.154	0.141	93	0.001	0.082	0.075	92
	8	20	45	35	0.007	0.150	0.122	89	0.005	0.163	0.136	90	0.001	0.195	0.163	90	0.006	0.109	0.086	89

[1]Percentage of subject who died before the end of follow-up; [2]Percentage of subject who got truncated at the end of follow-up; [3]Percentage of subject who got censored before the end of follow-up; ESE: empirical standard error; ASE: asymptotic standard error CP: 95% coverage probability.

4 APPLICATION

We applied the proposed methods to quantify the effects of patient characteristics on mean pre-transplant survival time among end-stage renal disease (ESRD) patients wait listed for deceased-donor kidney transplantation. Data were obtained from the Scientific Registry of Transplant Recipients (SRTR), a national population based organ transplant registry. The SRTR data system includes data on all donors, wait-listed candidates, and transplant recipients in the U.S., as submitted by the members of the Organ Procurement and Transplantation Network (OPTN), and has been described elsewhere. The Health Resources and Services Administration (HRSA), U.S. Department of Health and Human Services provides oversight to the activities of the OPTN and SRTR contractors.

The study population consisted of patients wait listed for deceased-donor kidney transplantation in the United States between 01/01/2009 and 12/31/2014. Only adult patients (age$\geq$18 at listing) not previously transplanted were included in the study cohort. For each patient, follow-up began on the date of wait listing (WL), and continued until the earliest of death, receipt of a kidney transplant, loss to follow-up or the end of the observation period (12/31/2014). The known censoring time was given by the time interval between the date of wait listing and 12/31/2014. Random right censoring occurred at the earliest of kidney transplantation and loss to follow-up. Note that patients were censored upon receipt of a kidney transplant since survival in the absence of kidney transplantation (i.e., pre-transplant survival) was of interest. Under the kidney allocation rules in effect during our observation period, such censoring can validly be considered to be independent of pre-transplant death time, conditional on the covariate vector. Note that the treatment of transplantation as independent censoring is not appropriate for some solid organ type; e.g., liver (see [20; 21].

Covariates, each determined at the time of wait listing, included the following: age, gender, race, blood type, years on dialysis (prior to wait listing), serum albumin (g/dL), height (cm), weight (kg), and the following comorbid conditions: diabetes, hypertension, cerebrovascular disease (CVD), chronic obstructive pulmonary disease (COPD) and peripheral vascular disease (PVD).

Cox regression was used to estimate the IPCW weight, W_i, with the covariate vector being equal to that included in the death model. As is common when using IPCW, some very large values occurred for the weight function, owing to very small probability of remaining uncensored. Unduly large estimated weights can increase the variance of $\widehat{\beta}$ by inappropriately assigning excessive weight to a subject. One frequently applied solution is to cap the weight. Since we found that 99% of subjects had $\widehat{W_i} < 2.25$, we capped the weight function at 2.25.

The study cohort consisted of $n = 56,518$ patients, of which 9,676 were observed to die (17.1%); 11,520 were were randomly censored (20.4%); and 35,322 were censored by end-of-study (62.5%). Elements of the estimated regression parameter are listed in Table 2, with the time units being months post-wait listing. Mean pre-transplant survival time decreases by approximately 4.5 months for every 10 year increase in age, all other covariates being equal ($p < 0.0001$). Every racial minority subgroup outlives Caucasian patients ($p < 0.0001$) by $\geq$7 months, with the longest mean survival estimated to be for Asian patients. Mean survival time decreases by ≈ 2 months per year increase in the time interval between dialysis-initiation and wait listing. This time interval is perhaps a surrogate for unmeasured health status factors delaying the patient being deemed medically suitable for transplantation. Note that hypertension status is based on the receipt of a prescription for medication for the condition. From this perspective, the apparently protective effect of hypertension may derive from a reduced mortality risk attributable to treatment. The estimated

intercept (≈ 4 years) applies to a patient with covariate vector 0.Since we subtracted, from each numerical covariate, a round number approximately equal to its respective median, the intercept corresponds to a patient with the following characteristics: age 50, male, Caucasian, Type O, 3 years prior dialysis, serum albumin of 3 g/dL, 170 cm tall, weighing 80 kg, and free of all five listed comorbidities. The reference level of each categorical covariate was the mode.

