Lifetime Data Anal (2006) 12:169–190
DOI 10.1007/s10985-006-9002-4

Analysis of longitudinal health-related quality of life data with terminal events

Zhezhen Jin · Mengling Liu · Steven Albert · Zhiliang Ying

Received: 1 June 2004 / Accepted: 31 January 2006
Published online: 1 July 2006

Abstract Longitudinal health-related quality of life data arise naturally from studies of progressive and neurodegenerative diseases. In such studies, patients' mental and physical conditions are measured over their follow-up periods and the resulting data are often complicated by subject-specific measurement times and possible terminal events associated with outcome variables. Motivated by the "Predictor's Cohort" study on patients with advanced Alzheimer disease, we propose in this paper a semiparametric modeling approach to longitudinal health-related quality of life data. It builds upon and extends some recent developments for longitudinal data with irregular observation times. The new approach handles possibly dependent terminal events. It allows one to examine time-dependent covariate effects on the evolution of outcome variable and to assess nonparametrically change of outcome measurement that is due to factors not incorporated in the covariates. The usual large-sample properties for parameter estimation are established. In particular, it is shown that relevant parameter estimators are

Z. Jin (✉)
Department of Biostatistics, Columbia University, New York, NY, USA
e-mail: zjin@biostat.columbia.edu

M. Liu
Division of Biostatistics, School of Medicine, New York University, New York, NY, USA
e-mail: mengling.liu@med.nyu.edu

S. Albert
Department of Behavioral & Community Health Sciences, University of Pittsburgh, Pittsburgh, PA, USA
e-mail: smalbert@pitt.edu

Z. Ying
Department of Statistics, Columbia University, New York, NY, USA
e-mail: zying@stat.columbia.edu

asymptotically normal and the asymptotic variances can be estimated consistently by the simple plug-in method. A general procedure for testing a specific parametric form in the nonparametric component is also developed. Simulation studies show that the proposed approach performs well for practical settings. The method is applied to the motivating example.

Keywords Censoring · Health-related quality of life · Inverse probability weighting · Semiparametric regression · Terminal event

1 Introduction

In sociomedical studies of progressive and neurodegenerative diseases, such as the Alzheimer's disease and the Parkinson disease, health-related quality of life (HRQoL) is a very important outcome. Measurement of health-related quality of life, however, is complicated because it is not directly observable and is often assessed by various self or proxy rated questionnaires on daily activity, daily affect, functional ability, cognitive performance etc. Albert et al. (1999) showed that the daily activity measure is significantly correlated with a variety of other indicators of HRQoL, such as daily affect, functional ability, cognitive performance and clinician global ratings of dementia severity, which makes the daily activity as a good surrogate measurement of HRQoL. Patients with progressive and neurodegenerative diseases typically experience difficulties in carrying out daily activities. The level of difficulty usually increases over time, resulting in more dependence on caregivers. An important aspect is the evolution of the capability to conduct routine and normal daily activities and how the deterioration is related to various factors. To address these issues, daily activities are often measured longitudinally over a period of time in HRQoL studies.

An interesting example is the ''Predictor's Cohort'' study. Between 1988 and 1998, Columbia University, the Johns Hopkins University and the Massachusetts General Hospital jointly conducted an observational longitudinal study to examine quality of life in people with advanced Alzheimer's disease (AD). One of the key aims of the study was to determine the relationship between dementia severity and the HRQoL outcomes. The enrolled cohort had to meet DSM-III criteria for dementia and NINDS-ADRDA criteria for probable AD. Each enrolled patient was scheduled to be evaluated every 6 months. However, the actual follow-up times varied among patients and the follow-up of some patients was terminated before the end of study because of death. Fifteen distinct outdoor and indoor activities which are within the capacity of patients with dementia receiving supervision and aid were measured. Albert et al. (1996) showed that correlations between the items were high and the items formed a single component in an exploratory factor analysis and suggested to measure the daily activity by the frequency of the 15 activities. The dementia severity was assessed by the score of the modified Mini-Mental State Exam (mMMS) (Stern et al. 1987). Albert et al. (1996, 2000) contain details about the study background as well as instruments for daily activities.

