Received: 4 June 2017 | Revised: 1 September 2017 | Accepted: 6 October 2017

DOI: 10.1002/sim.7547

RESEARCH ARTICLE

WILEY Statistics in Medicine

Time-to-event data with time-varying biomarkers measured only at study entry, with applications to Alzheimer's disease

Catherine Lee | Rebecca A. Betensky | for the Alzheimer's Disease Neuroimaging Initiative

Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Correspondence
Catherine Lee, Division of Research, Kaiser Permanente, 2000 Broadway, Oakland, CA 94612, USA.
Email: cal373@mail.harvard.edu

Present Address
Catherine Lee, Division of Research, Kaiser Permanente, 2000 Broadway, Oakland, CA 94612, USA

Funding information
NIH, Grant/Award Number: T32NS048005, R01NS094610, P50AG005134, P01AG036694, P30AG10161, R01AG17917 and U01 AG024904; DOD, Grant/Award Number: W81XWH-12-2-0012

Relating time-varying biomarkers of Alzheimer's disease to time-to-event using a Cox model is complicated by the fact that Alzheimer's disease biomarkers are sparsely collected, typically only at study entry; this is problematic since Cox regression with time-varying covariates requires observation of the covariate process at all failure times. The analysis might be simplified by using study entry as the time origin and treating the time-varying covariate measured at study entry as a fixed baseline covariate. In this paper, we first derive conditions under which using an incorrect time origin of study entry results in consistent estimation of regression parameters when the time-varying covariate is continuous and fully observed. We then derive conditions under which treating the time-varying covariate as fixed at study entry results in consistent estimation. We provide methods for estimating the regression parameter when a functional form can be assumed for the time-varying biomarker, which is measured only at study entry. We demonstrate our analytical results in a simulation study and apply our methods to data from the Rush Religious Orders Study and Memory and Aging Project and data from the Alzheimer's Disease Neuroimaging Initiative.

KEYWORDS

Cox model, delayed entry, left truncation, survival analysis, time-dependent covariates, time origin

1 | INTRODUCTION

Alzheimer's disease (AD) is the most common type of dementia in older adults, characterized by loss of cognition and functional impairment. Its prevalence is estimated to be 5.5 million[1] in 2017, with an estimated incidence of 476 000 in 2016 for people aged 65 years or older. Much of Alzheimer's research over the past decade has been dedicated to the identification and validation of biomarkers that reflect the underlying neuropathology of the disease. Biomarkers will help identify subjects early in their disease course, when therapeutic interventions might be successful. They also will help identify surrogate end points for progression, which will allow for clinical trials that do not require long follow-up to a clinical AD diagnosis.

Several models have been proposed for time-varying AD biomarkers such as Pittsburgh compound B–positron emission tomography (PIB-PET), fluorodeoxyglucose–positron emission tomography (FDG-PET), and magnetic resonance imaging

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

 Copyright © 2017 John Wiley & Sons, Ltd.

measures (MRI).[2,3] Relating these biomarker trajectories to time-to-event using a Cox regression model would allow for the prediction of disease progression. However, such an analysis presents a challenge, as imaging biomarkers are typically sparsely collected. This is problematic since the Cox regression with time-varying covariates requires observation of the covariate process at all failure times. The analysis could be simplified by treating the time-varying covariate collected only at study entry as a fixed baseline covariate representative of a subject's disease state at the time of measurement. However, in longitudinal AD studies, the time at study entry tends to be arbitrary, and the interpretability of such an analysis is tenuous.[4,5] In a time-to-event analysis, a well-chosen time origin should align subjects with respect to risk, so that conditional on baseline covariates subjects are comparable.[6] In addition, the choice of time origin should address the goals of the analysis and provide meaningful estimates.[4,7-12] A natural choice of time origin in AD is pathological disease onset, a point at which individuals are comparable with respect to risk. However, disease onset precedes clinical symptoms by decades, and thus, the exact time of onset is unknown. As an alternative, the time of first clinical diagnosis of mild cognitive impairment (MCI) is a reasonable and measurable time origin. Birth is a plausible time origin since disease risk increases with age; however, individuals at a given age may not necessarily be comparable with respect to risk, even after conditioning on fixed covariates. In summary, while using study entry as the time origin would enable the use of the time-varying biomarker as a fixed covariate in the analysis, it would jeopardize the interpretability of the analysis.

Several authors have considered the use of an incorrect time origin of study entry in a proportional hazards model with a fixed-time covariate. Assuming the true time origin is birth, Korn and colleagues[8] showed that fitting a proportional hazards model that uses study entry as the time origin does preserve the true regression coefficient corresponding to a fixed baseline covariate, provided that age at entry is included as a covariate in the model and the baseline hazard on the age scale follows an exponential form. Thiebaut and Benichou[13] and Pencina et al[11] investigated the practical ramifications of this result when the baseline hazard is not of exponential functional form (eg, Weibull or piecewise Weibull) and found that the bias in the regression coefficient was not large for a wide range of true values.

In the case of time-varying covariates, some simulation studies have suggested an extension of the result from Korn et al[8]; however, this has not been established analytically. Cnaan and Ryan[4] argued that using an incorrect time origin would preserve the regression parameter of interest provided that survival on the age scale is exponentially distributed. In a simulation study, Thiebaut and Benichou[13] found that use of study entry as the time origin results in unbiased estimates for a Cox model with constant baseline hazard function and a fully observed time-varying binary covariate. Through simulations, Thiebaut and Benichou[13] and Griffin et al[14] illustrated that bias can result when an incorrect choice of time origin is used in a Cox model with time-varying covariates when the true baseline hazard is Weibull. There are no papers that consider estimation based on a time-varying covariate that is observed only at study entry, with the time origin taken to be either onset or study entry.

There is a considerable literature on dealing with sparse observation of time-varying covariates[15-21] in survival regression models; standard methods include imputation, likelihood estimation, and joint modeling. These methods are not applicable to many current Alzheimer's studies because of the their very sparse observation of the time-varying imaging biomarker. Relevant to this setting, Sperrin and Buchan[12] proposed a 2-stage method for analyzing survival data with time-varying covariates measured at a single, arbitrary time point. They assume that the covariate process can be expressed as the sum of a population-level function and individual time-independent errors. The first stage involves estimating the population parameters of the covariate function and then calculating for each subject the residual error relative to the observed value. In the second stage, the estimated residual for each subject is included as a covariate in a survival model. They provided an analytical argument for the use of birth as the time origin, rather than study entry. They showed through simulations that their 2-stage residual method has superior predictive ability, as measured by Brier and logarithmic scores, when birth is used as the origin and the estimated residuals and age are included in the model as covariates, compared with 2 competing models that treat the time-varying covariate as fixed: one from study entry that adjusts for age as a covariate and another from birth that does not adjust for age.

In this paper, we consider the analysis of time-to-event data in conjunction with a time-varying covariate that is observed only at study entry. We assume the true time origin for the risk model is known and precedes study entry. We extend the results of Korn et al to the case of time-varying biomarkers and derive the conditions for which treating a time-varying biomarker as fixed at study entry results in an unbiased regression coefficient estimator. We then extend the work of Sperrin and Buchan,[12] who proposed methods for relating a time-varying covariate with fixed-time residuals to time-to-event when the covariate is measured only at study entry, by incorporating time-varying residuals. Furthermore, whereas Sperrin and Buchan focused on predictive ability, we investigate bias and standard error estimation. We use the sigmoidal AD biomarker model[22] as the basis for this approach. We support our analytical results and methods with simulation studies. Finally, we apply our methods to data from the Rush Religious Orders Study and Memory and Aging Project and

from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and demonstrate the impact of different treatments of the time-varying biomarker.

2 | NOTATION AND MODELS

For subject i, let U_i denotes the time from disease onset to event, let C_i denotes the time from disease onset to censoring, let E_i denotes the time from disease onset to study entry, and let $Z_i(u)$ denotes a time-varying covariate. We use the notation $\bar{Z}_i(u)$ to denote the history of Z up to and including time u. Then $T_i = U_i - E_i$ represents the time from study entry to event. The observed time from onset to event is $Y_i = \min(U_i, C_i)$, and $\delta_i = I(Y_i \leq C_i)$ is the status indicator. We assume that the time-varying covariates are collected only at study entry. The observed data are thus $(E_i, Y_i, Z_i(E_i), \delta_i)$.

We assume that the time-varying covariate, or biomarker, $Z_i(u)$, after transformation via $f(\cdot)$, satisfies a general additive error model, $f(Z_i(u)) = f(S(u)) + \varepsilon_i(u)$, where $S(u)$ is the population-level biomarker trajectory and $\varepsilon_i(u)$ is an individual-level time-dependent deviation. We assume that the individual-level deviations are defined by $\varepsilon_i(u) = a_i + b_i u$, so that

$$f(Z_i(u)) = f(S(u)) + a_i + b_i u, \tag{1}$$

where a_i and b_i are randomly distributed, both with mean 0; consequently, $\varepsilon_i(u)$ has mean 0. When f is the identity function, we refer to (1) as the *absolute error model* since the individual-level deviations are on the same scale as the population-level trajectory. When f is the identify function and $b_i \equiv 0$, the covariate model in (1) coincides with the covariate model provided by Sperrin and Buchan.[12] When $f(x) = \text{logit}(x/M)$, for some constant M, we refer to (1) as the *logit error model*.

In the setting of AD, many population-level biomarker trajectories are plausibly modeled using a sigmoidal function[2]

$$S(u) = \frac{M}{1 + e^{-\lambda(u - T_0)}}, \tag{2}$$

where the upper bound M is assumed to be known as the maximum value of the biomarker and the parameters λ and T_0, which dictate the shape of the sigmoid, are not known. When $T_0 = 0$, $S(0) = M/2$; the value of T_0 determines the lateral shift of this curve. The parameter λ determines the rate at which the sigmoid accelerates to its maximum value with larger values indicating a faster rate of increase; at $u = T_0$, the slope of the sigmoid is $M\lambda/4$.

