Lifetime Data Anal (2015) 21:1–19
DOI 10.1007/s10985-014-9306-8

Ornstein–Uhlenbeck threshold regression for time-to-event data with and without a cure fraction

Roger Erich · Michael L. Pennell

Received: 6 March 2013 / Accepted: 23 July 2014 / Published online: 6 August 2014

Abstract In this paper we propose a threshold regression (TR) model for time to event data related to subject health using a latent Ornstein–Uhlenbeck (OU) process that fails once it hits a boundary value for the first time. Baseline covariates are incorporated into the analysis using a log-link function for the initial state of the health process. The model provides clinically meaningful covariate effects and does not require the proportional hazards assumption of the commonly used Cox model. Unlike TR models based on the Wiener process, the OU model allows increments in the health process to depend on previous values and drifts toward a state of equilibrium or homeostasis, which are present in many biological applications. We also extend our model to incorporate a cure rate for applications with improper survival functions, such as time to tumor recurrence in a cancer clinical trial. Our models are applied to overall and relapse-free survival data of melanoma patients undergoing definitive surgery.

Keywords Cancer clinical trial · First hitting time model · Gaussian process · Mixture model · Nonproportional hazards · Survival analysis

Electronic supplementary material The online version of this article (doi:10.1007/s10985-014-9306-8) contains supplementary material, which is available to authorized users.

R. Erich
U.S. Air Force Institute of Technology, Wright-Patterson Air Force Base, OH 45433, USA

M. L. Pennell (✉)
Division of Biostatistics, College of Public Health, The Ohio State University, 246 Cunz Hall, 1841 Neil Avenue, Columbus, OH 43210, USA
e-mail: pennell.28@osu.edu

1 Introduction

In clinical studies of mortality or disease progression, death or disease onset/recurrence can be viewed as the outcome of a series of genetic and physiological events where a subject's health deteriorates until it reaches a low-point at which the adverse event occurs. Take for instance a subject who has been diagnosed with melanoma. The cancer begins in Stage 0 (also known as melanoma in-situ) which is treatable and has a high survival rate. However, if untreated, the cancer will progress through invasive (Stage I/II), regional (Stage III), and then distant (Stage IV) metastasis at which point the chance of survival is low even with treatment (American Cancer Society 2012).

The idea that an event such as death or tumor occurrence can be viewed as the culmination of a series of changes in health is the basis of a relatively new class of survival models known as first hitting time (FHT) models. In these models, subject health is modeled using a latent stochastic process that fails once it reaches a boundary for the first time. See Lee and Whitmore (2006) for a recent review of FHT models. A popular FHT model involves modeling subject health using a Wiener process with drift, and it results in an inverse-Gaussian distribution for the time-to-event (Cox and Miller 1968; Chhikara and Folks 1989). The Wiener process model has been applied to biomedical (Horrocks and Thompson 2004), engineering (Park and Padgett 2005) and social science data (Lancaster 1972). In an approach known as threshold regression (TR), the parameters of the Wiener process (drift, variance, initial state) may be related to covariates using generalized link functions (see, for example, Lee et al. 2000, 2004; Lee and Whitmore 2006; Aalen and Gjessing 2001; Aalen et al. 2008). Threshold regression models have a couple of advantages over the commonly used Cox regression model. First, by relating covariates to parameters of the health process, TR results in clinically meaningful covariate effects (Pennell et al. 2010). Second, the effects of covariates on the hazard vary with time avoiding the proportional hazards assumption of the Cox model.

In a 2004 paper, Aalen and Gjessing (2004) reviewed several survival models based on the Ornstein–Uhlenbeck (OU) process. The OU process is a modification of the Wiener process that includes drift toward an equilibrium state or homeostasis. Many biological processes tend to a state of homeostasis; for instance, physiological processes ensure that body temperature (Blessing 1987) and ion levels in the blood (Chiras 2005) are maintained at healthy levels. If untreated, the health of a high grade cancer patient would decline until death, which can be viewed as a state of unhealthy homeostasis. Thus, in some contexts, the tendency to drift toward a state of homeostasis makes the OU process a more attractive choice for modeling patient health than the Wiener process. Also, the OU process avoids the independent increments assumption of the Wiener process; often, one would expect changes in health over a certain time interval to depend on the current state of health. The Ornstein–Uhlenbeck process has been used extensively in modeling longitudinal data (see, for example, Taylor and Law 1998; Li et al. 2007) but has received less attention in the context of survival analysis. Several authors have developed mathematical models for the hazard rate using a quadratic function of an OU process (Woodbury and Manton 1977; Myers 1981; Yashin 1985; Yashin et al. 2007). A few authors have also studied the first hitting time distribution of the OU process (Thomas 1975; Ricciardi and Sato 1988; Larralde

2004) but, to our knowledge, there are no papers on the use of the OU process in the context of threshold regression.

In this paper, we propose two threshold regression models based on the Ornstein–Uhlenbeck process. In our first model, the latent health process is modeled using an OU process whose initial status is modeled as a log-linear function of covariates whose coefficients may be interpreted in terms of relative baseline health. We then extend the approach to accommodate a cure rate, in which case a portion of the population will never experience the event of interest. Cure rate models are commonly used in cancer clinical trials (see Ibrahim et al. 2001 for a review). For example, in a study of an experimental surgical procedure, $100 \times p$ percent of patients may remain cancer free (i.e., are "cured") and the remaining $100(1 - p)$ percent will eventually experience a tumor recurrence. We model latent cancer progression following treatment (e.g., surgery) using a mixture of two processes: (1) among cured patients, cancer progression is modeled using a non-stochastic process which remains at its post-treatment state and (2) in situations where treatment did not successfully eradicate cancer, the Ornstein–Uhlenbeck process is used to model cancer recurrence. Covariates are related to the cure rate through a logit link and the initial status of the OU process through a log-link. A related approach has been used to extend standard parametric (Farewell 1982; Yamaguchi 1992) and semiparametric (Kuk and Chen 1992) survival models to include a cure rate. However, the conceptual appeal of our model and the ability to accommodate non-proportional hazards make our approach more attractive for certain applications including cancer studies.

The remainder of the paper is organized as follows: In Sect. 2, we present our stochastic model for health and Ornstein–Uhlenbeck threshold regression (OU-TR) model. In Sect. 3, we present the OU-TR mixture model for survival data with a cure rate. In Sect. 4, we summarize simulation results for each model and apply the models to overall and relapse free survival of melanoma patients enrolled in a clinical trial in Sect. 5. In Sect. 6, we discuss our findings and areas of future research.