Statistical analysis of longitudinal HRQoL studies, however, is quite challenging due to the limited study follow-up and presence of terminating event, such as death. Linear and linear mixed effects models, such as pattern mixture models, and generalized estimating equation (GEE) have been used in the analysis of HRQoL (Fayers and Machin 2000; Fairclough 2002). These models do not take into account the censoring and their validity often depends on parametric model assumptions. Standard survival analysis techniques have also been adopted for the analysis of HRQoL with definition of a "milestone" event. However, such an approach ignores intermediate information and interpretation of analysis results depends critically on the definition of the "milestone" (Albert et al. 2000). In clinical trials of chronic diseases, such as AIDS and cancer, quality-adjusted survival analysis (Q-TWiST) has been developed (Cole et al. 1993; Zhao and Tsiatis 1997). Unfortunately, the Q-TWiST approach is difficult to apply to the analysis of HRQoL due to the difficulty in defining the disease stages in patients who are already in advanced stages. Also as pointed out by the associate editor, the Q-TWiST approach focuses on developing health states based on categorized HRQoL scores, see Cole et al. (2004). Practically, it is also noted that progression of the diseases and risk of death are related, so that limiting analyses to survivors is likely to underestimate or overestimate the quality-of-life impact of disease. Recently, Dupuy and Mesbah (2002) focused on joint modeling of event time and nonignorable missing longitudinal data with combination of a first-order Markov model for covariate and time-dependent Cox model for event time. Although their approach can handle longitudinally measured covariates, it cannot deal with longitudinally observed outcome variables. Hence there is clearly a need for an analysis method that can directly model original HRQoL longitudinal data with ease of interpretation as well as taking the limited study follow-up with possibly irregular individual observation times and terminating event into account.

The present paper is aimed at investigating a semiparametric modeling approach for analysis of longitudinal HRQoL data. The semiparametric modeling approach for longitudinal data has been studied extensively; see Moyeed and Diggle (1994), Zeger and Diggle (1994), Hoover et al. (1998), Lin and Ying (2001) among others. In this paper, we extend the approach of Lin and Ying (2001) by incorporating the presence of terminating event. We apply the technique of inverse probability weighting after adjusting possibly time-dependent covariates; see Robins and Rotnitzky (1992) and Ghosh and Lin (2002). Our method takes into account the fact that subjects who die cannot have further HRQoL measurements. We also develop a simple test for checking whether or not the nonparametric component actually follows a specific parametric form.

In the next section, we introduce notation and model specification. Estimation procedure is presented in Sect. 3 and its asymptotic properties and plug-in variance estimation are addressed in Sect. 4. In Sect. 5, a general procedure for testing a specific parametric form of the nonparametric component is presented. Simulation results and the analysis of the dataset from the ‘‘Predictor’s Cohort’’ study are presented in Sect. 6. Some concluding remarks are given in Sect. 7.

2 Notation and model specification

We shall use n to denote the number of study subjects. For the ith subject, $i = 1,\ldots, n$, we use $Y_i(t)$ to denote the outcome measurement and $X_i(t)$ to denote a $p \times 1$ vector of covariates at time t. The subject is followed until the time of either censoring or failure (death), whichever occurs first. Let C_i be the censoring time, S_i be the survival time, $\tilde{S}_i = \min(C_i, S_i)$ and $\Delta_i = I(S_i \le C_i)$. During the follow-up period of the ith subject, the measurements of $Y_i(\cdot)$ are taken at certain time points. Let $N_i^*(t)$ be the number of such measurements up to time t if there is neither censoring nor failure. The covariate process $X_i(t)$, however, is assumed to be observed for $t \in [0, \tilde{S}_i]$. Thus, the observations consist of $\{\tilde{S}_i, \Delta_i, N_i^*(t), Y_i(t)dN_i^*(t), X_i(t), t \in [0, \tilde{S}_i]\}$; $i = 1,\ldots, n$. Notice that $Y_i(t)dN_i^*(t)$ implies that $Y_i(\cdot)$ is observed only at jump points of the process $N_i^*(\cdot)$. In this paper, it is assumed that $N_i^*(\cdot)$ depends on only part or full of the covariate process $X_i(\cdot)$, which implies that the survival or censoring times are noninformative to observation schedules conditional on the covariate process; see (3.5).