The covariate function $S(u)$ should be selected on the basis of prior studies and subject matter understanding. Scatter plots of the cross-sectional data may also give insight into to the functional form of the covariate process. Choice of the error function, $f(\cdot)$, also requires consideration. The absolute error model may not be appropriate for covariates that are constrained by a maximum measured value, such as positron emission tomography amyloid, as it may lead to individual-specific trajectories that exceed the maximum value. The logit error model with time-varying deviations has an attractive interpretation when the population-level function is assumed to be the sigmoid in (2). The individual-specific deviations modify the acceleration parameter, λ, and the shift parameter, T_0, so that each subject is allowed to have an individual-specific trajectory. In particular, inserting the sigmoidal function (2) into the logit error model (1) yields an individual-specific sigmoidal covariate function, $Z_i(u) = M/[1 + e^{-\tilde{\lambda}(u - \widetilde{T_0})}]$, where $\tilde{\lambda}_i = \lambda + b_i$ and $\widetilde{T_{0,i}} = (\lambda T_0 - a_i)/(\lambda + b_i)$. The logit error model assumes that the time-varying covariate is linear on the logit scale, which may be appropriate only in some settings.

We assume that the true hazard function for the event of interest conditional on the history of Z, with time measured from onset, is

$$\lambda_{U_i}(u|\overline{Z}_i(\cdot)) = \lambda_{0U}(u) e^{\beta g(Z_i(u))}, \tag{3}$$

for some function g and baseline hazard function, $\lambda_{0U}(u)$. In this paper, we consider the case $g = f$. In general, g should be defined to best capture the nature of the biomarker trajectory and its relationship to the hazard of the event.

3 | IMPLICATIONS OF USING STUDY ENTRY AS THE TIME ORIGIN FOR COX MODELS WITH TIME-VARYING PREDICTORS

When an appropriate time origin precedes study entry and the covariate of interest is not time varying, Korn et al[8] provided sufficient conditions under which fitting a model that uses study entry as the time origin preserves the desired estimated

regression coefficient corresponding to the fixed covariate. They assumed that while the true survival model is given by

$$\lambda_{U_i}(u|x_i, E_i) = \lambda_{0U}(u)e^{\beta x_i},$$

where x_i is a fixed covariate, the data are fit instead to the model,

$$\lambda_{T_i}(t|x_i, E_i) = \lambda_{0T}(t)e^{\xi x_i + \gamma E_i}, \tag{4}$$

which uses study entry as the time origin. It is clear that $\xi = \beta$ provided that $\lambda_{0U}(u) = c\exp\{\gamma u\}$ and $\lambda_{0T}(t) = c\exp\{\gamma t\}$, for some positive constant c. Under these conditions, $\hat{\xi}$ based on fitting the Cox model specified in (4) will be a consistent estimator of β.

We now extend the results of Korn et al to the case of a continuous time-varying covariate, $w_i(u)$, and provide conditions under which using study entry as the time origin and/or treating a time-varying covariate as fixed will preserve the Cox regression parameter of interest. Suppose that the true survival model is

$$\lambda_{U_i}(u|\overline{w_i}(\cdot), E_i) = \lambda_{0U}(u)e^{\beta w_i(u)}. \tag{5}$$

We assume further that E and U are continuous random variables and that the time-varying biomarker $w(u)$ is continuous and defined on the support of U. Without loss of generality, we assume that the support of E is contained in the support of U. In addition, we assume that $\beta \neq 0$, and the study entry times, E_i, and biomarkers, $w_i(u)$, are not constant across individuals.

We consider 3 possible models that might be fit to these data. First, we consider the model that uses study entry as the time origin and adjusts for study entry time as a covariate

$$\lambda_{T_i}(t|\overline{w_i}(\cdot), E_i) = \lambda_{0T}(t)e^{\xi w_i(t+E_i) + \gamma E_i}, \tag{6}$$

where $\bar{w}(t)$ denotes the history of w up to and including time t. In Theorem 1 part (a) below, we derive conditions on the true baseline hazard function under which the coefficient of $w_i(t + E_i)$ is equal to β. Assuming only that $\lambda_{0U}(u) = ce^{\gamma u}$, the Cox partial likelihood derived from (6) can be used for estimation of β. Note that this extends the fixed covariate case (Korn et al[8]).

Next, we consider models that treat the time-varying predictor as fixed, using either onset or study entry as the origin, and derive conditions under which we obtain a consistent estimator of β. If the value of the time-varying covariate is known only at study entry, it can be included in a Cox model as a fixed covariate in a second model with onset as the time origin:

$$\lambda_{U_i}(u|\overline{w_i}(\cdot), E_i) = \lambda_{1U}(u)e^{\alpha w_i(E_i)}. \tag{7}$$

Note that estimation of this model must adjust for delayed entry (left truncation). If time from onset to study entry is known, we can include E_i as a covariate in a third model using study entry as the time origin:

$$\lambda_{T_i}(t|\overline{w_i}(\cdot), E_i) = \lambda_{2T}(t)e^{\zeta w_i(E_i) + \eta E_i}. \tag{8}$$

The case where there is no covariate adjustment for study entry time is included in this model when $\eta = 0$. Alternatively, the time of study entry could be adjusted for through stratification or through covariate adjustment by a nonlinear function of E_i, which we do not consider here. In Theorem 1 parts (b) and (c), we derive conditions under which the coefficients of $w_i(E_i)$ in (7) and (8) are equal to β.

Theorem 1. *Suppose the true hazard regression model is given by (5).*

(a) (Korn et al[8] extension) Consider the model in (6): $\beta = \xi \iff \lambda_{0U}(u) = ce^{\gamma u}$ and $\lambda_{0T}(t) = ce^{\gamma t}$, for some positive constant c. In practice, if $\lambda_{0U}(u) = ce^{\gamma u}$, (6) can be used to estimate β without specifying that $\lambda_{0T}(t) = ce^{\gamma t}$.

(b) Consider the model in (7): $\beta = \alpha \iff w_i(u) = w_i(E_i) + g(u)$, for some continuous function g that is zero on the support of E. In practice, if $w_i(u) = w_i(E_i) + g(u)$, (7) can be used to estimate β in an analysis that adjusts for left truncation.

(c) Consider the model in (8): $\beta = \zeta \iff w_i(t + E_i) - w_i(E_i) = k(t)$, $\lambda_{0U}(u) = ce^{\eta u}$ and $\lambda_{2T}(t) = ce^{\beta k(t) + \eta t}$, for $c > 0$. In practice, under these conditions on λ_{0U} and $w_i(t + E_i) - w_i(t)$, (8) can be used to estimate β even if $\lambda_{2T}(t) \neq ce^{\beta k(t) + \eta t}$.

The proof for Theorem 1 is given in the Appendix. Theorem 1 provides guidelines for implementing Cox regression with a time-varying covariate using study entry as the origin or when the time-varying covariate is measured only at study entry. Theorem 1 part (a) shows that study entry can be used as the origin if the covariate process is fully observed

or if its functional form is known, as long as the baseline hazard is of exponential functional form. These conditions are reasonable in some settings: In many instances, a covariate process can be modeled using subject-level knowledge,[23] and constant hazard functions have been empirically shown to model incidence in many epidemiological studies,[24] and exponential-form (Gompertz) baseline hazards are widely used to model mortality.

If the time-varying covariate is observed only at study entry and its functional form is known, then the covariate process can be treated as fixed in an analysis from onset, as long as the covariate function, $w_i(u)$, satisfies the condition in Theorem 1 part (b). An example of such a function is $w_i(u) = a_i + m \cdot (h(u) - h(L))1\{u > L\}$, where L is the upper bound of the support of E and h is continuous; this is a time-varying continuous covariate function that is constant on the support of E but can vary for values of u outside of the support of E. Since we are mainly concerned with $w(u)$ on the support of E, Theorem 1 part (b) indicates that treating a time-varying (ie, nonconstant) biomarker as fixed using onset as the origin is not valid.

Theorem 1 part (c) shows that if the time-varying covariate is observed only at study entry, an analysis using study entry as the origin is valid only for special cases of the baseline hazard and covariate process. As discussed above, the baseline hazard assumptions are realistic in some settings. One example of a time-varying covariate that satisfies the condition in Theorem 1 part (c) is $w_i(t) = a_i + a_1 t$, a linear function with a subject-specific intercept. Thus, in some instances, it is reasonable to treat time-varying covariates as fixed in an analysis from study entry.

In Theorem 1 parts (a) and (c), when the true baseline hazard function is of an exponential form, the results of the theorem are contingent upon the incorporation of study entry time, E, into the model as a linear covariate. It can be shown, using arguments analogous to the proofs of Theorem 1 parts (a) and (c), that adjustment of study entry time through a transformation of E results in biased estimation; however, the extent of the bias is unknown. However, correct specification of E may be important if entry time is highly correlated with the biomarker. For the case of a fixed covariate, Pencina et al[11] illustrated through simulations that in an analysis from study entry, adjusting for E using a quadratic term and also through stratified methods can result in negligible bias.

The focus of this paper is the impact of the time-varying covariate $w_i(u) = f(S(u)) + a_i + b_i u$, as defined in (1), on the hazard function through the Cox model defined in (5). Theorem 1 part (a) allows for valid estimation of β using study entry as the time origin provided that f, a_i, and b_i are known, E_i is included in the model as a covariate, and $\lambda_{0U}(u) = c\exp\{\gamma u\}$. If the covariate process is known only at study entry, then it is not valid to treat the time-varying covariate as fixed in an analysis from onset because $w_i(u) - w_i(E_i) = f(S(u)) - f(S(E_i)) + b_i(u - E_i)$ is a function that depends on i and violates the condition of part (b) of the Theorem.

The implications of Theorem 1 part (c) differ for the 2 examples of f introduced in Section 2. First, consider the absolute error model with $w_i(u) = S(u) + a_i + b_i u$, where $S(u)$ is the sigmoid in (2). Theorem 1 part (c) establishes that using study entry as the origin and treating the time-varying covariate as fixed will yield biased estimates of β. To see this, Theorem 1 part (c) requires that $w_i(t+E_i) - w_i(E_i)$ is independent of i; however, in this case $w_i(t+E_i) - w_i(E_i) = S(t+E_i) - S(E_i) + b_i t$ is not independent of i. This is the case even if $b_i \equiv 0$. Alternatively, for the logit error model with $w_i(u) = \text{logit}[S(u)/M] + a_i + b_i u$, if $b_i \equiv b$, using study entry as the origin and treating the time-varying biomarker as fixed is valid, provided that the baseline hazard is of an exponential form, since $w_i(t + E_i) - w_i(E_i) = (\lambda + b)t$ is independent of i. If $b_i \neq b$, this condition is not satisfied.