The marginal mean survival time is estimated at $\widehat{\mu} = 41.6$ months, with estimate standard error 0.37 months.

We evaluated the adequacy of the model, in terms of risk discrimination, through the C index, also known as the Index of Concordance [22]. This quantity is essentially computed as the fraction of patient pairs for which the ordering of the death times is concordant with the ordering of the fitted mean survival times. We computed $\mathbb{C} = 0.661$, indicating that that the proposed model correctly orders patient-pairs approximately twice as often as it does so incorrectly. This is a respectable result, particularly given that $\mathbb{C} = 0.660$ was computed for a Cox regression model based on the same data and covariate vector. Results were very similar when we computed the C index based on 10-fold cross-validation, in which case $\mathbb{C} = 0.658$ for the tobit model and $\mathbb{C} = 0.659$ for the Cox model. The consistency between the 'internal' and cross-validation versions of $\mathbb{C}$ is likely due to the sample size (and number of deaths) being so large. Similar findings were reported by Schaubel et al. [23] in the context of end-stage liver disease.

5 CONCLUSION

The proposed method can be used to evaluate covariate effects on mean survival time, while accounting for both end-of-study and random censoring. The proposed approach models mean survival time directly by using it explicitly as response vari-

Table 2: Analysis of pre-kidney transplant survival: Estimated regression parameter (time unit: months)

Covariate	$\widehat{\beta}$	$\widehat{SE}$	p
Intercept	47.27	1.14	<0.0001
Age (per 10 years)	-4.47	0.24	<0.0001
Gender:			
Female	2.91	0.64	<0.0001
Male	0	..	..
Race:			
Asian	10.65	1.11	<0.0001
African American	7.80	0.57	<0.0001
Hispanic	7.28	0.70	<0.0001
Other	9.57	1.87	<0.0001
Caucasian	0	..	..
Blood Type:			
A	-1.65	0.52	0.001
B	-0.87	0.69	0.203
AB	-2.98	1.23	0.015
O	0	..	..
Years on dialysis (prior to WL[a])	-1.95	0.10	<0.0001
Albumin (per g/dL)	12.14	0.41	<0.0001
Height (per 10 cm)	-1.46	0.34	<0.0001
Weight (per 5 kg)	0.77	0.07	<0.0001
Comorbid conditions:			
Diabetes	-8.27	0.50	<0.0001
Hypertension	3.06	0.51	<0.0001
CVD[b]	-1.52	1.17	0.19
COPD[c]	-5.49	1.49	0.0002
PVD[d]	-5.74	0.86	<0.0001

Notes: (a) WL = wait listing (b) CVD = cerebrovascular disease (c) COPD = chronic obstructive pulmonary disease (d) PVD = peripheral vascular disease

ate in a Tobit regression. This would often be much more convenient than what is currently a frequently applied alternative: modeling the hazard (e.g., through Cox regression), combining the regression parameter and cumulative baseline hazard estimates, then transforming and finally integrating the subject-specific survival curve. Moreover, in cases where mean survival time is of interest, the regression parameter is directly relevant, while hazard regression procedures yield covariate effects which apply directly to the hazard function, but which apply indirectly to any non-linear function of the hazard function, such as the survival function or its integration.

The implementation of the proposed method is computationally convenient, since the Cox model and Tobit model are widely available in standard statistical software packages, such that coding effort is reduced.

Although the methodology in this report was motivated by data originating from the health sciences and medicine, the methods could be applied to other fields such as economics, sociology and engineering. For example, the average time of using a credit card before closing, and how long an airplane engine can function before failure.

A limitation of the proposed method is the reliance on the Normal distribution for the failure times. It is possible that results are somewhat robust to non-normality. However, it is well known that survival data often exhibit right-skewness.

The most popular alternative to modeling mean survival time directly would be to model the restricted mean survival time. Methods for modeling the restricted mean directly have gained increased attention in the biostatistical literature recently [24]. The main drawback to such methods is the need to select a truncation time. In particular, different investigators could prefer different restriction times, with the inference and, hence, conclusions drawn possibly depending on which time is chosen. That said, extrapolation is inherent when the mean is modeled based on censored data; concern along these lines is largely mitigated by modeling the restricted mean.

In the end, it can be safely stated that mean and restricted mean are both interesting and useful bases for the analysis of censored survival times.