First, we relate the outcome measurement $Y(\cdot)$ to covariate $X(\cdot)$ by

$$Y_i(t) = \alpha(t) + \beta^{\mathrm{T}} X_i(t) + \epsilon_i(t), \tag{2.1}$$

where $\alpha(\cdot)$ is a completely unspecified function of time, β is an unknown $p \times 1$ parameter vector, and $\epsilon_i(\cdot)$ is a mean 0 measurement error independent of $X_i(\cdot)$. Note that β reflects the linear effect of covariate $X(\cdot)$ on $Y(\cdot)$, while $\alpha(\cdot)$ reflects the change of $Y(\cdot)$ due to factors other than $X(\cdot)$ as well as baseline trend. Model (2.1) contains two nonparametric components. The stochastic part, represented by the error process $\epsilon(\cdot)$, is nonparametric since except for zero mean its distribution can be arbitrary. The trend function $\alpha(\cdot)$ is also completely unspecified. Moyeed and Diggle (1994) and Zeger and Diggle (1994) studied the model (2.1) with $\epsilon(\cdot)$ being a zero-mean stationary Gaussian process. Lin and Ying (2001) relaxed this assumption to any zero-mean process.

The counting process $N^*(\cdot)$ can be handled naturally by using the proportional means model as in Pepe and Cai (1993), Lawless and Nadeau (1995) and Lin et al. (2000). Specifically, we model $N^*(\cdot)$ by

$$E\{dN_i^*(t)|Z_{i1}(t)\} = \exp\{\gamma^{\mathrm{T}} Z_{i1}(t)\}d\Lambda(t), \tag{2.2}$$

where $\Lambda(\cdot)$ is a completely unspecified function of time, γ is unknown $q \times 1$ parameter vector, $Z_{i1}(\cdot)$ is a known covariate vector which is a part of $X_i(\cdot)$. Note that equation (2.2) specifies the mean frequency of observation times for the i-th subject. It provides a way of modeling irregular and possibly subject-specific observation time points. When the observation time is independent of the process, for example, observations are taken at prespecified time points, γ takes value 0 in model (2.2), and the process $N^*(\cdot)$ only depends on the unspecified baseline cumulative intensity function $\Lambda(\cdot)$.

The analysis of longitudinal data with model specifications (2.1) and (2.2) was first introduced by Lin and Ying (2001). Since the approach of Lin and Ying (2001) does not take terminating event into account, it is not valid in the presence of terminating events. Specifically, the validity of the approach of Lin and Ying (2001) relies on the assumption that

$$E[Y_i(t)|X_i(t), T_i \geq t] = E[Y_i(t)|X_i(t)],$$

which essentially assumes the noninformativeness of the censoring. With the existence of terminal event which may be correlated with the responses, this assumption may not hold. To overcome this difficulty, the survival time is modeled based on data $\{\tilde{S}_i, \Delta_i, X_i(\cdot)\}$. Here, we use the Cox proportional hazards model for the survival time S (Cox 1972):

$$\lambda_i(t|Z_{i2}(t)) = \lambda_d(t)\exp\{\xi^{\mathrm{T}} Z_{i2}(t)\}, \tag{2.3}$$

where ξ is unknown $r \times 1$ parameter vector, $\lambda_d(\cdot)$ is completely unspecified function of time, and $Z_{i2}(\cdot)$ is a known covariate vector which is a part of $X_i(\cdot)$.

3 Parameter estimation

Let $\delta_i(t) = I\{C_i \geq t\}$ and $\mathcal{A}(t) = \int_0^t \alpha(s)d\Lambda(s)$. When C_i is always observed prior to S_i, then

$$M_i(t;\mathcal{A},\beta,\gamma) = \int_0^t \left[\delta_i(s)(Y_i(s) - \beta^{\mathrm{T}} X_i(s))dN_i^*(s) - \delta_i(s)\exp\{\gamma^{\mathrm{T}} Z_{i1}(s)\}d\mathcal{A}(s)\right] \tag{3.1}$$

is a mean-zero stochastic process because the second term is essentially the conditional expectation of the first term. Following Lin and Ying (2001), β, $\mathcal{A}(\cdot)$, and γ can be estimated with following estimating equations:

$$\sum_{i=1}^n M_i(t;\mathcal{A},\beta,\gamma) = 0, \tag{3.2}$$

$$\sum_{i=1}^n \int_0^\infty W(t)\delta_i(t)\left[X_i(t) - \bar{X}(t;\gamma)\right]\left[Y_i(t) - \bar{Y}(t;\gamma) - \beta^{\mathrm{T}}(X_i(t) - \bar{X}(t;\gamma))\right]dN_i^*(t) = 0, \tag{3.3}$$