Note that the results of Theorem 1, parts (a) and (c), apply even when $\beta = 0$. In contrast, Theorem 1 part (b) does not apply to this case. If $\beta = 0$, then the true survival model is

$$\lambda_{U_i}(u|\overline{w_i}(\cdot), E_i) = \lambda_{0U}(u).$$

If we fit the model in (7), it is easy to see that $\alpha = 0 \iff \lambda_{0U}(u) = \lambda_{1U}(u)$. In practice, this implies that fitting (7) when $\beta = 0$ will yield $\hat{\alpha} = 0$ without restrictions on $w_i(u)$.

4 | METHODS

On the basis of our analysis in Section 3, it is not advisable to treat time-varying biomarkers measured only at study entry as fixed covariates for the time origin of disease onset. And for the time origin of study entry, this can be done only under strong assumptions on the baseline hazard function and the covariate function. Thus, lacking full observation of the biomarkers over time, we require an approach that will allow us to estimate the full biomarker trajectory based on cross-sectional observation of it and, importantly, does not require restrictive assumptions on the baseline hazard function of the functional form of the covariate process.

Assuming that the time-varying biomarker follows the general error model in (1) and the true hazard can be modeled as in (3) for $g = f$, we provide methods for estimating the regression parameter β when the fully observed time-varying biomarker is measured only at study entry. We note that the components of the error, a_i and b_i, cannot be simultaneously estimated when only a single biomarker measurement is available for each subject. We thus consider the cases $a_i \equiv 0$ and $b_i \equiv 0$ separately.

4.1 | Case: $b_i \equiv 0$

The covariate process in this case is defined as $f(Z_i(u)) = f(S(u; \lambda, T_0)) + a_i$, with individual-level deviations that are fixed over time. We assume that $E(a_i) = 0$. We assume that the value of M, which represents the maximum value of the sigmoidal predictor in (2), is known. To estimate the regression parameter β, we first estimate the parameters λ and T_0 of the population-level sigmoid $S(u; \lambda, T_0)$ using least squares to minimize

$$\sum_i \left[f(Z_i(E_i)) - f(S(E_i; \lambda, T_0)) \right]^2$$

with respect to λ and T_0. We then estimate the a_i as $\widehat{a_i} = f(Z_i(E_i)) - f(S(E_i; \hat{\lambda}, \widehat{T_0}))$ and the time-varying covariate process as $\widehat{f(Z_i(u))} = f(S(u; \hat{\lambda}, \widehat{T_0})) + \widehat{a_i}$. The true hazard function $\lambda_{U_i}(u | f(\overline{Z}_i(\cdot)))$ can then be approximated as

$$\lambda_{U_i}\left(u \middle| \widehat{f(Z_i(\cdot))}\right) = \lambda_{0U}(u) \exp\left\{\beta \cdot \widehat{f(Z_i(u))}\right\} = \lambda_{0U}(u) \exp\left\{\beta \cdot \left[f(S(u; \hat{\lambda}, \widehat{T_0})) + \widehat{a_i}\right]\right\} = \widetilde{\lambda_{0U}}(u) \exp\{\beta \cdot \hat{a}_i\}. \tag{9}$$

The model in (9) shows that the desired regression parameter β can be estimated by fitting a Cox model with $\hat{a}_i$ included as a fixed covariate.

Note that if $f(Z_i(u))$ is linear in u, as is the case with the sigmoid-based logit error model, where

$$f(Z_i(u)) = \text{logit}[Z_i(u)/M] = \lambda(u - T_0) + a_i,$$

then the model based on $f(\widehat{Z_i(t + E_i)}))$ is equal to that based on $f(Z_i(E_i))$. In this case,

$$f(\widehat{Z_i(t + E_i)}) = \hat{\lambda}(t + E_i - \hat{T}_0) + \hat{a}_i = \hat{\lambda}t + \widehat{f(Z_i(E_i))} = \hat{\lambda}t + f(Z_i(E_i)),$$

by the definition of $\hat{a}_i$. Thus, fitting a Cox model from study entry with $\widehat{f(Z_i(t + E_i))}$ as a time-varying covariate is equivalent to fitting the model from entry with the fixed covariate $f(Z_i(E_i))$, since the population-level $\hat{\lambda}t$ term will be absorbed into the baseline hazard function and will not affect estimation.

4.2 | Case: $a_i \equiv 0$

The covariate process in this case is defined as $f(Z_i(u)) = f(S(u); \lambda, T_0) + b_i u$, with time-varying individual-level deviations. At each E_i, it follows that

$$\frac{f(Z_i(E_i))}{E_i} = \frac{f(S(E_i; \lambda, T_0))}{E_i} + b_i, \tag{10}$$

and thus, λ and T_0 can be estimated via least squares by minimizing

$$\sum_i \left[\frac{f(Z_i(E_i))}{E_i} - \frac{f(S(E_i; \lambda, T_0))}{E_i} \right]^2$$

with respect to λ and T_0. We then estimate b_i as

$$\widehat{b_i} = \frac{f(Z_i(E_i))}{E_i} - \frac{f(S(E_i; \hat{\lambda}, \widehat{T_0}))}{E_i}.$$

Note that estimation of b_i uses only those entry times E_i that are nonzero; in some settings, it is possible that a time origin, such as disease onset, may coincide with study entry for some subjects. We estimate the time-varying covariate process as $\widehat{f(Z_i(u))} = f(S(u; \hat{\lambda}, \widehat{T_0})) + \widehat{b_i} u$. The true hazard function $\lambda_{U_i}(u | f(\overline{Z}_i(\cdot)))$ is approximated as

$$\lambda_{U_i}\left(u|\widehat{f(Z_i(\cdot))}\right) = \lambda_{0U}(u)\exp\left\{\beta\cdot\widehat{f(Z_i(u))}\right\} = \lambda_{0U}(u)\exp\left\{\beta\cdot\left[f(S(u;\hat{\lambda},\widehat{T_0}))+\hat{b}_i u\right]\right\} = \widetilde{\lambda_{0U}}(u)\exp\{\beta\cdot\hat{b}_i u\}. \quad (11)$$

The regression parameter β can be estimated using Cox regression with the time-varying covariate $\hat{b}_i u$.

4.3 | Adjusting sigmoid for disease severity

In the AD setting, individual sigmoidal trajectories might be explained by individual-level covariates (eg, APOE4 allele status, gender) and not just simply by unexplained individual deviations. We propose to incorporate this into the sigmoidal model by allowing the lateral shift of the population-level sigmoid, given by T_0, to be a function of covariates. Let w_i denotes a vector of individual-specific covariates reflecting disease severity. We model the individual-specific lateral shifts as $T_{0,i} = \gamma'\mathbf{w}_i$, for a vector of coefficients γ. Then the sigmoid is redefined as

$$S(u,\mathbf{w}_i;\lambda,\gamma) = \frac{M}{1+e^{-\lambda(u-\gamma'\mathbf{w}_i)}},$$

additional individual deviations are modeled as $f(Z_i(u,\mathbf{w}_i)) = f(S(u,\mathbf{w}_i;\lambda,\gamma)) + a_i + b_i u$, and the true hazard model is assumed to be

$$\lambda_{U_i}(u|\, f(\overline{Z}_i(\cdot))) = \lambda_{0U}(u)\exp\{\beta f(Z_i(u,\mathbf{w}_i;\lambda,\gamma))\}. \quad (12)$$

This can be implemented for any f. In the case of the logit error model (1),

$$\begin{aligned}\text{logit}\left[\frac{Z_i(u,\mathbf{w}_i)}{M}\right] &= \text{logit}\left[\frac{S(u,\mathbf{w}_i);\lambda,\gamma}{M}\right] + a_i + b_i u\\ &= \lambda(u-\gamma'\mathbf{w}_i) + a_i + b_i u,\end{aligned}$$

so that the individual-specific lateral shifts are given by $\gamma'\mathbf{w}_i$. The model in (12) then simplifies as

$$\begin{aligned}\lambda_{U_i}(u|\text{logit}(\overline{Z}_i(\cdot)/M)) &= \lambda_{0U}(u)\exp\left\{\beta\left[\text{logit}(Z_i(u,\mathbf{w}_i;\lambda,\gamma)/M)\right]\right\}\\ &= \lambda_{0U}(u)\exp\left\{\beta\left[\lambda(u-\gamma'\mathbf{w}_i)+a_i+b_i u\right]\right\}\\ &= \widetilde{\lambda_{0U}}(u)\exp\left\{\beta[a_i+b_i u-\lambda\gamma'\mathbf{w}_i]\right\}.\end{aligned}$$

As discussed above, both error components a_i and b_i cannot be estimated simultaneously. The estimation procedure is analogous to that described in Sections 4.1 and 4.2 and differs only regarding the parameters of disease severity adjusted population-level sigmoid; least squares is used to estimate λ and γ. For example, for the case $b_i \equiv 0$, we use least squares to minimize

$$\sum_i\left[f(Z_i(E_i)) - f(S(E_i,\mathbf{w}_i;\lambda,\gamma))\right]^2$$

with respect to λ and γ.

5 | SIMULATIONS

We simulated left-truncated time-to-event data such that the time from onset to event, U_i, satisfies the hazard model in (3) and the transformed time-varying predictor, $f(Z_i(u))$, follows the logit error model in (1) with sigmoidal S in (2), upper bound M, and $f(x) = \text{logit}(x/M)$. Study entry times, E_i, were generated independently of the survival times, U_i. The time-varying covariates were available only from the times of study entry. The data, $(E_i, U_i, f(Z_i(E_i)))$, were generated to achieve specified truncation probabilities, denoted $\pi_{\text{trunc}} = P(U_i < E_i)$. Parameters of the covariate process were chosen to reflect Pittsburgh compound B–positron emission tomography measurements of the AD biomarker, Aβ; hazard regression parameters were chosen so that time-to-AD diagnosis values align with those observed in AD studies. The data were generated in the following steps:

(1) Select $\beta \in \{0, 0.25, 0.5, 1, 1.5\}$, $n \in \{100, 250\}$, $M = 3$, $\lambda \in \{0.1, 0.25, 0.4\}$, $T_0 = 40$, $\sigma_a \in \{0.5, 1\}$, $\sigma_b \in \{0.01, 0.05\}$.
(2) Generate error components, $a_i \sim N(0, \sigma_a^2)$ and set $b_i = 0$, or generate $b_i \sim N(0, \sigma_b^2)$ and set $a_i = 0$.