2 Ornstein–Uhlenbeck threshold regression

2.1 Stochastic model for health

In a biomedical study of time to an event, let there be n subjects observed starting at baseline (time $t = 0$) with the ith ($i = 1, \ldots, n$) subject observed over $(0, T_i)$, where T_i (the observation time) is either that subject's event time (S_i) or censoring time (U_i). We assume independent censoring meaning that

$$\mathrm{P}(t \leq T_i < t + dt, T_i = S_i | T_i \geq t, \text{past}) = \mathrm{P}(t \leq S_i < t + dt | S_i \geq t), \quad (2.1)$$

where "past" denotes the event history up to time t for all n subjects and dt is some small increment in time (Aalen et al. 2008).

In our model, we assume that subject health is characterized by a latent stochastic process which takes on a value of $X_i(t)$ for $0 \leq t \leq T_i$. The event of interest occurs when $X_i(t)$ equals some predetermined value a for the first time, also known as the

first hitting time of the process. For instance consider a study of high-grade cancer patients. Suppose a subject dies once their health process reaches its lowest possible state, a. Upon enrollment, the health of the ith subject is $X_i(0) = x_{0i} > a$. If untreated or administered an unsuccessful treatment, that patient's health should decline fairly rapidly with death occurring once health has declined to a at time S_i (i.e., if the sample path is continuous, $X_i(s_i) = a$ given $S_i = s_i$).

Several options exist for the stochastic process $X_i(t)$ as discussed in Lee and Whitmore (2006). Probably the most popular choice is the Wiener process with drift where the first hitting time follows the inverse Gaussian distribution (Chhikara and Folks 1989), which is easy to work with computationally. Lee et al. (2004) also argue that, in many circumstances, the Wiener process is a realistic model for health status because it allows bidirectional changes in health which can occur on a daily or hourly basis. However, the Wiener process is not without its limitations. For instance, the model assumes normal, independent increments in the health status. While the independence assumption improves tractability, biologically, it often makes more sense for changes in health to depend on the patient's current health status.

The Ornstein–Uhlenbeck (OU) process addresses the independent increment limitation of the Wiener process. The OU process is a mean reverting modification of the Wiener process in that it has a propensity to drift back in the direction of a fixed level. Let $W_i(t)$, denote the value of a standard Brownian motion process (mean zero, volatility of one) for subject i at time t. For every small increment in time, dt, the change in the Wiener process $W_i(t + dt) - W_i(t) \sim N(0, dt)$. The OU process, represented by $X_i(t)$, is defined by the stochastic differential equation

$$dX_i(t) = (a - bX_i(t))dt + \sigma dW_i(t), \tag{2.2}$$

which has three parameters a, b, and σ, where a and b are real numbers and $\sigma > 0$. Assuming $b > 0$ and a fixed initial health status, $X_i(0) = x_{0i}$, the change in $X_i(t)$ over $(t, t + dt)$ is dependent on the current value of the process and has drift toward a/b (i.e., $E(X_i(t)) = a/b$ as $t \to \infty$) but is agitated by the Gaussian noise contained in $dW_i(t)$, often called white noise (Aalen and Gjessing 2004). The size of the increments in the OU process are dependent on the distance between the current state of the process and the point of homeostasis through the term $(a - bX_i(t))dt$, or equivalently $([a/b - X_i(t)]/b)dt$, in Eq. (2.2). Thus, the increments in the process tend to get smaller as the process approaches the point of homeostasis. Figure 1 compares a sampled path of an OU process to a Wiener process with the same initial state. The method for generating the sample path from the OU process is described in Section B of the Supplementary Materials. The values of the Wiener process can change drastically from one time point to the next; in contrast, changes in the Ornstein–Uhlenbeck process are more modest and the process drifts toward its homeostasis level at 0.

The tendency of the OU process to revert to a point of equilibrium is similar to the homeostatic nature of many biological processes. An example of a biological process that exhibits homeostasis is kidney function. Kidneys get rid of extra water and ions from blood through passage of urine. Thus, the kidneys carry out homeostatic regulation by removing waste or excess products from the body.

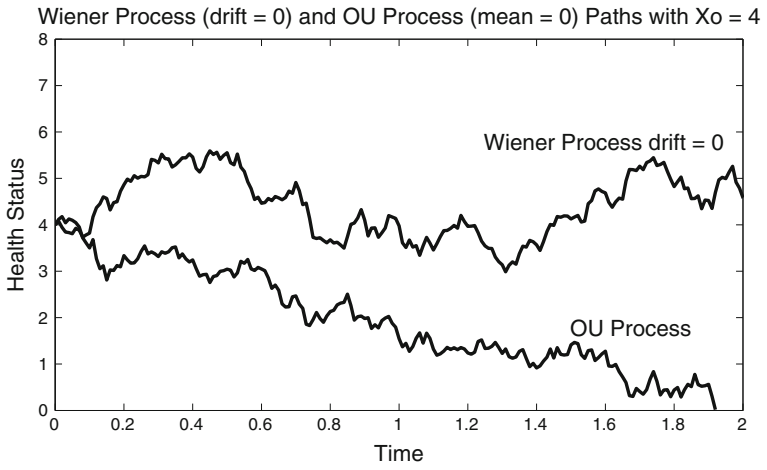

Fig. 1 Sample paths from the Wiener and Ornstein–Uhlenbeck processes

2.2 First hitting time distribution

2.2.1 Ricciardi–Sato distribution

In this paper, we are interested in using the Ornstein–Uhlenbeck process to model the time required for a subject with initial health $x_{0i} > a/b$ to reach the point of homeostasis, a/b, defined in Eq. 2.2. If the event under study is a positive one (e.g., discharge from a hospital or intensive care unit) the threshold a/b is regarded as a point of healthy homeostasis. In other situations, we may be modeling time to an unhealthy homeostasis. For instance, the health of a subject with high grade cancer will progress rapidly toward tumor recurrence or death if untreated or if treatment is unsuccessful.

Unfortunately, under the general OU model defined in Eq. 2.2, the distribution of the first hitting time (i.e., time for the process to reach a/b from x_{0i}) does not have a closed form. However, Ricciardi and Sato (1988) found that the distribution does have a closed form when $a = 0$, $\sigma^2 = 2$, and $b = 1$; in this case the barrier for the latent process is zero and the probability density function of the first hitting time t is

$$f(t|x_{0i}) = \sqrt{\frac{2}{\pi}} x_{0i} \frac{e^{2t}}{(e^{2t}-1)^{3/2}} \exp\left(-\frac{x_{0i}^2}{2(e^{2t}-1)}\right) \tag{2.3}$$

and the corresponding survival function is

$$S(t|x_{0i}) = 2\Phi\left(\frac{x_{0i}}{\sqrt{e^{2t}-1}}\right) - 1, \tag{2.4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function. We will henceforth refer to this distribution as the Ricciardi–Sato distribution. Note that there is one free parameter which determines the shape of the Ricciardi–Sato distribution, the