Future work includes consideration of dependent censoring and non-constant variance, both of which frequently occur in clinical and epidemiologic data, often as manifestations of unmeasured covariates.

6 ACKNOWLEDGMENTS

This work was supported in part by National Institutes of Health Grant R01 DK070869. The data reported here have been supplied by the Minneapolis Medical Research Foundation (MMRF) as the contractor for the Scientific Registry of Transplant Recipients (SRTR). The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy of or interpretation by the SRTR or the U.S. Government. The authors also wish to thank the Associate Editor and two Referees, whose thoughtful comments improved the manuscript considerably.

References

[1] Cox DR. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society Series B.* 1972;34:187-220.

[2] Wei G, Schaubel DE. Estimating cumulative treatment effects in the presence of nonproportional hazards. *Biometrics.* 2008;64:724-732.

[3] Struthers CA, Kalbfleisch JD. Misspecified proportional hazards models. *Biometrika.* 1986;73:363-369.

[4] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association.* 1958;53:457-481.

[5] Efron B. The two sample problem with censored data. *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability.* 1967;4:831-852.

[6] Gill RD. Censoring and stochastic integrals. *Statistica Neerlandica.* 1980;34:124.

[7] Brown BW, Hollander M, Korwar RM. *Nonparametric tests of independence for censored data, with applications to heart transplant studies. Reliability and Biometry: Statistical Analysis of Lifelength.* Philadelphia: Society for Industrial and Applied Mathematics; 1974.

[8] Gelber RD, Goldhirsch A, Cole BF. Parametric extroplation for survival estimates with applications to quality of life evaluation of treatments. *Controlled Clinical Trials.* 1993;14:485-499.

[9] Moeschberger ML, Klein JP. A comparison of several methods of estimating the survival function when there is extreme right censoring. *Biometrics.* 1985;41:253-259.

[10] Gong Q, Fang L. Asymptotic properties of mean survival estimate based on the Kaplan-Meier curve with an extrapolated tail. *Pharmaceutical Statistics.* 2012;11:135-140.

[11] Irwin JO. The standard error of an estimate of expectational life. *Journal of Hygiene.* 1949;47:188-189.

[12] Zucker D. Restricted mean life with covariates: Modification and extension of a useful survival analysis method. *Journal of American Statistical Association.* 1998;93:702-709.

[13] Chen P, Tsiatis AA. Causal inference on the difference of the restricted mean lifetime between two groups. *Biometrics.* 2001;57:1030-1038.

[14] Zhang M, Schaubel DE. Estimating differences in restricted mean lifetime using ovservational data subject to dependent censoring. *Biometrics.* 2011;67:740-749.

[15] Zhang M, Schaubel DE. Double-robust semiparametric estimator for differences in restricted mean lifetimes in observational studies. *Biometrics.* 2012;68:999-1009.

[16] Tobit J. Estimation of relationships for limited dependent variables. *Econometrica.* 1958;26:24-36.

[17] Robins JM, Rotnitzky A. Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell N., Dietz K., Farewell B., eds. *AIDS Epidemiology - Methodological Issues*, Boston: Birkhäuser 1992 (pp. 297-331).

[18] Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) Log-rank tests. *Biometrics.* 2000;56:779-788.

[19] Schaubel DE, Zhang H, Kalbfleisch JD, Shu X. Semiparametric methods for survival analysis of case-control data subject to dependent censoring. *Canadian Journal of Statistics.* 2014;42:365-383.

[20] Gong Q, Schaubel DE. Partly conditional estimation of the effect of a time-dependent factor in the presence of dependent censoring. *Biometrics.* 2013;69:338-347.

[21] Gong Q, Schaubel DE. Estimating the average treatment effect on survival based on observational data and using partly conditional modeling. *Biometrics.* 2017;73:134-144.

[22] Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine.* 1996;15(4):361-387.

[23] Schaubel DE, Guidinger MK, Biggins SW, et al. Survival benefit-based deceased-donor liver allocation. *American Journal of Transplantation.* 2009;9(4 Pt 2):970-981.

[24] Tian L, Zhao L, Wei LJ. Predicting the restricted mean event time with the subject's baseline covariates in survival analysis. *Biostatistics.* 2014;15:222-233.