(3) Generate $V_i \sim \text{Unif}(0,1)$ and solve for U_i in

$$V_i = \exp\left\{-\int_0^{U_i} \lambda_{U_i}(w)dw\right\} = \exp\left\{-\int_0^{U_i} \lambda_{0U}(w)\exp\{\beta \cdot [f(S(w)) + a_i + b_i w]\}dw\right\}. \quad (13)$$

This can be easily done when the integral in (13) can be calculated in closed form.[25] This is the case under the logit error model (1) where $f(S(u)) = \lambda(u - T_0) + a_i + b_i u$. We considered 3 separate baseline hazard functions, $\lambda_{0U}(u)$, defined below: constant, exponential form, normal density kernel. The constant and exponential-form baseline hazard functions appear in Theorem 1 and were chosen to confirm our theoretical results. The normal density kernel baseline hazard function was chosen because it permits a closed form solution for U_i and violates the requirements of the theorem and thus also confirms our theoretical results.

Baseline hazard functions:

(*a*) (Constant baseline hazard) $\lambda_{0U}(u) = c$, where $c = 0.0154, 0.01, 0.004, 0.001, 0.001$ corresponding to $\beta = 0, 0.25, 0.5, 1, 1.5$.

(*b*) (Exponential-form baseline hazard) $\lambda_{0U}(u) = ce^{\psi u}$, where $c > 0$ and $\psi \neq 0$ with $c = 7 \times 10^{-8}$ and $\psi = 0.2$.

(*c*) (Normal density kernel baseline hazard) $\lambda_{0U}(u) = m\exp\left\{-c(u-a)^2\right\}$, where $(c, a, m) = (0.001, 60, 1)$ for the fixed-deviations case ($b_i \equiv 0$) and $(c, a, m) = (0.1, 70, 100)$ for the time-varying deviations case ($a_i \equiv 0$). The hazard function can be expressed as follows:

$$\lambda_{U_i}(u) = m\exp\left\{-c(u-a)^2 + \beta[\lambda(u - T_0) + a_i + b_i u]\right\} = \phi(u; \mu_i, \sigma^2) \cdot K_i,$$

where $\phi(u; \mu_i, \sigma^2)$ is the normal density with mean $\mu_i = a + \beta(\lambda + b_i)/(2c)$ and variance $\sigma^2 = 1/(2c)$ and

$$K_i = m\exp\left\{\beta\left[a(\lambda + b_i) + a_i + \frac{\beta(\lambda + b_i)^2}{4c} - \lambda T_0\right]\right\}\sqrt{2\pi\sigma^2}.$$

(4) Generate study entry times $E_i \sim \text{Uniform}(0, R_{\text{trunc}})$, where R_{trunc} is chosen to achieve $P(U_i < E_i) = \pi_{\text{trunc}} = 0.2, 0.4$ and is a function of the parameters of the distribution of U. Values of R_{trunc} used can be found in Table S1.

(5) Retain the first n observations that satisfy $E_i < U_i$; ie, incorporate delayed entry to this study (left truncation).

(6) Calculate $f(Z_i(E_i)) = f(S(E_i)) + a_i + b_i E_i = \lambda(E_i - T_0) + a_i + b_i E_i$, where either a_i or b_i is zero.

(7) Estimate λ and T_0 using least squares as in Sections 4.1 and 4.2.

(8) Estimate a_i or b_i using $\hat{\lambda}$ and $\widehat{T_0}$ as in Sections 4.1 and 4.2.

(9) Define $\widehat{f(Z_i(u))} = \hat{\lambda}(u - \widehat{T_0}) + \hat{a}_i$ or $\widehat{f(Z_i(u))} = \hat{\lambda}(u - \widehat{T_0}) + \hat{b}_i u$, as appropriate.

(10) Fit the following models, adjusting for left truncation when the origin is onset:

$$\lambda_{U_i}(u|f(\overline{Z}_i(\cdot)), E_i) = \lambda_{0U}(u)e^{\beta_1 f(Z_i(u))} \quad (A)$$

$$\lambda_{U_i}(u|f(\overline{Z}_i(\cdot)), E_i) = \lambda_{0U}(u)e^{\beta_2 \widehat{f(Z_i(u))}} \quad (A')$$

$$\lambda_{U_i}(u|f(Z_i(E_i)), E_i) = \lambda_{1U}(u)e^{\alpha f(Z_i(E_i))} \quad (B)$$

$$\lambda_{T_i}(t|f(Z_i(E_i)), E_i) = \lambda_{1T}(t)e^{\zeta_1 f(Z_i(E_i))} \quad (C)$$

$$\lambda_{T_i}(t|f(Z_i(E_i)), E_i) = \lambda_{2T}(t)e^{\zeta_2 f(Z_i(E_i)) + \eta E_i} \quad (D)$$

$$\lambda_{T_i}(t|f(\overline{Z}_i(\cdot)), E_i) = \lambda_{3T}(t)e^{\xi_1 f(\widehat{Z}_i(t+E_i))} \quad (E)$$

$$\lambda_{T_i}(t|f(\overline{Z}_i(\cdot)), E_i) = \lambda_{4T}(t)e^{\xi_2 f(\widehat{Z}_i(t+E_i)) + \gamma E_i}. \quad (F)$$

Model (A) is the true survival model, where $f(Z_i(u)) = \lambda(u - T_0) + a_i + b_i u$. Model (A) corresponds to the proposed method in Section 4 where $\widehat{f(Z_i(u))}$ is defined in step 9. Model (B) corresponds to Theorem 1 part (b), and models (C) and (D) correspond to Theorem 1 part (c), where $f(Z_i(E_i))$ is from step 3. Models (E) and (F) correspond to Theorem 1 part (a), where $\widehat{f(Z_i(t+E_i))} = \hat{\lambda}(t + E_i - \widehat{T_0}) + \hat{a}_i + \hat{b}_i(t + E_i)$. Given the linear form of $f(Z_i(u))$ with $b_i \equiv 0$, estimation via models (E) and (F) is equivalent to that via models (C) and (D), respectively.

(11) Repeat 5000 times.

The statistical package R[26] version 3.1.2 was used for all data analyses. Nonlinear least squares minimization was executed using the `nls` function in the `stats` package[27] with convergence parameters set to `maxiter=1e4` and `minFactor=1e-10` and with the default tolerance `tol = 1e-5`. Hazard regression was implemented using the function `coxph` in the `survival` package.[28] Cox models with time-varying predictors were fit using the time-transform option of `coxph` or using the counting process style input. R code is available in the Supporting Information.

The simulation results are given in Table 1 for varying values of β, $\lambda = 0.25$, $\sigma_a = 0.5$ and $\sigma_b = 0.05$, and $\pi_{\text{trunc}} = 0.2$ for the fixed (a_i) and time-varying (b_i) individual-level deviations models, respectively. Means and standard deviations of simulated estimates are presented. Type I error and power were taken to be the proportion of simulations, which rejected the test statistic with probability less than 0.05 for $\beta = 0$ and $\beta > 0$, respectively. Type I error and power corresponding to $\beta = 0.5$ are given in Tables 2 and 3, respectively, for the parameter values used in Table 1. For the case of fixed deviations (a_i) and a normal density kernel baseline hazard, the parameters originally chosen for the baseline hazard function to align with the AD setting result in a baseline hazard function that closely resembles an exponential and thus do not serve the purpose of illustrating results under deviations from the exponential. Thus, the parameters of the normal density kernel baseline hazard were then chosen to illustrate this contrast, $(c, a, m) = (0.05, 30, 1)$, and the results for this parameterization of the normal kernel density baseline hazard function in the fixed deviations case are given in Table 1.

The simulation results in column 1 of Table 1 confirm that fitting the true survival model, in (A), gives consistent estimates of β. The second column demonstrates that the proposed method described in Section 4, which corresponds to model (A), performs nearly as well as fitting the true model in column 1 for all baseline hazard distributions considered (constant, exponential form, and normal density kernel), with slightly higher standard errors, because of estimation of the covariate process.

When the covariate process follows the general error model in (1), Theorem 1 part (b) guarantees that treating a time-varying biomarker as fixed at its value at study entry will always produce biased estimates of β if the true origin is used, for nonzero values of β. This analysis corresponds to fitting model (B), and the results are listed in column 3 of Table 1 for $\beta \neq 0$ for both fixed and time-varying deviations and for all 3 baseline hazards considered and are as expected. When $\beta = 0$, we expect the estimated regression coefficient to be zero (see discussion following Theorem 1); this is seen in the table, as well.

Theorem 1 part (c) provides conditions for when using an incorrect origin of study entry and treating the time-varying predictors as fixed results in consistent estimates of β. For the fixed deviation models (ie, a_i), the discussion following Theorem 1 ensures that inference is valid if study entry is used as the time origin and the time-varying biomarker is treated as fixed (with or without covariate adjustment for study entry time), as in models (C) and (D), provided that the baseline hazard is constant. Similarly, if the baseline hazard follows an exponential form and both the value of the time-varying biomarker at study entry and study entry time are included in the model as covariates, as in (D), estimates of β are unbiased. For the time-varying deviations models (b_i), Theorem 1 part (c) guarantees that estimates of β will be biased if study entry is used as the time origin and the time-varying biomarker is treated as a fixed covariate (with or without covariate adjustment for study entry time), as in (C) and (D). These theoretical findings are illustrated through our simulation results in columns 4 and 5.

Theorem 1 part (a) corresponds to models (E) and (F). It asserts that for a constant baseline hazard, the model that uses study entry as the origin and includes $f(Z_i(t + E_i))$ as a time-varying covariate, with or without covariate adjustment for study entry time, will give consistent estimates of β for fixed or time-varying deviations. For a baseline hazard that is of an exponential form, fitting the model that uses study entry as the origin, including both $f(Z_i(t+E_i))$ and study entry time as covariates, will give consistent estimates of β. These results are supported through our simulations when $f(\widehat{Z_i(t + E_i)})$ is used to approximate $f(Z_i(t + E_i))$. When the baseline hazard does not follow an exponential form, Theorem 1 part (a) guarantees that estimates of β will be biased if a model is fit using study entry as the origin, including both $f(Z_i(t + E_i))$ and study entry time as covariates. For the case of the normal density kernel baseline hazard and fixed deviations, this is confirmed through our simulation results, when $f(\widehat{Z_i(t + E_i)})$ is used to approximate $f(Z_i(t + E_i))$, with greater bias corresponding to larger values of β. For the case of the normal density kernel baseline hazard and time-varying deviations, our simulation results indicate that the bias increases with β.