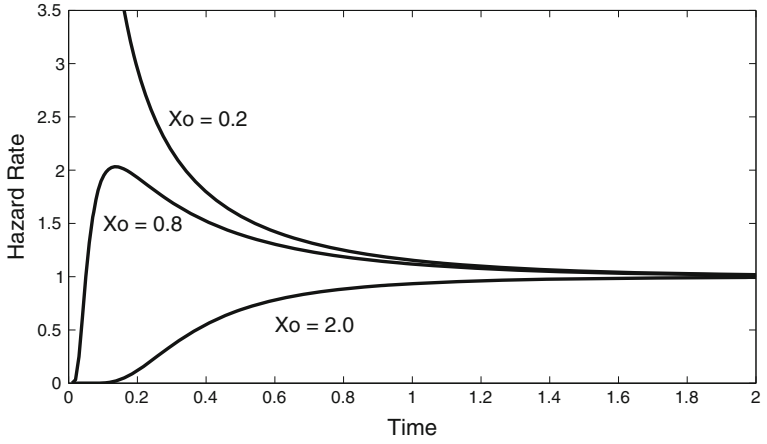

Fig. 2 Effects of initial health (x_0) on the shape of the hazard function of the Ricciardi–Sato distribution. The upper hazard function ($x_0 = 0.2$) starts out at zero, has a strong initial increase and then is generally decreasing (the *top of the curve* was cut off to avoid obscuring the shape of the other two curves)

initial health status or x_{0i}. Figure 2 shows how the shape of the hazard function, $f(t|x_{0i})/S(t|x_{0i})$, changes with the initial state of the health process. This figure is a redrawing of Figure 1 in Aalen and Gjessing (2004). The hazard rises from zero and approaches an asymptote at one as t increases. It possesses an intermediate peak hazard rate greater than one when x_0 is small, specifically, when $x_0 < 1.713$.

The Ricciardi–Sato distribution can be converted to a couple of well known distributions through transformation. First, the transformation $V = e^{2T} - 1$ converts a Ricciardi–Sato random variable T to the first hitting time distribution of a Brownian motion process on time scale V that starts at distance x_0 from the barrier (Chhikara and Folks 1989). Second, the transformation $U = (e^{2T} - 1)^{-1}$ converts the Ricciardi–Sato distribution to a gamma distribution with mean x_0^{-2} and variance $2x_0^{-4}$.

2.2.2 Scaled Ricciardi–Sato distribution

Unfortunately, the Ricciardi–Sato distribution has just one free parameter which limits its flexibility. Of particular concern to us is that it lacks a scale parameter which means that the fit of the distribution to a set of data depends on the time unit (i.e., days, weeks, months, or years). To address this limitation, we propose a simple modification to the Ricciardi–Sato distribution. Let T denote the survival time whose units are established by the analyst. We model $T = T^*/\alpha$ where T^* follows the Ricciardi–Sato distribution and $\alpha > 0$ is an unknown scale parameter which converts the time scale chosen by analyst to the appropriate time scale of the latent health process. The conversion of the time scale is somewhat related to previous methods for converting calendar time to operational time in Wiener process FHT models (see, for example, Lee et al. 2004). Under our simple transformation, the probability distribution function of T becomes

$$f_{\mathrm{sc}}(t|x_{0i}, \alpha) = \sqrt{\frac{2}{\pi}}\alpha x_{0i} \frac{e^{2\alpha t}}{(e^{2\alpha t} - 1)^{3/2}} \exp\left(-\frac{x_{0i}^2}{2(e^{2\alpha t} - 1)}\right) \tag{2.5}$$

and the corresponding survival function is

$$S_{\text{sc}}(t|x_{0i}, \alpha) = 2\Phi\left(\frac{x_{0i}}{\sqrt{e^{2\alpha t} - 1}}\right) - 1. \tag{2.6}$$

Our proposed survival models are based on this distribution, which we will henceforth call the scaled Ricciardi–Sato distribution.

2.3 Regression model

Several authors have proposed threshold regression (TR) models in which generalized link functions are used to relate parameters of the health process to covariates. Such TR models have been proposed for first hitting time models based on the Wiener (Lee et al. 2000, 2004; Lee and Whitmore 2006; Aalen and Gjessing 2001; Aalen et al. 2008) and gamma (Lawless and Crowder 2004) processes, but to our knowledge, none have been proposed for the OU process. Assuming that survival times follow the scaled Ricciardi–Sato distribution, we propose a regression model which uses a log-link to relate covariates to the initial state of the latent OU health process:

$$\log(x_{0i}) = \gamma_0 + \gamma_1 z_{i1} + \cdots + \gamma_r z_{ir}, \tag{2.7}$$

where $z_{i1}, \ldots, z_{ir}$ are a set of r covariates measured on the ith subject and $\gamma_1, \ldots, \gamma_r$ are their respective regression coefficients with γ_0 denoting the intercept (i.e., initial health assuming all $z_{ik} = 0$). With the inclusion of fixed covariates, the independent censoring assumption in Eq. (2.1) is modified slightly:

$$\begin{aligned} \mathrm{P}(t \le T_i < t + dt, T_i = S_i | T_i \ge t, \text{past}^*) \\ = \mathrm{P}(t \le S_i < t + dt | S_i \ge t, z_{i1}, \ldots, z_{ir}), \end{aligned} \tag{2.8}$$

where "past*" denotes the event history up to time t and covariate values for all n subjects.

An attractive feature of our model is that it yields clinically meaningful covariate effects. For example, suppose a researcher is interested in modeling time to tumor recurrence following surgery to remove a malignant melanoma. Let $z_{ik} = 1$ if the patient was a high grade case of melanoma (Stage III or IV) and 0 if a moderate-low grade case (Stage 0–II). The value of e^{γ_k} would thus indicate how healthy (in terms of the latent process) a high grade patient was immediately following surgery relative to a low-grade case. Equivalently, the covariate effects allow comparison of the baseline proximity to the threshold which triggers the event. For instance, in our example, if $e^{\gamma_k} = 0.5$ this means that a high grade patient was 50 % closer to a tumor recurrence immediately following surgery than a moderate–low grade case.

Under special circumstances, the coefficients could also be interpreted in terms of effects on median survival. To see this, we first note that the median survival given a set of predictors $\mathbf{z}_i = (z_{i1}, \ldots, z_{ir})'$ $(M_{\mathbf{z}_i})$ is

$$M_{\mathbf{z}_i} = \frac{1}{2\alpha} \log\left(\frac{x_{0i}^2 + [\Phi^{-1}(0.75)]^2}{[\Phi^{-1}(0.75)]^2}\right), \tag{2.9}$$

where $\Phi^{-1}(0.75)$ is the 75th percentile from the standard normal distribution. When $x_{0i} \geq 2$

$$M_{\mathbf{z}_i} \approx \frac{1}{2\alpha} \log\left(\frac{x_{0i}^2}{[\Phi^{-1}(0.75)]^2}\right),$$

or equivalently,

$$M_{\mathbf{z}_i} \approx \frac{1}{\alpha}\Big[\gamma_0 + \gamma_1 z_{i1} + \cdots + \gamma_r z_{ir} - \log\{\Phi^{-1}(0.75)\}\Big].$$

Thus, when all other predictors are held constant, the difference in median survival corresponding to a one-unit difference in z_j is approximately γ_j/α. Therefore, when x_0 is moderate-to-large in value, the coefficients are proportional to a difference in median survival times; when the data follow the Ricciardi–Sato distribution (i.e., $\alpha = 1$) the covariate effects could be interpreted directly as a difference in medians.