$$\sum_{i=1}^n \int_0^\infty \left(Z_{i1}(t) - \bar{Z}_1(t;\gamma)\right)\delta_i(t)dN_i^*(t) = 0, \tag{3.4}$$

where $W(\cdot)$ is a possibly data-dependent weight function, and $\bar{X}(t;\gamma) = \sum_{i=1}^n \delta_i(t)X_i(t)e^{\gamma^T Z_{i1}(t)}/\sum_{i=1}^n \delta_i(t)e^{\gamma^T Z_{i1}(t)}$, $\bar{Y}(t;\gamma) = \sum_{i=1}^n \delta_i(t)Y_i(t)e^{\gamma^T Z_{i1}(t)}/\sum_{i=1}^n \delta_i(t)e^{\gamma^T Z_{i1}(t)}$, $\bar{Z}_1(t;\gamma) = \sum_{i=1}^n \delta_i(t)Z_{i1}(t)e^{\gamma^T Z_{i1}(t)}/\sum_{i=1}^n \delta_i(t)e^{\gamma^T Z_{i1}(t)}$

The weight function $W(\cdot)$ can be chosen analogous to that in weighted logrank statistic in univariate survival analysis, such as $W(\cdot) \equiv 1$ or functions of the survival function. Note that the γ is estimated by (3.4), β is estimated by (3.3) and $\mathcal{A}(\cdot)$ is estimated by (3.2).

Since censoring times C_is are not always known due to the terminating event, such as death in the Alzheimer's "Predictor's Cohort" study, the $\delta_i(t)$s are not always available and estimating Eqs. (3.2)–(3.4) cannot be implemented. Specifically, the $\delta_i(t)$ is unknown if $\min\{C_i,t\} > S_i$. It is assumed that

$$E[dN_i^*(t)|X_i(t), Y_i(t), \tilde{S}_i > t] = E[dN_i^*(t)|Z_{i1}(t)], \tag{3.5}$$

which implies that the process $N_i^*(\cdot)$ depends on X_i, Y_i, C_i and S_i only essentially through Z_{i1}. In particular, the assumption (3.5) is satisfied if observation schedule is prespecified and any deviation from the schedule happens in a completely random fashion as in a typical clinical trial setting. Let $\delta_i^D(t) = I\{\tilde{S}_i \geq t\}$. Note that $\delta_i^D(t)$ is always observable. We replace $\delta_i(t)$ in (3.2)–(3.4) using $\delta_i^D(t)$ with the inverse probability of survival technique (Robins and Rotnitzky 1992; Ghosh and Lin 2002). This is based on the fact that $\delta_i(t)$ and $\delta_i^*(t) \equiv \delta_i^D(t)/\bar{F}(t|Z_{i2})$ have the same conditional expectation when C_i and S_i are conditionally independent conditioning on Z_{i2}, where $\bar{F}(t|Z_{i2}) = P(S_i \geq t|Z_{i2})$. The $\bar{F}(t|Z_{i2})$ can be estimated by the fit of model (2.3). Specifically, let $\hat{\xi}$ be the maximum partial likelihood estimator of ξ and

$$\hat{\Lambda}_d(t) = \sum_{i=1}^n \frac{I\{\tilde{S}_i \leq t\}\Delta_i}{\sum_{j=1}^n \delta_j^D(\tilde{S}_i)e^{\hat{\xi}^T Z_{j2}(\tilde{S}_i)}}$$

be the Breslow estimator of $\Lambda_d(t) = \int_0^t \lambda_d(t)\,dt$ in the model (2.3), then $\hat{\bar{F}}(t;\hat{\xi}|Z_{i2}) = \exp\{-\int_0^t e^{\hat{\xi}^T Z_{i2}(u)}d\hat{\Lambda}_d(u)\}$ is an estimator of $\bar{F}(t|Z_{i2})$ and $\delta_i^*(t)$ can be approximated by $\hat{\delta}_i^*(t) \equiv \delta_i^D(t)/\hat{\bar{F}}(t;\hat{\xi}|Z_{i2})$. Let

$$M_i^*(t;\mathcal{A},\beta,\gamma) = \int_0^t \left[\delta_i^*(s)(Y_i(s) - \beta^T X_i(s))dN_i^*(s) - \delta_i^*(s)\exp\{\gamma^T Z_{i1}(s)\}d\mathcal{A}(s)\right]. \tag{3.6}$$