For the models with fixed deviations (a_i), notice that fitting models (C) and (E) give identical results. The same is true for models (D) and (F). As noted at the end of Section 4.1, this is expected when the form of the covariate model is linear in u, as is the case with the sigmoid-based logit error model.

Table 2 shows that the type I error is within 0.06 for all analyses that use onset as the time origin. For the analyses that use study entry as the origin, the type I error is less than 0.06 for baseline hazard functions that are constant, or follow

TABLE 1 Bias and analytical standard errors of Cox parameter estimates for the logit error model for specified covariate model and baseline hazard function using $n = 250$ and 5000 iterations

		Onset origin			Study entry origin			
Covariate model	β	**A:** $f(Z_i(u))$	**A′:** $f(\widehat{Z_i(u)})$	**B:** $f(Z_i(E_i))$	**C:** $f(Z_i(E_i))$	**D:** $f(Z_i(E_i)) + E_i$	**E:** $f(\widehat{Z_i(t+E_i)})$	**F:** $f(\widehat{Z_i(t+E_i)}) + E_i$
Fixed deviations, a_i								
Normal density–based	0	0.001 (0.131)	0.001 (0.132)	0 (0.034)	1.306 (0.108)	0.002 (0.169)	1.306 (0.108)	0.002 (0.169)
hazard	0.25	0.003 (0.132)	0.003 (0.132)	−0.234 (0.033)	1.188 (0.104)	0.059 (0.168)	1.188 (0.104)	0.059 (0.168)
	0.5	0.004 (0.135)	0.004 (0.135)	−0.471 (0.032)	1.033 (0.117)	0.121 (0.174)	1.033 (0.117)	0.121 (0.174)
	1	0.006 (0.144)	0.002 (0.144)	−0.951 (0.031)	0.770 (0.158)	0.191 (0.174)	0.770 (0.158)	0.191 (0.174)
	1.5	0.008 (0.158)	−0.005 (0.157)	−1.438 (0.030)	0.436 (0.260)	0.162 (0.176)	0.436 (0.260)	0.162 (0.176)
Constant hazard	0	−0.002 (0.129)	−0.002 (0.129)	0 (0.031)	0 (0.030)	−0.002 (0.130)	0 (0.030)	−0.002 (0.130)
	0.25	0.002 (0.131)	0.002 (0.132)	−0.248 (0.014)	0.001 (0.019)	0.001 (0.132)	0.001 (0.019)	0.001 (0.132)
	0.5	0.005 (0.135)	0.003 (0.135)	−0.495 (0.014)	0.002 (0.030)	0.004 (0.136)	0.002 (0.030)	0.004 (0.136)
	1	0.008 (0.145)	0.003 (0.145)	−0.990 (0.015)	0.004 (0.055)	0.009 (0.146)	0.004 (0.055)	0.009 (0.146)
	1.5	0.009 (0.156)	−0.003 (0.157)	−1.484 (0.016)	0.006 (0.082)	0.010 (0.157)	0.006 (0.082)	0.010 (0.157)
Exponential-form	0	0 (0.131)	0.001 (0.131)	0 (0.012)	0.729 (0.041)	0 (0.135)	0.729 (0.041)	0 (0.135)
hazard	0.25	0.002 (0.131)	0.003 (0.131)	−0.248 (0.014)	0.708 (0.053)	0.002 (0.135)	0.708 (0.053)	0.002 (0.135)
	0.5	0.005 (0.135)	0.004 (0.135)	−0.495 (0.014)	0.688 (0.065)	0.004 (0.138)	0.688 (0.065)	0.004 (0.138)
	1	0.008 (0.145)	0.006 (0.145)	−0.989 (0.016)	0.652 (0.089)	0.009 (0.148)	0.652 (0.089)	0.009 (0.148)
	1.5	0.009 (0.155)	0.003 (0.156)	−1.483 (0.017)	0.615 (0.112)	0.011 (0.157)	0.615 (0.112)	0.011 (0.157)
Time-varying deviations, b_i								
Normal density–based	0	0 (0.021)	0 (0.021)	−0.001 (0.013)	0.556 (0.052)	0 (0.044)	0.289 (0.020)	0 (0.026)
hazard	0.25	0.002 (0.025)	0.001 (0.026)	−0.206 (0.014)	0.360 (0.057)	0.122 (0.049)	0.063 (0.021)	0.014 (0.029)
	0.5	0.003 (0.034)	0 (0.035)	−0.434 (0.014)	0.156 (0.062)	0.133 (0.061)	−0.166 (0.022)	0.021 (0.040)
	1	0.005 (0.056)	−0.022 (0.078)	−0.917 (0.016)	−0.267 (0.071)	−0.118 (0.087)	−0.632 (0.023)	−0.011 (0.069)
	1.5	0.006 (0.083)	−0.065 (0.150)	−1.411 (0.016)	−0.696 (0.078)	−0.523 (0.108)	−1.102 (0.024)	−0.165 (0.133)
Constant hazard	0	0 (0.015)	0 (0.015)	0 (0.029)	0.001 (0.028)	−0.001 (0.081)	0 (0.013)	0 (0.015)
	0.25	0.001 (0.025)	0 (0.025)	−0.204 (0.013)	−0.014 (0.019)	0.155 (0.046)	0.001 (0.018)	0.001 (0.023)
	0.5	0.002 (0.035)	−0.002 (0.036)	−0.436 (0.014)	−0.113 (0.028)	0.244 (0.061)	−0.001 (0.029)	0.001 (0.032)
	1	0.004 (0.057)	−0.021 (0.078)	−0.921 (0.015)	−0.474 (0.044)	0.222 (0.083)	−0.010 (0.055)	−0.006 (0.057)
	1.5	0.007 (0.083)	−0.052 (0.137)	−1.410 (0.017)	−0.890 (0.055)	0.029 (0.104)	−0.026 (0.087)	−0.019 (0.087)
Exponential-form	0	0 (0.018)	0 (0.018)	0 (0.012)	0.372 (0.031)	0 (0.032)	0.221 (0.017)	0 (0.019)
hazard	0.25	0.001 (0.025)	0.001 (0.025)	−0.206 (0.013)	0.364 (0.047)	0.109 (0.042)	0.103 (0.022)	0.001 (0.024)
	0.5	0.002 (0.035)	−0.001 (0.036)	−0.433 (0.015)	0.348 (0.060)	0.152 (0.056)	−0.012 (0.027)	0.001 (0.034)
	1	0.006 (0.057)	−0.016 (0.073)	−0.913 (0.016)	0.181 (0.067)	0.088 (0.080)	−0.237 (0.040)	−0.001 (0.056)
	1.5	0.007 (0.080)	−0.049 (0.132)	−1.404 (0.018)	−0.195 (0.078)	−0.117 (0.099)	−0.464 (0.055)	−0.011 (0.085)

TABLE 2 Type I error for the logit error model for specified covariate model and baseline hazard function using $n = 250$ and 5000 iterations

	Onset origin			Study entry origin			
Covariate model	**A:** $f(Z_i(u))$	**A′:** $f(\widehat{Z_i(u)})$	**B:** $f(Z_i(E_i))$	**C:** $f(Z_i(E_i))$	**D:** $f(Z_i(E_i)) + E_i$	**E:** $f(\widehat{Z_i(t+E_i)})$	**F:** $f(\widehat{Z_i(t+E_i)}) + E_i$
Fixed deviations, a_i							
Normal density–based hazard	0.051	0.052	0.054	1	0.124	1	0.124
Constant hazard	0.048	0.048	0.052	0.053	0.048	0.053	0.048
Exponential-form hazard	0.054	0.054	0.048	1	0.055	1	0.055
Time-varying deviations, b_i							
Normal density–based hazard	0.061	0.060	0.050	1	0.060	1	0.070
Constant hazard	0.056	0.056	0.044	0.049	0.053	0.050	0.058
Exponential-form hazard	0.050	0.052	0.052	1	0.052	1	0.051

TABLE 3 Power corresponding to $\beta = 0.5$ for the logit error model for specified covariate model and baseline hazard function using $n = 250$ and 5000 iterations

	Onset origin			Study entry origin			
Covariate model	**A:** $f(Z_i(u))$	**A′:** $f(\widehat{Z_i(u)})$	**B:** $f(Z_i(E_i))$	**C:** $f(Z_i(E_i))$	**D:** $f(Z_i(E_i)) + E_i$	**E:** $f(\widehat{Z_i(t+E_i)})$	**F:** $f(\widehat{Z_i(t+E_i)}) + E_i$
Fixed deviations, a_i							
Normal density–based hazard	0.97	0.97	0.15	1	0.98	1	0.98
Constant hazard	0.97	0.97	0.07	1	0.97	1	0.97
Exponential-form hazard	0.97	0.97	0.07	1	0.96	1	0.96
Time-varying deviations, b_i							
Normal density–based hazard	1	1	0.99	1	1	1	1
Constant hazard	1	1	0.99	1	1	1	1
Exponential-form hazard	1	1	0.99	1	1	1	1

an exponential form, provided that study entry time is included as a covariate (models (D) and (E); the type I error is inflated for all other analyses from study entry. This follows from Theorem 1 parts (a) and (c), which hold for $\beta = 0$; specifically, when $\beta = 0$ and the conditions of Theorem 1 parts (a) and (c) are not met, then the estimated regression coefficients are expected to be nonzero. Table 3 shows that the null hypothesis test is well powered, under the simulation parameters specified in Table 1, for all analyses except for when onset is used as the time origin and the time-varying covariate is treated as fixed (model (B). In this case, the population-level covariate function is absorbed by the baseline hazard function, and the impact of the covariate is due to the individual-specific deviations, as in (9); it is likely that $\beta \cdot a_i$ is close to zero for all subjects so that the likelihood of rejecting the null is also small.

In additional simulations (not shown), variations in the truncation probability, π_{trunc}, had little effect on bias, type I error, and power. Larger values of σ_a and σ_b, the standard deviations of a_i and b_i, were associated with smaller standard errors and greater power, but had little effect on type I error. Varying λ, the acceleration parameter of the sigmoid $S(u)$, gave comparable results with respect to bias, type I error, and power; however, greater values of β were required to see bias in the case of a normal density–based baseline hazard and a covariate process with time-varying deviations (b_i). Reducing the sample size to $n = 100$ resulted in larger standard errors and less power, but had little effect on type I error and bias, except for the case of a normal density–based baseline hazard and time-varying deviations (b_i) where greater values of β were needed to see bias.