2.4 Methods of estimation and inference

Estimation of parameters occurs through maximum likelihood. After solving for the MLEs, asymptotic standard errors are calculated using the observed Fisher information and may be used in performing Wald tests. Likelihood ratio tests of coefficients are also simple to perform using our model. Technical details regarding the methods and software used to obtain our estimates are available in Section A of the Supplementary Materials.

3 Ornstein–Uhlenbeck threshold regression mixture model

3.1 General framework

In cancer clinical trials, the hope is that the treatment under study (e.g., drug or surgical procedure) will prevent recurrence of the disease. Thus, it is expected that a certain proportion of the study population (p) will be cured following treatment. Therefore cure rate models, models which accommodate a proportion of the population which will never experience the event, are frequently used in cancer studies (Ibrahim et al. 2001).

Motivated by studies evaluating cancer treatments, we extend the OU-TR model presented in Sect. 2 to incorporate a cure rate. Our method is related to previous extensions of the Cox model (Kuk and Chen 1992) and parametric survival models (Farewell 1982; Yamaguchi 1992) in that we model the study population using a mixture of two distributions: one corresponding to the cured subjects and one corresponding to the subjects who will experience the event of interest if followed long enough. Let Y_i be a latent variable denoting group membership; $Y_i = 1$ if subject i is not susceptible to

experiencing the event under study (or is cured), $Y_i = 0$ if subject i will (eventually) experience the event. As in Sect. 2, let $X_i(t^*)$ denote level of the latent health process of subject i at time t^* (i.e., time on the scale appropriate for the latent process) and $X_i(0) = x_{0i}$ denote the initial value (e.g., the value immediately following surgery to remove a malignant mass). Given Y_i, we assume

$$X_i(t^*)|Y_i = Y_i x_{0i} + (1 - Y_i)X^*_{0i}(t^*) \tag{3.1}$$

where $X^*_{0i}(t^*)$ is the health process assuming that subject i will eventually experience the event, which follows the OU process defined in Sect. 2. Lee and Whitmore (2006) and Balka et al. (2009) proposed related extensions of gamma and Wiener process FHT models respectively. In words, our model means that the health of "cured" subjects remains at x_{0i} over the time period under study. Subjects who are not "cured" follow the OU stochastic process defined in Sect. 2 and revert to the process mean (0) at which point the event occurs. This model is attractive for modeling time to tumor recurrence (or cancer relapse) following surgery to remove a malignant mass. In this setting, $X_i(t^*)$ is used to model latent cancer progression. If the surgery is successful, the patient's cancer has been eradicated and $X_i(t^*)$ will remain at its post surgery state (x_{0i}). However, if the surgery was unsuccessful, cancer will progress and $X_i(t^*)$ will eventually hit the threshold of 0 at which point another tumor appears.

When a subject doesn't experience an event by time t_i, the following could be true: (1) The subject will never experience the event ($Y_i = 1$) or (2) The subject will eventually experience the event ($Y = 0$), but did not during the observation period. Thus, the marginal survival function for this model is

$$S^*(t_i|p_i, x_{0i}, \alpha) = p_i + (1 - p_i)S_{\text{sc}}(t_i|x_{0i}, \alpha), \tag{3.2}$$

where $S_{\text{sc}}(t_i|x_{0i}, \alpha)$ is the survival function of the scaled Ricciardi–Sato distribution (Eq. 2.6). The portion of the survival function $S^*(t_i|p_i, x_{0i}, \alpha)$ that captures subjects in category (1) is p_i and $(1 - p_i)S_{\text{sc}}(t_i|x_{0i}, \alpha)$ captures those in category (2). Note that $S^*(\infty|p_i, x_{0i}) = p_i > 0$ indicating an improper survival function.

3.2 Regression model

The model presented in the previous section assumes a common cure rate in the population. However, in most patient populations, one would expect it to vary with patient characteristics. For example, the chance that surgery to remove a malignant mass is successful probably depends on the stage of cancer. Also, as in Sect. 2, baseline health is likely to vary from patient-to-patient in ways that can be quantified by measured characteristics (e.g., age at the time of surgery). Thus, to accommodate patient-to-patient variability in the parameters of our mixture model, we include covariate effects in two different places. As in Sect. 2.3, the initial health status (x_{0i}) is linked to a set of covariates $z_{i1}, \ldots, z_{ir}$ through the log-link function defined in Eq. 2.7. As in the OU-TR model, these coefficients may be interpreted in terms of relative initial health or relative proximity to the threshold at baseline.

A set of subject-specific covariates $\mathbf{w}_i = (w_{i1}, \ldots, w_{iq})$ is also linked to the cure rate (p_i) which is modeled as the logistic function:

$$p = \frac{\exp\{\beta_0 + \beta_1 w_{i1} + \cdots + \beta_q w_{iq}\}}{1 + \exp\{\beta_0 + \beta_1 w_{i1} + \cdots + \beta_q w_{iq}\}}, \tag{3.3}$$

where β_0 is the intercept and $\beta_1, \ldots, \beta_q$ are the coefficients of the covariates. The regression parameters are interpreted as in any multiple logistic regression model for a binary response. For instance, let $w_{ik} = 1$ if subject i has high grade melanoma and 0 if subject i has low-moderate grade melanoma. In this example, e^{β_k} is the ratio of the odds of one's cancer being cured by surgery for high grade versus low-moderate grade patients. Farewell (1982), Kuk and Chen (1992), and Taylor (1995) also used a logistic function for the cure rate, though their models for survival conditional on $Y_i = 0$ (not being cured) are substantially different from ours; Farewell used a Weibull model, Kuk and Chen used the Cox model, and Taylor used Kaplan–Meier estimates. Balka et al. (2009) proposed a mixture model which used a Wiener process TR model to model survival conditional on $Y_i = 0$ and Xiao et al. (2012) extended their approach to allow the cure rate to depend on covariates through a logit link. As mentioned in Sect. 2.1, the OU process is often a more a realistic model for health than the Wiener process because the increments of the OU process are dependent on the current state of the process and it reverts to a point of homeostasis.