Then $M_i^*(t;\mathcal{A},\beta,\gamma)$ is still a mean-zero stochastic process. By replacing $\delta_i(t)$ with $\hat{\delta}_i^*(t)$ in (3.2)–(3.4), we obtain the following estimating equations for β, $\mathcal{A}(\cdot)$, and γ:

$$\sum_{i=1}^n \hat{M}_i^*(t;\mathcal{A},\beta,\gamma) = 0, \tag{3.7}$$

$$\sum_{i=1}^{n}\int_{0}^{\infty}W(t)\hat{\delta}_i^*(t)\left[X_i(t)-\hat{\bar{X}}^*(t;\gamma)\right]\left[Y_i(t)-\hat{\bar{Y}}^*(t;\gamma)-\beta^{\mathrm{T}}(X_i(t)-\hat{\bar{X}}^*(t;\gamma))\right]dN_i^*(t)=0, \tag{3.8}$$

$$\sum_{i=1}^{n}\int_{0}^{\infty}\left(Z_{i1}(t)-\hat{\bar{Z}}_1^*(t;\gamma)\right)\hat{\delta}_i^*(t)dN_i^*(t)=0, \tag{3.9}$$

where $\hat{\bar{X}}^*(t;\gamma)=\sum_{i=1}^{n}\hat{\delta}_i^*(t)X_i(t)e^{\gamma^{\mathrm{T}}Z_{i1}(t)}/\sum_{i=1}^{n}\hat{\delta}_i^*(t)e^{\gamma^{\mathrm{T}}Z_{i1}(t)}$,
$\hat{\bar{Y}}^*(t;\gamma)=\sum_{i=1}^{n}\hat{\delta}_i^*(t)Y_i(t)e^{\gamma^{\mathrm{T}}Z_{i1}(t)}/\sum_{i=1}^{n}\hat{\delta}_i^*(t)e^{\gamma^{\mathrm{T}}Z_{i1}(t)}$,
$\hat{\bar{Z}}_1^*(t;\gamma)=\sum_{i=1}^{n}\hat{\delta}_i^*(t)Z_{i1}(t)e^{\gamma^{\mathrm{T}}Z_{i1}(t)}/\sum_{i=1}^{n}\hat{\delta}_i^*(t)e^{\gamma^{\mathrm{T}}Z_{i1}(t)}$ and

$$\hat{M}_i^*(t;\mathcal{A},\beta,\gamma)=\int_0^t\left[\hat{\delta}_i^*(s)(Y_i(s)-\beta^{\mathrm{T}}X_i(s))dN_i^*(s)-\hat{\delta}_i^*(s)\exp\{\gamma^{\mathrm{T}}Z_{i1}(s)\}d\mathcal{A}(s)\right]. \tag{3.10}$$

Denote the solution of (3.9) by $\hat{\gamma}$. Then $\Lambda(t)$ can be estimated by

$$\hat{\Lambda}(t;\hat{\xi},\hat{\gamma})=\sum_{i=1}^{n}\int_0^t\frac{\hat{\delta}_i^*(t)dN_i^*(t)}{\sum_{j=1}^{n}\hat{\delta}_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}}. \tag{3.11}$$

It is noted that an approach similar to (3.9) and (3.11) was used by Ghosh and Lin (2002) in the analysis of recurrent events data with a terminal event.