The analytical standard errors from the Cox model do not account for the variability from estimating the sigmoidal parameters and may therefore be underestimated. Bootstrapped standard errors would account for the 2 sources of variation; however, bootstrapping within this simulation study is not computationally feasible. Because of this, we present the analytical standard errors. In Section 6, where we apply our methods to Alzheimer's data, we compare the analytical and bootstrapped standard errors.

6 | APPLICATION TO ALZHEIMER'S DATA

We applied our methodology to 2 Alzheimer's studies: one from the Rush Alzheimer's Disease Center (RADC) and the other from the ADNI. We chose to study the impact of a single biomarker on disease progression. Recently, Capuano and

colleagues[3] at Rush University provided convincing evidence for a sigmoidal trajectory for a measure of global cognition in the RADC dataset. This complements the existing literature on sigmoidal biomarker models for AD proposed by Jack and colleagues.[2] On the basis of this theoretical framework for Alzheimer's biomarkers and empirical evidence, we assumed a sigmoidal trajectory (2) for the time-varying biomarker in both analyses.

We fit models (A)-(E), and (F). We also fit the model that adjusts the lateral shift on the basis of an individual's disease severity in Section 4.3. To account for estimation of the covariate process in variance estimation, we used a simple bootstrap with 5000 resamples. For each study, we describe below the analyses, study population, biomarker, error model, origin, end point, and study entry definitions.

6.1 | Global cognition

We obtained a limited dataset from the RADC that combines data from the Religious Orders Study and the Memory and Aging Project.[29,30] The biomarker was an average of a battery of 19 cognitive test z scores,[31] which was taken to be a measure of global cognitive function. The origin was defined as birth, the study entry was defined as the time of baseline visit, and the end point was defined to be time of death or time of last cognitive testing based on the Mini-mental State Examination,[32] which is the last time the patient is known to be alive. Only subjects with at least one biomarker measurement and who carried a diagnosis of MCI at the time of biomarker measurement were included in the study. The global measure of cognition was translated by 2 units to achieve positive values. Gender and body mass index[33] were included as covariates in the model that adjusts the lateral shift on the basis of an individual's disease severity.

In choosing an error model, we used graphical diagnostics to determine goodness-of-fit of the absolute error and logit error models based on the sigmoid, $S(u)$, in (2). The absolute error model assumes that $Z(u) = S(u; \lambda, T_0) + a_i + b_i u$. We used least squares to estimate λ and T_0 based on data $(E_i, Z(E_i))$ and plotted the residuals, $Z(E_i) - S(E_i; \hat{\lambda}, \widehat{T_0})$, against study entry time, E_i. We saw that the residuals were centered around zero with nearly constant variance, indicating that the sigmoid-based absolute error model with fixed deviations aligned well with the observed biomarker data. We also considered the sigmoid-based logit error model that assumes that the transformed biomarker, $\text{logit}(Z_i(u)/M)$, is linear in u. We plotted the transformed observed biomarker values collected at study entry, $\text{logit}(Z_i(E_i)/M)$, against time at study entry, E_i, and saw a nonlinear, decreasing trend that accelerates with time. On the basis of these diagnostics, we assumed the global measure of cognition followed a sigmoid-based absolute error model with fixed deviations. As a sensitivity analysis, we also fit the sigmoid-based absolute error model with time-varying deviations, the sigmoid-based logit error model with both fixed and time-varying deviations, and the following linear absolute error model,

$$Z_i(u) = \lambda(u - T_0) + a_i + b_i u, \tag{14}$$

where λ and T_0 play similar roles as they do in the sigmoidal function. The corresponding severity adjusted model is given by

$$Z_i(u) = \lambda(u - \gamma' w_i) + a_i + b_i u, \tag{15}$$

where w_i is a vector of individual-specific covariates reflecting disease severity, which alter the lateral shift of the linear function.

Descriptive statistics are given in Table 4. A total of 737 subjects were included in the analysis, of whom 473 experienced the event. A plot of the cross-sectional global cognition score at time of baseline visit with a fitted loess curve is displayed in Figure 1. Although on the basis of cross-sectional, and not longitudinal data, a sigmoidal functional form is discernible, and a clear downward trend is present indicating a decline in global cognition with increasing age.

The results of the analysis are displayed in Table 5. For each of the 6 covariate models assumed (absolute error, sigmoid a_i, etc), the estimated regression coefficient of the biomarker, considering 2 different origins and treating the biomarker as time varying or fixed, is presented. Note that models (B), (C), and (D), which include the time-varying biomarker value at study entry as a fixed covariate, do not involve estimation of the covariate function and are thus listed with the same results for fixed (a_i) as for time-varying deviations (b_i). For these models, the 2-stage method proposed by Sperrin and Buchan[12] is comparable with our method. Assuming the sigmoid-based absolute error model with fixed deviations, a_i, the hazard ratio comparing the biomarker level over a one unit difference in the biomarker is 0.60 (exp(−0.51)) based on our proposed method (A). It is 0.49 (exp(−.72)) based on (E) that assumes the origin is study entry and treats the biomarker as a fixed covariate, adjusting for time between the origin and study entry. This difference highlights the importance of selection of the analytic approach for answering the clinical question. Given that there is support for the sigmoid-based

TABLE 4 Descriptive statistics of subjects included in the Rush Alzheimer's Disease Center analyses

	n	%
Number of subjects	737	
Male	218	29.6
	Mean	**SD**
BMI	26.6	5.1
Global cognition score	0.42	0.46
Age at baseline visit	80.9	7.5
Time from birth to last observation	87.8	7.2
Time from study entry to last visit	6.9	4.8

Abbreviations: BMI, body mass index.

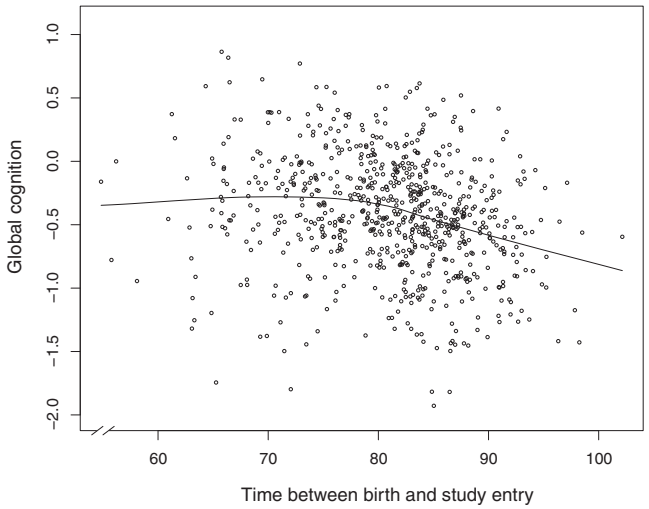

FIGURE 1 Plot of global cognition at study entry versus age at study entry with fitted loess curve

TABLE 5 Estimated regression coefficients corresponding to the biomarker (for 2 choices of origin and treating the biomarker as time varying or fixed) and bootstrapped standard errors (5000 replications) of Rush Alzheimer's Disease Center analysis with global cognition biomarker

	Birth origin			**Study entry origin**			
Covariate model	**A′:** $f(\widehat{Z_i(u)})$	**B:** $f(Z_i(E_i))$	**Sev Adj**	**C:** $f(Z_i(E_i))$	**D:** $f(Z_i(E_i))+E_i$	**E:** $f(\widehat{Z_i(t+E_i)})$	**F:** $f(\widehat{Z_i(t+E_i)})+E_i$
Absolute error, sigmoid, a_i	−0.51 (0.10)	−0.36 (0.10)	−0.35 (0.10)	−0.82 (0.10)	−0.72 (0.11)	−0.66 (0.09)	−0.72 (0.10)
Absolute error, sigmoid, b_i	−0.27 (0.09)	−0.36 (0.10)	−0.30 (0.10)	−0.82 (0.10)	−0.72 (0.11)	−0.61 (0.09)	−0.65 (0.10)
Logit error, sigmoid, a_i	−0.60 (0.07)	−0.40 (0.07)	−0.38 (0.07)	−0.93 (0.07)	−0.81 (0.08)	−0.67 (0.06)	−0.81 (0.07)
Logit error, sigmoid, b_i	−0.56 (0.07)	−0.40 (0.07)	−0.33 (0.07)	−0.93 (0.07)	−0.81 (0.08)	−0.58 (0.06)	−0.74 (0.07)
Absolute error, linear trend, a_i	−0.54 (0.10)	−0.36 (0.10)	−0.19 (0.09)	−0.82 (0.10)	−0.72 (0.11)	−0.59 (0.09)	−0.72 (0.10)
Absolute error, linear trend, b_i	−0.50 (0.09)	−0.36 (0.10)	−0.15 (0.09)	−0.82 (0.10)	−0.72 (0.11)	−0.52 (0.09)	−0.65 (0.10)

absolute error model with fixed deviations based on our diagnostics, and it does not require any assumptions on the baseline hazard, we would report the hazard ratio of 0.60.