Note that we use different notation for the covariates for p_i and x_{0i} because they need not be the same. For instance, if a treatment is not expected to have an immediate effect on health (or the cancer progression process if modeling time to tumor recurrence) then one would not expect it to affect x_{0i} (health immediately following the start of treatment) but, if successful, it should affect the cure rate. However, some covariates may affect both aspects; for example, stage of melanoma probably affects both patient health immediately following surgery to remove the malignant mass and the chances that the surgery was successful. The ability to distinguish between these two types of covariate effects is a strength of our model. No interpretation problems are caused by including a predictor in both parts of the model, with the exception that when $x_0 \geq 2$, the coefficient in the initial health status model cannot be interpreted as a difference in median survival times.

3.3 Methods of estimation and inference

Maximum likelihood methods are again used to estimate the parameters in the OU-TR mixture model and asymptotic standard errors are obtained using observed Fisher information. More details are available in Section A.2 of the Supplementary Materials.

4 Simulation studies

Simulation studies were performed to evaluate the bias of the maximum likelihood estimates of regression coefficients and the scale parameter (α) and validity of their standard errors under various coefficient values and censoring rates. The description

of the simulation methods and their results are provided in Section B (OU-TR) and Section C (OU-TR Mixture Model) of the Supplementary Materials. Bias was small ($<10\,\%$) for each model regardless of the censoring rate. In addition, the standard errors based on Fisher information were similar to the empirical standard deviations of the estimates, supporting our methods for standard error estimation.

We also conducted some simulations examining the robustness of our models to assumption violations. For the OU-TR model, we examined sensitivity to misspecification of the parameters a, b, and σ^2 in the OU process (Section B of Supplementary Materials). We found that misspecification of b and σ^2 was not problematic as covariate effects and survival probability estimates were unbiased. However, when a was misspecified, covariate effects were severely biased and survival probability estimates were slightly biased. For the OU-TR mixture model, we examined robustness to the assumption that the cure rate p is independent of the initial health status X_{0i} conditional on any predictors shared across models (Section C of Supplementary Materials). We found that omission of a predictor common to both p and x_{0i} results in biased estimates of α and the coefficients of x_{0i}. However, similar results were obtained by omitting a predictor that is only related to x_{0i}. Thus, the bias observed when omitting a predictor common to x_{0i} and p may be due to unexplained heterogeneity in x_{0i}, not an unexplained relationship between x_{0i} and p.

5 Application of OU-TR models to survival of melanoma patients

5.1 Data and methods

We applied our OU-TR models to survival data from the E1690 clinical trial analyzed in Kirkwood et al. (2000). We analyzed a subset of data from this trial published in Ibrahim et al. (2001). This dataset or a variation thereof was also used in a paper by Chen et al. (2002), who used a Bayesian approach to model cure rates for this disease. High-risk melanoma patients undergoing definitive surgery were randomized into two groups; one that received treatment with Interferon Alfa-2b and the other did not receive treatment. We considered only the control subjects in our analysis in order to develop natural history models for melanoma patients undergoing definitive surgery. Survival data were recorded as time since entry into the study, with the surgery occurring soon afterward; for more details about the study design, see Kirkwood et al. (2000).

We performed two analyses. We first modeled time to death, or overall survival (OS), using the OU-TR model. Potential explanatory variables for x_0 included the subject's age at the start of the clinical trial, sex, Breslow score, performance score ($0=$ fully active, $1=$ restricted in activity), and nodal category. The nodal category is broken down into four groups: N0 denotes no spread to nearby lymph nodes, N1 indicates spread to one nearby lymph node, N2 denotes spread to two or three nearby lymph nodes or the melanoma has spread to nearby skin or near a lymph node area without actually reaching the lymph nodes and N3 indicates spread to four or more lymph nodes or spread to lymph nodes which are clumped together or spread to nearby skin or toward a lymph node area and into the lymph node(s) (American Cancer Society

2012). We used all subsets variable selection to select the set of explanatory variables for x_0 which produced the smallest Integrated Brier Score (IBS), which is a measure of inaccuracy of survival predictions. We used the definition of the IBS provided in Graf et al. (1999); the details are in Section D of the Supplementary Materials.

We next modeled relapse free survival, using our OU-TR mixture model and the same set of clinical predictors considered in modeling OS. Relapse free survival (RFS) is the length of time following cancer treatment (surgery in this case) until signs or symptoms of that cancer return. With five potential predictors, each of which could be in the regression model for p, x_0, or both, the number of possible models for all subsets model selection was enormous. Thus, to reduce the dimensionality of our model selection, we used a two step approach. First, we fit the OU-TR model to the RFS data and used an all subsets approach to select the set of explanatory variables for x_0 which produced the smallest IBS; the selected predictors were used to model x_0 in the OU-TR mixture model. We then performed all subsets model selection to identify the set of predictors for the cure rate (p) which produced the smallest IBS. A total of 198 patients with complete covariate information were used in the analysis of OS (88 events) and RFS (121 events).

Conceptually, our OU-TR models are a good match for these data. Since the data are from high risk cancer patients, who are presumably very frail, it is natural to think of their overall health approaching an unhealthy homeostasis (death) thus supporting the use of the OU-TR model for OS. In the RFS analysis, cancer progression is modeled using a latent stochastic process $X(t)$. It is expected that the surgery eradicated the cancer for some patients and hence their cancer progression process remained stable at its initial state ($X(0) = x_0$). For other patients, the surgery was unsuccessful in which case cancer quickly progressed to a tumor recurrence. Thus, the cancer progression process of this second group can be thought of as being pulled toward an unhealthy homeostasis making the OU process a reasonable model. In addition, one would expect that the chances of being cured by surgery (p) and the state of the process at the time of entry (x_0) to depend on factors related to patient health and severity of the cancer.

5.2 Results

5.2.1 Overall survival

All subsets variable selection revealed that the best model (in terms of IBS) contained three predictors: age, Breslow score, and nodal category. However, the coefficients of age and Breslow score were not significantly different from zero (Wald p values $= 0.36$ and 0.91, respectively) and dropping these predictors resulted in a minor increase in the IBS (from 0.207 to 0.210). Thus, in the interest of parsimony we dropped age and Breslow score and report the results from a model containing a single predictor for x_0: nodal category.