For the birth origin analyses, the estimates from the model that treats the biomarker as fixed (B) are smaller in magnitude by approximately 30% than those from the model that estimates the full covariate trajectory (A), with the exception of the absolute error model with time-varying deviations (second row) where the effect is increased in magnitude by 30%. The estimates from the severity adjusted analyses using birth as the origin (column 3) are also moderately different from the estimated time-varying model (A), with more of a difference seen with the linear trend biomarker models, in which the change in magnitude of the regression estimates is more than 65%. This suggests that subject-specific adjustments to the sigmoidal model are potentially important. This is seen as well with respect to the covariate function parameter

TABLE 6 Analytical standard errors versus bootstrapped standard errors (5000 replications) for Rush Alzheimer's Disease Center analysis

	Birth origin						Study entry origin							
	$\widehat{f(Z_i(u))}$		$f(Z_i(E_i))$		Sev. Adj.		$f(Z_i(E_i))$		$f(Z_i(E_i))+E_i$		$\widehat{f(Z_i(t+E_i))}$		$\widehat{f(Z_i(t+E_i))}+E_i$	
	Wald	Boot	Wald	Boot	Wald	Boot	Wald	Boot	Wald	Boot	Wald	Boot	Wald	Boot
Absolute error, sigmoid, a_i	0.06	0.10	0.06	0.10	0.06	0.10	0.06	0.10	0.06	0.11	0.06	0.09	0.06	0.10
Absolute error, sigmoid, b_i	0.05	0.09	0.06	0.10	0.06	0.10	0.06	0.10	0.06	0.11	0.05	0.09	0.06	0.10
Logit error, sigmoid, a_i	0.06	0.07	0.07	0.07	0.07	0.07	0.06	0.07	0.07	0.08	0.07	0.06	0.07	0.07
Logit error, sigmoid, b_i	0.06	0.07	0.07	0.07	0.07	0.07	0.06	0.07	0.07	0.08	0.06	0.06	0.06	0.07
Absolute error, linear trend, a_i	0.06	0.10	0.06	0.10	0.06	0.09	0.06	0.10	0.06	0.11	0.06	0.09	0.06	0.10
Absolute error, linear trend b_i	0.05	0.09	0.06	0.10	0.06	0.09	0.06	0.10	0.06	0.11	0.05	0.09	0.06	0.10

TABLE 7 Descriptive statistics of subjects included in the Alzheimer's Disease Neuroimaging Initiative analyses

	n	%
Number of subjects	786	
Male	465	59
One or more APOE4 alleles	408	52
	Mean	**SD**
Hippocampal volume	6735	1158
Age at first hippocampal measurement	73	7
Time from birth to last visit	76	8
Time from study entry to last visit	3	2
Converted to AD	280	

Abbreviation: AD, Alzheimer's disease.

estimates given in Table S2. When study entry is taken to be the origin, there are moderate differences across the 4 models that we fit (columns 4 through 7), and these estimates can vary dramatically from the estimates that are obtained when using birth as the origin (columns 1 through 3).

In summary, for this data example, the choice of time origin and treatment of time-varying covariates do impact inference and thought must be given to the implications of these choices. Although the magnitude of the estimates for the models differ slightly, the direction of the effect among all models using both origins agrees.

As mentioned in Section 5, the analytical standard errors from the Cox model do not account for the variability in estimating the sigmoidal parameters and are likely to be underestimated. Table 6 compares the analytical standard errors with the bootstrapped standard errors. For the logit error models, the analytical standard errors are comparable with those that are obtained via bootstrapping. For the biomarker models based on the sigmoid with an absolute error model and based on the linear model, the analytical standard errors are greatly underestimated, suggesting that bootstrapped standard errors should be used.

6.2 | Hippocampal volume

We also analyzed data from the ADNI database (adni.loni.usc.edu). Launched in 2003, the ADNI is a longitudinal study aimed at using biomarkers of AD and clinical and neuropsychiatric assessments to measure the progression of MCI and AD.[34]

The origin was defined as birth, the study entry was defined as the first biomarker measurement, and the end point was defined to be the time of the first diagnosis of AD in the ADNI study. Only subjects with at least one biomarker measurement and who carried a diagnosis of MCI at the time of biomarker measurement were included in the study. Gender and APOE4 allele status were included as covariates in the model that adjusts the lateral shift on the basis of an individual's disease severity. After considering several potential markers in a preliminary analysis, we chose to focus on hippocampal volume, where a clear downward trend was seen in the cross-sectional plot. We standardized and shifted the values by 3 units so that all resulting biomarkers quantities were positive. Descriptive statistics for the study population are given in Table 7.

TABLE 8 Estimated regression coefficients corresponding to the biomarker (for 2 choices of origin and treating the biomarker as time varying or fixed) and bootstrapped standard errors (5000 replications) of Alzheimer's Disease Neuroimaging Initiative analysis with hippocampal volume biomarker

	Birth origin			Study entry origin			
Covariate model	**A′:** $f(\widehat{Z_i(u)})$	**B:** $f(Z_i(E_i))$	**Sev Adj**	**C:** $f(Z_i(E_i))$	**D:** $f(Z_i(E_i))+E_i$	**E:** $f(\widehat{Z_i(t+E_i)})$	**F:** $f(\widehat{Z_i(t+E_i)})+E_i$
Absolute error, sigmoid, a_i	−0.78 (0.07)	−0.75 (0.07)	−0.76 (0.07)	−0.72 (0.07)	−0.83 (0.08)	−0.83 (0.08)	−0.83 (0.08)
Absolute error, sigmoid, b_i	−0.74 (0.07)	−0.75 (0.07)	−0.74 (0.07)	−0.72 (0.07)	−0.83 (0.08)	−0.81 (0.07)	−0.81 (0.08)
Logit error, sigmoid, a_i	−0.83 (0.14)	−0.80 (0.13)	−0.81 (0.13)	−0.86 (0.09)	−0.93 (0.12)	−0.93 (0.12)	−0.93 (0.12)
Logit error, sigmoid, b_i	−0.83 (0.13)	−0.80 (0.13)	−0.80 (0.13)	−0.86 (0.09)	−0.93 (0.12)	−0.92 (0.12)	−0.92 (0.12)
Absolute error, linear trend, a_i	−0.79 (0.07)	−0.75 (0.07)	−0.73 (0.07)	−0.72 (0.07)	−0.83 (0.08)	−0.82 (0.08)	−0.83 (0.08)
Absolute error, linear trend, b_i	−0.78 (0.07)	−0.75 (0.07)	−0.72 (0.07)	−0.72 (0.07)	−0.83 (0.08)	−0.80 (0.07)	−0.81 (0.08)

We used graphical diagnostics, as described for the Rush data, to chose an error model for these data. A plot of $\text{logit}(Z(E_i)/M)$ against E_i resulted in a clear linear trend with constant variability. In considering an absolute error model based on the sigmoid in (2) and linear trends in (14), neither of the residual plots had mean zero. Therefore, we assumed a sigmoid-based logit error model for this analysis. We fit the sigmoid- and linear trend-based absolute error models as sensitivity analyses.

The results of the analysis are reported in Table 8. There were only small differences seen across the 3 birth origin models and the 4 study entry origin models that were fit, regardless of the biomarker model chosen. This was likely due to short follow-up and only small variations in the measured biomarker levels but potentially a nonexponential functional form for the baseline hazard function.

7 | DISCUSSION

In settings such as AD, using study entry as the time origin in a time-to-event analysis is convenient because time-varying biomarkers are often measured only at study entry and thus can be used as fixed baseline covariates. If study entry is not an appropriate time origin, the question of interest is "under what conditions is it valid to use an incorrect time origin of study entry in a Cox model with time-varying covariates?" The answer to this question for the case of fixed covariates was established by Korn et al.[8] We have extended the work of Korn et al to the case of a time-varying covariate; we have shown that using an incorrect origin of study entry is valid only for special cases of the baseline hazard function provided that the functional form of the covariate process is fully observed. We further examined the implications of treating a continuous, time-varying covariate as fixed at its value at study entry for both the onset and study entry origins; we showed that this is only valid in certain cases of the baseline hazard and time-varying covariate. These theoretical findings underscore the necessity for full observation of time-varying covariates or good modeling of them. This is challenging unless good prior data are available.

For situations in which the full covariate function is not available, we developed methods for estimating regression coefficients when the time-varying covariate follows the general error model $f(Z_i(u)) = f(S(u)) + a_i + b_i u$, when only $f(Z_i(E_i))$ is observed; this is a generalization of the covariate model with fixed residuals given by Sperrin and Buchan.[12] While Sperrin and Buchan focused on predictive ability, we have examined the bias and standard errors of hazard model estimates.

Our analysis of the RADC data provides an example of where the choice of time origin and biomarker model can dramatically affect estimation in the presence of a time-varying covariate. Estimates from an analysis from study entry that treats the time-varying biomarker as fixed are quite different from the estimates of our proposed model, even when time between onset and study entry is adjusted for as a continuous covariate. On the other hand, our analysis of ADNI data provides an example of when incorrectly viewing a time-varying covariate as fixed had little effect on estimation that was likely due to covariate measurements that are nearly constant over a short follow-up period. However, the choice of time origin did have a noticeable effect, suggesting that the baseline hazard was not of the requisite exponential functional form to permit the use of study entry as the origin.

In this paper, we assume that the history of the biomarker is not used for case ascertainment. If it were used, such as through a requirement that subjects' value for the biomarker be below a threshold, we would potentially have to adjust

our estimation procedure in Section 4 to account for partial observation of the biomarker function, S. Furthermore, interpretation of results would be limited to the sample population.

A main limitation of our results in Theorem 1 parts (a) and (c) is that the time of onset must be known. While exact onset dates are not known within the setting of AD, there are intermediate milestone times that may be known up to short intervals of time. And given the relatively long period of time during which biomarkers such as cognitive tests or amyloid are nonelevated, the error around the exact onset date may not be critical. However, if we cannot trust the purported time origin and revert to using study entry as the origin, the contribution of our paper is to identify when that is a reasonable approach and when it is not. Specifically, our analytical results show that it is valid to use study entry as the origin and treat time-varying covariates as fixed only when the baseline hazard measured from onset of disease is constant and the continuous, time-varying covariate process satisfies the condition in Theorem 1 part (c).

The strength of this work is that our analytical results and methods can be applied to any setting in which time-varying covariates follow continuous trajectories. It is important to note that our results are particular to the Cox model in which time-dependent population-level expressions are absorbed into the baseline hazard function, as in (9). For example, the methods we proposed cannot be directly applied to accelerated failure time (AFT) regression, as the AFT model with time-varying covariates cannot be simplified in this way. However, Sperrin and Buchan[12] showed through simulations that including individual-specific deviations in an AFT regression model is superior in its predictive ability to the model that includes the time-varying covariate treated as fixed.

As serial biomarker measurements are increasingly collected in longitudinal studies such as the RADC studies and ADNI, possibilities exist for improved analyses via better parametric modeling, identification of both components of individual level deviation, a_i and b_i, and more efficient estimation of the dependence between the event time and biomarker. With 2 or more biomarker measurements per individual, our methods would allow for the estimation of the desired regression parameter whereas methods for Cox regression with missing data and joint modeling of survival and longitudinal data are not likely to be robust with so few covariate measurements.

ACKNOWLEDGEMENTS

The authors wish to thank Deborah Blacker for helpful feedback in the preparation of this manuscript. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). The ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc; Cogstate; Eisai Inc; Elan Pharmaceuticals, Inc; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc; Fujirebio; GE Healthcare; IXICO Ltd; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co, Inc; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. The ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

This research was supported in part by NIHT32NS048005, R01NS094610, P50AG005134, P01AG036694, and P30AG10161 and R01AG17917 (RADC). We thank the study participants and staff of the Rush Alzheimer's Disease Center.

ORCID

Catherine Lee http://orcid.org/0000-0003-0008-9052

REFERENCES

1. Association Alzheimer's. 2017 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2017;13(4):18.
2. Jack C, Knopman D, Jagust W, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol*. 2010;9(1):119-128.

3. Capuano A, Wilson R, Leurgans S, Dawson J, Bennett D, Hedeker D. Sigmoidal mixed models for longitudinal data. *Stat Methods Med Res*. 2016. https://doi.org/10.1177/0962280216645632
4. Cnaan A, Ryan L. Survival analysis in natural history studies of disease. *Stat Med*. 1989;8(10):1255-1268.
5. Keiding N, Knuiman M. Comments on 'Survival analysis in natural history studies of disease' by Cnaan and Ryan Stat Med. 1989 Oct;8(10):1255-68. *Stat Med*. 1990;9(10):1221-1222.
6. Cox DR, Oakes D. *Analysis of Survival Data*. New York: Chapman and Hall; 1984.
7. Fieberg J, DelGiudice G. What time is it? Choice of time origin and scale in extended proportional hazards models. *Ecology*. 2009;90(6):1687-1697.
8. Korn E, Graubard B, Midthune D. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. *Am J Epidemiol*. 1997;145(1):72-80.
9. Lamarca R, Alonso J, Gomez G, Munoz A. Left-truncated data with age as time scale: an alternative for survival analysis in the elderly population. *J Gerontol A Biol Sci Med Sci*. 1998;53(5):M337-M343.
10. Liestol K, Andersen P. Updating of covariates and choice of time origin in survival analysis: problems with vaguely defined disease states. *Stat Med*. 2002;21(23):3701-3714.
11. Pencina M, Larson M, D'Agostino R. Choice of time scale and its effect on significance of predictors in longitudinal studies. *Stat Med*. 2007;26:1343-1359.
12. Sperrin M, Buchan I. Modelling time to event with observations made at arbitrary times. *Stat Med*. 2013;32(1):99-109.
13. Thiébaut A, Bénichou J. Choice of time-scale in Cox's model analysis of epidemiologic cohort data: a simulation study. *Stat Med*. 2004;23(24):3803-3820.
14. Griffin B, Anderson G, Shih R, Whitsel E. Use of alternative time scales in Cox proportional hazard models: implications for time-varying environmental exposures. *Stat Med*. 2012;31(27):3320-3327.
15. Lin D, Ying Z. Cox regression with incomplete covariate measurements. *J Am Stat Assoc*. 1993;88(424):1341-1349.
16. Zhou H, Pepe M. Auxiliary covariate data in failure time regression. *Biometrika*. 1995;82(1):139-149.
17. Chen H, Little R. Proportional hazards regression with missing covariates. *J Am Stat Assoc*. 1999;94(447):896-908.
18. Wu L, Liu W, Yi G, Huang Y. Analysis of longitudinal and survival data: joint modeling, inference methods, and issues. *J Probab Stat*. 2012;2012.
19. Wang Y, Taylor J. Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *J Am Stat Assoc*. 2001;96(455):895-905.
20. Tsiatis A, DeGruttola V, Wulfsohn M. Modeling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *J Am Stat Assoc*. 1995;90(429):27-37.
21. Wulfsohn M, Tsiatis A. A joint model for survival and longitudinal data measured with error. *Biometrics*. 1997;53(1):330-339.
22. Jack C, Knopman D, Jagust W, et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol*. 2013;12(2):207-216.
23. Fisher L, Lin D. Time-dependent covariates in the Cox proportional-hazards regression model. *Annu Rev Public Health*. 1999;20:145-157.
24. Ingram D, Makuc D, Feldman J. Re: 'Time-to-event analysis of longitudinal follow-up of a survey: choice of the time scale'. *Am J Epidemiol*. 1997;146(6):528-529.
25. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. *Stat Med*. 2005;24(11):1713-1723.
26. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria; R Foundation for Statistical Computing; 2014.
27. Brazzale A. hoa: an R package bundle for higher order likelihood inference. *Rnews*. 2005;5/1 May 2005:20-27.
28. Therneau T. A package for survival analysis in S. version 2.38; 2015.
29. Bennett D, Schneider J, Arvanitakis Z, Wilson R. Overview and findings from the religious orders study. *Curr Alzheimer Res*. 2012;9(6):628-645.
30. Bennett D, Schneider J, Buchman A, Barnes L, Boyle P, Wilson R. Overview and findings from the rush memory and aging project. *Curr Alzheimer Res*. 2012;9(6):646-663.
31. Wilson R, Boyle P, Yu L, et al. Temporal course and pathologic basis of unawareness of memory loss in dementia. *Neurology*. 2015;85(11):984-991.
32. Folstein M, Folstein S, McHugh P. Mini-mental state?: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.
33. Buchman A, Wilson R, Bienias J, Shah R, Evans D, Bennett D. Change in body mass index and risk of incident Alzheimer disease. *Neurology*. 2005;65(6):892-897.
34. Mueller S, Weiner M, Thal L, et al. Ways toward an early diagnosis in Alzheimer's disease: the Alzheimer's Disease Neuroimaging Initiative (ADNI). *Alzheimers Dement*. 2005;1:55-66.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Lee C, Betensky RA, for the Alzheimer's Disease Neuroimaging Initiative. Time-to-event data with time-varying biomarkers measured only at study entry, with applications to Alzheimer's disease. *Statistics in Medicine*. 2018;37:914–932. https://doi.org/10.1002/sim.7547

APPENDIX

Proof of Theorem 1 part (a). (Korn et al[8] extension)
Substituting the definition of time from study entry, $T_i = U_i - E_i$, into the definition of the hazard function

$$\lambda_{T_i}(t|\overline{w_i}(\cdot), E_i) = \lim_{\delta \to 0} P(t < T_i \leq t + \delta | T_i \geq t, \overline{w_i}(\cdot), E_i)/\delta$$

gives

$$\begin{aligned}
\lambda_{T_i}(t|\overline{w_i}(\cdot), E_i) &= \lim_{\delta \to 0} P(t + E_i < T_i + E_i \leq t + \delta + E_i | T_i + E_i \geq t + E_i, \overline{w_i}(\cdot), E_i)/\delta \\
&= \lim_{\delta \to 0} P(t + E_i < U_i \leq t + \delta + E_i | U_i \geq t + E_i, \overline{w_i}(\cdot), E_i)/\delta, \\
&= \lambda_{U_i}(t + E_i | \overline{w_i}(\cdot), E_i) \\
&\overset{(5)}{=} \lambda_{0U}(t + E_i) e^{\beta w_i(t+E_i)}.
\end{aligned} \tag{A1}$$

If $\lambda_{0U}(u) = ce^{\gamma u}$ for some $c > 0$, then (A1) yields

$$\lambda_{T_i}(t|\overline{w_i}(\cdot), E_i) = ce^{\gamma t} e^{\beta w_i(t+E_i) + \gamma E_i}.$$

Thus, in comparison with model (6), it is clear that $\beta = \xi$ provided $\lambda_{0T}(t) = ce^{\gamma t}$.

If $\beta = \xi$, it follows that

$$\lambda_{0T}(t) = \lambda_{0U}(t + E_i) e^{-\gamma E_i},$$

because $\lambda_{T_i}(t) = \lambda_{0T}(t) e^{\xi w_i(t+E_i) + \gamma E_i}$ by (6) and $\lambda_{T_i}(t) = \lambda_{0U}(t + E_i) e^{\beta w_i(t+E_i)}$ by (A1). In order for $\lambda_{0T}(t)$ to be independent of i, it must be that $\lambda_{0U}(u) = ce^{\gamma u}$. □

Proof of Theorem 1 part (b). First, observe that the true survival model in (5) can be rewritten as

$$\lambda_{U_i}(u|\overline{w_i}(\cdot), E_i) = \lambda_{0U}(u) e^{\beta w_i(E_i)} e^{\beta[w_i(u) - w_i(E_i)]}. \tag{A2}$$

If β in (A2) is equal to α in (7), then

$$\lambda_{1U_i}(u) = \lambda_{0U}(u) e^{\beta[w_i(u) - w_i(E_i)]}. \tag{A3}$$

This implies that $w_i(u) - w_i(E_i) = g(u)$ for some function g that is independent of i; otherwise, $\lambda_{1U}(u)$ would depend on i, which is not the case. The condition $w_i(u) - w_i(E_i) = g(u)$ requires that the function g is zero on the support of E since it is zero at all study entry times, E_i. Note that $g(u)$ may be nonzero outside of the support of E.

If $w_i(u) = w_i(E_i) + g(u)$, then the true model given by (5) and fitted model given by (7) are equivalent and $\beta = \alpha$. □

Proof of Theorem 1 part (c). Note that the true survival model in (A1) can be rewritten as

$$\lambda_{T_i}(t|\overline{w_i}(\cdot), E_i) = \lambda_{0U}(t + E_i) e^{\beta[w_i(t+E_i) - w_i(E_i)]} e^{\beta w_i(E_i)}. \tag{A4}$$

If β in (A4) is equal to ζ in (8), then

$$\lambda_{2T}(t) = \lambda_{0U}(t + E_i) e^{\beta[w_i(t+E_i) - w_i(E_i)] - \eta E_i}. \tag{A5}$$

It must be that $\lambda_{0U}(t + E_i) e^{-\eta E_i} = h(t)$ and $w_i(t + E_i) - w_i(E_i) = k(t)$ for some functions h and k independent of i; otherwise, the right-hand side of (A5) would depend on i and cannot be equal to $\lambda_{2T}(t)$, an expression independent

of i. Since we are assuming that all hazard functions are continuous, $\lambda_{0U}(t+E_i)e^{-\eta E_i} = h(t)$ implies $\lambda_{0U}(u) = ce^{\eta u}$, for some $c > 0$.

If $w_i(t+E_i) - w_i(E_i) = k(t)$ and $\lambda_{0U}(u) = ce^{\eta u}$ for $c > 0$, then the true model in (A4) is

$$\lambda_{T_i}(t|\overline{w_i}(\cdot), E_i) = ce^{\eta t + \beta k(t)} e^{\beta w_i(E_i) + \eta E_i}.$$

Comparing this with (8), it must be that $\beta = \zeta$ and $\lambda_{2T}(t) = ce^{\eta t + \beta k(t)}$. □