The coefficient estimates are provided in Table 1. Patients with one or more positive nodes were considerably less healthy upon entering the study than patients with no positive nodes, i.e., the value of their latent health process was closer to 0 at time $= 0$. The greatest difference was seen between patients in category N2 and N0; patients

Table 1 Final OU-TR and Wiener TR models for overall survival of melanoma patients undergoing definitive surgery

Model		Parameter estimate	SE	p value
OU-TR	Intercept	−6.134	17.698	0.729
	N1	−0.406	0.134	0.002
	N2	−0.680	0.147	<0.001
	N3	−0.587	0.149	<0.001
	α	5.07×10^{-7}	1.80×10^{-5}	
Wiener	Intercept	0.711	0.114	<0.001
	N1	−0.393	0.138	0.004
	N2	−0.659	0.152	<0.001
	N3	−0.570	0.154	<0.001
	μ	0.072	0.067	0.285

Coefficients reported for the Wiener process model are for the regression model for $\log(X_0)$

in category N2 were $100(1 - e^{-0.681}) = 49\,\%$ less healthy (x_0 was 49 % closer to 0) at study entry than patients with no positive nodes. While the coefficient for N3 was less negative than that of N2, which would suggest (counter-intuitively) that N3 patients were healthier than N2 patients at the beginning of the study, the difference between these coefficients was not statistically significant ($p = 0.552$). The estimate of α was very small which suggests that the unit of time we used in the analysis (years) was much larger than the appropriate scale for the latent process. When the analysis was re-run using years $\times 0.0001$ as the unit of time, the new estimate of α was larger (0.0073) but the coefficients of the nodal categories, their standard errors, and survival estimates did not change.

The estimated survival curves from our OU-TR model and Kaplan–Meier estimates are provided in Fig. 3. Our model estimates agreed fairly well with Kaplan–Meier curves for N0 and N3. Agreement between our estimates and the Kaplan–Meier estimates for N2 was also good up to the last event time in that group, after which our model appeared to underestimate survival. Fit of our model was the worst for N1 as it failed to capture a plateau in the survival curve at approximately 3.5 years after enrollment.

For comparison purposes, we also fit a Wiener process TR model to the data using the STTHREG package in STATA (Xiao et al. 2012). The initial state of the Wiener process (x_0) was modeled as a log-linear function of nodal category (same as the OU-TR model) and we assumed a constant drift (μ). Extending the model to allow the drift to differ by nodal category did not significantly improve fit (likelihood ratio p value =0.26). The coefficient estimates for the Wiener process model are provided in Table 1. The effects of nodal category agree with those of the OU-TR model; patients whose cancer had spread to the lymphnodes were less healthy at study entry than patients with no spread to the lymphnodes, with category N2 having the smallest initial health value. The survival estimates from the Wiener process model were close to the OU-TR estimates (Fig. 3). The Wiener process model fit the N1 survival data better than the OU-TR model, the OU-TR model fit the data slightly better for categories N0 and N3, and the fit of the two models was comparable for N2. The IBS for the Wiener process model was 0.204, which was slightly better than the OU-TR model (0.210).

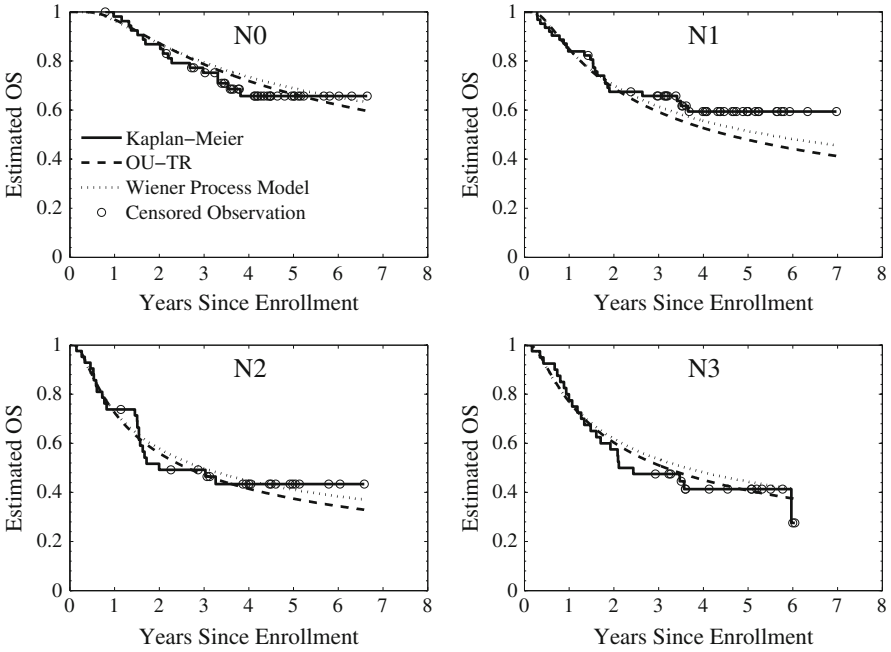

Fig. 3 Overall survival of melanoma patients undergoing definitive surgery. Estimates for the OU-TR model were calculated using the survival function of the Scaled Ricciardi–Sato distribution (Eq. 2.6)

5.2.2 Relapse free survival

In the first stage of our analysis, we found that the best OU-TR model contained just one explanatory variable: nodal category. Therefore, in all subsets selection of the best OU-TR mixture model, we compared models containing nodal category as the sole predictor for X_0 and all possible combinations of the available prognostic variables in the model for the cure rate. The model with the best Brier score (0.219) contained age, sex, Breslow score, and nodal category as predictors for p though a reduced model containing Breslow score and nodal category as the sole predictors for p had a Brier score which was only 1 % larger (0.222). Thus, in the interests of parsimony, we selected the reduced model as our final model.

The coefficient estimates and standard errors for our final model are provided in Table 2. The coefficients for the initial status of the process (x_0) allow comparisons of cancer progression at the time of entry into the study with smaller values of x_0 indicating that a patient was closer to a cancer relapse (smaller distance from 0, the point of homeostasis of the OU process). Not surprisingly, the coefficients became more negative as the nodal category increased in severity. For instance, a patient whose cancer had spread to one lymph node prior to surgery (category N1) had an x_0 which was $100(1 - e^{-0.446}) \approx 36\,\%$ smaller than a patient whose cancer had not spread to the lymph nodes (N0) or, equivalently, at the beginning of the study N1 patients were 36 % closer to the threshold at which a relapse happens than N0 patients. In the worst

Table 2 Final OU-TR mixture model and Wiener process TR model for relapse free survival of melanoma patients undergoing definitive surgery

		Parameter estimate	SE	p value
OU-TR mixture model				
X_0	Intercept	−0.720	0.468	0.124
	N1	−0.446	0.171	0.009
	N2	−0.462	0.176	0.009
	N3	−0.522	0.172	0.002
p	Intercept	1.421	1.097	0.195
	Breslow	−0.400	0.213	0.060
	N1	−0.999	0.854	0.242
	N2	−2.257	1.295	0.081
	N3	−3.075	1.402	0.028
	α	0.155	0.141	
Wiener process				
X_0	Intercept	0.166	0.216	0.441
	Breslow	−0.035	0.026	0.181
	N1	−0.527	0.176	0.003
	N2	−0.668	0.204	0.001
	N3	−0.712	0.198	<0.001
μ	Intercept	0.245	0.173	0.157
	Breslow	−0.015	0.021	0.467
	N1	0.142	0.181	0.433
	N2	−0.012	0.204	0.952
	N3	−0.155	0.203	0.446

case scenario, nodal category N3, patients were $100(1 - e^{-0.522}) \approx 41\,\%$ closer to the cancer relapse threshold upon entering the study than patients in category N0. The cure rate also decreased as Breslow score increased and spread to the lymph nodes increased. For instance, the cure rates for each nodal category at the median Breslow score (3.35 mm) were 0.520 (N0), 0.285 (N1), 0.102 (N2), and 0.048 (N3). The cure rates were considerably smaller at 75th percentile of Breslow score (5.025 mm): 0.357 (N0), 0.170 (N1), 0.055 (N2), and 0.025 (N3).

As in the OS analysis, we also fit a Wiener process model, which this time contained Breslow score and nodal category as predictors of the initial state (via a log link function) and drift (via an identify link function). A cure rate can occur in the Wiener process model via a drift away from the threshold (Balka et al. 2009). Thus, both predictors were kept in the model for the drift regardless of significance to allow them to be related to a possible cure rate. The coefficient estimates from the Wiener process model are in Table 2. Similar to the OU-TR mixture model, there was no effect of Breslow score on the initial status of the latent Wiener process and proximity to the tumor recurrence threshold immediately following surgery increased as spread to the lymph nodes increased. There was no effect of either Breslow score or nodal category on the drift of the Wiener process.

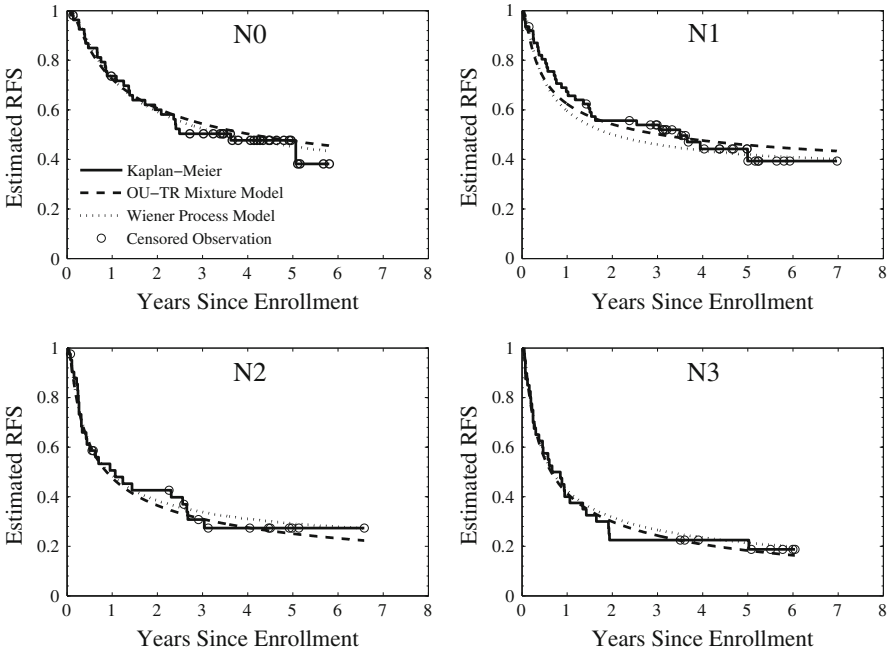

Fig. 4 Relapse free survival of melanoma patients undergoing definitive surgery. Estimates from the OU-TR mixture and Wiener process models were calculated using the median Breslow score in each nodal category. The survival function for the OU-TR mixture model was calculated using Eq. 3.2

Estimated survival curves from our OU-TR mixture model and the Wiener process model were compared to Kaplan–Meier estimates in Fig. 4. It is difficult to determine visually which model provided a better overall fit because over certain time frames, the OU-TR estimates were closer to the KM curve while over other time frames fit of the Wiener process appeared superior. However, the IBS for the OU-TR mixture model (0.222) was considerably less than that of the Wiener process model (0.308) suggesting that, globally, our model provided a better representation of the data.

6 Discussion

Motivated by applications in cancer, we proposed two innovative threshold regression (TR) models based on the Ornstein–Uhlenbeck process. Our first model, the Ornstein–Uhlenbeck threshold regression (OU-TR) model, models latent health using a stochastic process whose baseline level is modeled using a log-linear function of covariates. Increments in the process over time are dependent on the current state of the process (thus avoiding the independent increments assumption of Wiener process models) with an event occurring once the process has reached a point of homeostasis. We applied the OU-TR model to overall survival data of high risk melanoma patients undergoing definitive surgery and found that, for the most part, the model adequately described the survival experience of these patients. We next proposed an OU-TR

mixture model which generalized our first model to include a cure rate. This second model provides two sets of covariate effects: the effects on baseline health (x_0) and probability of being cured by treatment (p). The model was applied to the relapse free survival data of the melanoma patients and demonstrated good fit. Sample Matlab code for fitting the OU-TR and OU-TR mixture models are available upon request from the corresponding author.

Conceptually, the OU-TR model is most appropriate for survival data of patients whose health should gravitate toward a point of homeostasis. For instance, one would expect a high grade cancer patient's health to gravitate toward its nadir at which point death occurs. However, extending the Ricciardi–Sato distribution to include a scale parameter (α) allows our model to be applied to a wide variety of data scenarios because if the event of interest is slow to develop, α will adjust accordingly.

The OU-TR mixture model should provide good fit in scenarios where treatment either immediately cures a subject or is unsuccessful and the patient's health will steadily decline until an event occurs; a perfect example, is time to cancer recurrence of patients who underwent definitive surgery (i.e., our data example). However, if prolonged exposure to a treatment (e.g., chemotherapy) is needed to cure a disease, our model may not be a good fit.

Our method assumed that patient health follows an OU process with fixed parameter values ($a = 0, \sigma^2 = 2$, and $b = 1$). Our simulation studies suggest that the assumptions on σ^2 and and b are not problematic because misspecification is accounted for by an adjustment in the value of intercept of the initial health status (γ_0) and scale (α), respectively. However, the covariate effect estimates in the initial health status are not robust to misspecification of the parameter a. Thus, we acknowledge that greater flexibility may be obtained if we were able to estimate a from the data. However, the lack of a closed form for the first hitting time distribution is a major hurdle for such a model.

Our model also assumed a constant time scale (α) across subjects. This assumption was reasonable for our motivating application given that it involved short-term follow-up of a population of frail patients. A constant scale may make less sense in applications involving lengthy follow-up (e.g., a study of time to death over an entire lifetime) of subjects with very different exposure profiles (e.g., smokers versus non-smokers in a study of lung cancer). This limitation could be addressed by adding a regression model for α, but this would complicate interpretation. Since the initial status is affected by the time scale, the covariate effects in x_0 are only interpretable if the time scale remains constant across subjects. Furthermore, the effects of covariates on the time scale would not have a nice interpretation in terms of the latent health process; for instance, explaining what a difference in scale means to a clinician would be a difficult task.

If one is uninterested in the interpretation of covariate effects on the latent process and if follow-up is lengthy (e.g., a study of time to death over an entire lifetime) and among subjects with very different exposure profiles (e.g., smokers versus non-smokers in a study of lung cancer), it may be a good idea to model the scale because it would probably improve fit. This extension may be difficult though due to identifiability problems. For instance, in our analysis of overall survival data from the E1690 clinical trial, we found that the scale and intercept of x_0 were weakly identifiable from

each other (see Section A.1 of the Supplemenary Materials for more details). Thus, it is reasonable to expect more identifiability problems under a more complicated model for α.

The OU-TR mixture model could also potentially be extended to allow more flexible timing of being cured. For instance, in studies with a set dosing schedule, one might consider a model where $100p_k$ % of subjects would be cured following the kth treatment and the remaining subjects' health would continue to decline toward the threshold. Such a model could borrow features of previous extensions of Wiener process TR models to allow time-varying covariates (Lee et al. 2010) and coefficients (Li and Lee 2011).

Finally, in our simulations we found that unexplained heterogeneity in the initial health status can result in biased parameter estimates and under-estimation of standard errors. Thus, we are currently extending our model to include a random effect in the model for x_0 to account for unmeasured covariates.

Acknowledgments This work was supported by the National Institute of Occupational Safety and Health [OH008649].

References

Aalen OO, Gjessing HK (2001) Understanding the shape of the hazard rate: a process point of view. Stat Sci 16:1–22

Aalen OO, Gjessing HK (2004) Survival models based on the Ornstein–Uhlenbeck process. Lifetime Data Anal 10:407–423

Aalen OO, Borgan O, Gjessing HK (2008) Survival and event history analysis. Springer, New York

American Cancer Society (2012) Melanoma skin cancer. http://www.cancer.org/Cancer/SkinCancer-Melanoma/DetailedGuide/melanoma-skin-cancer-staging. Retrieved 26 Mar 2012

Balka J, Desmond A, McNicholas P (2009) Review and implementation of cure models based on first hitting times for Wiener processes. Lifetime Data Anal 15:147–176

Blessing W (1987) The lower brainstem and bodily homeostasis. Oxford University Press, New York

Chen M-H, Ibrahim J, Sinha D (2002) Bayesian inference for multivariate survival data with a cure fraction. J Multivar Anal 80:101–126

Chhikara RS, Folks JL (1989) The inverse Gaussian distribution: theory, methods, and applications. Marcel Dekker, New York

Chiras D (2005) Hum Biol. Jones and Bartlett Learning, Sudbury, MA

Cox DR, Miller HD (1968) The theory of stochastic processes. Chapman and Hall, London

Farewell VT (1982) The use of mixture models for the analysis of survival data with long-term survivors. Biometrics 38:1041–1046

Graf E, Schmoor C, Sauerbrei W, Schumacher M (1999) Assessment and comparison of prognostic classification schemes for survival data. Stat Med 18:2529–2545

Horrocks J, Thompson M (2004) Modelling event times with multiple outcomes using the Wiener process with drift. Lifetime Data Anal 10:29–49

Ibrahim JG, Chen M, Sinha D (2001) Bayesian survival analysis. Springer, New York

Kirkwood J, Ibrahim J, Sondak V, Richards J, Flaherty L, Ernstoff M, Smith T, Rao U, Steele M, Blum R (2000) High and low-dose interferon alfa-2b in high-risk melanoma: first analysis of intergroup trial E1690/S9111/C9190. J Clin Oncol 18:2444–2458

Kuk AYC, Chen C-H (1992) A mixture model combining logistic regression with proportional hazards regression. Biometrika 79:531–541

Lancaster T (1972) A stochastic model for the duration of a strike. J R Stat Soc Ser A 135:257–271

Larralde H (2004) A first passage time distribution for a discrete version of the Ornstein–Uhlenbeck process. J Phys A Math Gen 37:3759–3767

Lawless J, Crowder M (2004) Covariates and random effects in a gamma process model with application to degradation and failure. Lifetime Data Anal 10:213–227
Lee MLT, DeGruttola V, Schoenfeld D (2000) A model for markers and latent health status. J R Stat Soc Ser B 62:747–762
Lee MLT, Garshick E, Whitmore GA, Laden F, Hart J (2004) Assessing lung cancer risk to rail workers using a first hitting time regression model. Environmetrics 15:1–12
Lee MLT, Whitmore GA (2006) Threshold regression for survival analysis: modeling event times by a stochastic process reaching a boundary. Stat Sci 21:501–513
Lee MLT, Whitmore GA, Rosner B (2010) Threshold regression for survival data with time-varying covariates. Stat Med 29:896–905
Li Q, Shen X, Pearl D (2007) Bayesian modelling of the dynamics of hepatotoxicity. Stat Med 26:3591–3611
Li J, Lee MLT (2011) Analysis of failure time using threshold regression with semi-parametric varying coefficients. Stat Neerl 65:164–182
Myers L (1981) Survival functions induced by stochastic covariate processes. J Appl Probab 18:523–529
Park C, Padgett WJ (2005) Accelerated degradation models for failure based on geometric Brownian motion and gamma processes. Lifetime Data Anal 11:511–527
Pennell ML, Whitmore GA, Lee MLT (2010) Bayesian random-effects threshold regression with application to survival data with non proportional hazards. Biostatistics 11:111–126
Ricciardi L, Sato S (1988) First-passage-time density and moments of the Ornstein–Uhlenbeck process. J Appl Probab 25:43–57
Taylor JMG (1995) Semi-parametric estimation in failure time mixture models. Biometrics 51:899–907
Taylor J, Law N (1998) Does the covariance structure matter in longitudinal modelling for the prediction of future CD4 counts? Stat Med 17:2381–2394
Thomas M (1975) Some mean first-passage time approximations for the Ornstein–Uhlenbeck process. J Appl Probab 12:600–604
Woodbury MA, Manton KG (1977) A random-walk model of human mortality and aging. Theor Popul Biol 11:37–48
Xiao T, Whitmore G, He X and Lee M-L T (2012) Threshold regression for time-to-event analysis: the stthreg package. Stata J 12:257–283
Yamaguchi K (1992) Accelerated failure-time regression models with a regression model of surviving fraction: an application to the analysis of "permanent employment" in Japan. J Am Stat Assoc 87:284–292
Yashin A (1985) Dynamics in survival analysis: conditional Gaussian property versus Cameron–Martin formula. In: Krylov N, Liptser R, Novikov A (eds) Statistics and control of stochastic processes. Springer, New York, pp 466–475
Yashin A, Arbeev K, Akushevich I, Kulminski A, Akushevich L, Ukraintseva S (2007) Stochastic model for analysis of longitudinal data on aging and mortality. Math Biosci 208:538–551