To summarize, the implementation of parameter estimation can be carried out as follows:

Step 1: Fit model (2.3) with data $\{\tilde{S}_i,\Delta_i,X_i(\cdot)\}$, and obtain estimates $\hat{\xi}$ and $\hat{\Lambda}_d(\cdot)$;
Step 2: Evaluate $\hat{\bar{F}}(t;\hat{\xi})$ and $\hat{\delta}_i^*(t)$ based on the results in step 1, then fit model (2.2) with estimating Eq. (3.9) for γ and with expression (3.11) for $\Lambda(\cdot)$;
Step 3: Plug $\hat{\gamma}$, $\hat{\delta}_i^*(\cdot)$ and $\hat{\Lambda}(t;\hat{\gamma})$ in estimating Eqs. (3.7) and (3.8) and solve for estimates $\hat{\beta}$ and $\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})$. Specifically, the estimating Eq. (3.8) with $\hat{\gamma}$ yields

$$\hat{\beta}=\left[\sum_{i=1}^{n}\int_0^{\infty}W(t)\{X_i(t)-\hat{\bar{X}}^*(t;\hat{\gamma})\}^{\otimes 2}\hat{\delta}_i^*(t)dN_i^*(t)\right]^{-1}$$
$$\times\sum_{i=1}^{n}\int_0^{\infty}W(t)\{X_i(t)-\hat{\bar{X}}^*(t;\hat{\gamma})\}\{Y_i(t)-\hat{\bar{Y}}^*(t;\hat{\gamma})\}\hat{\delta}_i^*(t)dN_i^*(t), \tag{3.12}$$

where $a^{\otimes 2}$=aa^{T}, and the estimating Eq. (3.7) with $\hat{\beta}$ and $\hat{\gamma}$ leads to

$$\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})=\sum_{i=1}^{n}\int_0^t\frac{\{Y_i(s)-\hat{\beta}X_i(s)\}\hat{\delta}_i^*(s)dN_i^*(s)}{\sum_{j=1}^{n}\hat{\delta}_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}}. \tag{3.13}$$

In addition, the variance-covariance matrix for $\hat{\beta}$ is estimated by a plug-in procedure given in the next section, where large sample properties are derived. Note that the estimation of β is done without a direct estimation of the unknown baseline function $\alpha(\cdot)$, as pointed out by a referee.

4 Large sample properties and variance–covariance estimation

To present large sample properties, we introduce more notations. Let

$$M_i^S(t;\xi) = I\{\tilde{S} \le t\}\Delta_i - \int_0^t \delta_i^D(u) e^{\xi^{\mathrm{T}} Z_{i2}(u)} d\Lambda_d(u),$$

$$\mathcal{M}_i(t;\gamma,\xi) = \int_0^t \delta_i^*(u)[dN_i^*(u) - e^{\gamma^{\mathrm{T}} Z_{i1}} d\Lambda(u)].$$

Define $R_{xz}^{(2)}(t;\gamma) = \frac{1}{n}\sum_{i=1}^n \delta_i^*(t) X_i(t) Z_{i1}^{\mathrm{T}}(t) e^{\gamma^{\mathrm{T}} Z_{i1}(t)}$, and for k = 0, 1, 2, define

$$R_{z1}^{(k)}(t;\gamma) = \frac{1}{n}\sum_{i=1}^n \delta_i^*(t) Z_{i1}^{\otimes k}(t) e^{\gamma^{\mathrm{T}} Z_{i1}(t)}$$

$$R_{z2}^{(k)}(t;\xi) = \frac{1}{n}\sum_{i=1}^n \delta_i^D(t) Z_{i2}^{\otimes k}(t) e^{\xi^{\mathrm{T}} Z_{i2}(t)}$$

$$R_{x}^{(k)}(t;\gamma) = \frac{1}{n}\sum_{i=1}^n \delta_i^*(t) X_i^{\otimes(k)}(t) e^{\gamma^{\mathrm{T}} Z_{i1}(t)}$$

and denote the expectation of $R_{z1}^{(k)}(t;\gamma)$, $R_{z2}^{(k)}(t;\xi)$, $R_x^{(k)}(t;\gamma)$ and $R_{xz}^{(2)}(t;\gamma)$ as $r_{z1}^{(k)}(t;\gamma), r_{z2}^{(k)}(t;\xi), r_x^{(k)}(t;\gamma)$ and $r_{xz}^{(2)}(t;\gamma)$ respectively for $k = 0, 1, 2$. Moreover, let $\bar{x}(t;\gamma)$ be the limit of $\bar{X}(t;\gamma)$, $\bar{y}(t;\gamma)$ be the limit of $\bar{Y}(t;\gamma)$ and $\bar{z}_1(t;\gamma)$ is the limit of $\bar{Z}_1(t;\gamma)$.

In addition to assumptions of (3.5) and the conditional independence between C_i and S_i conditioning on Z_{i2}, we impose the following regularity conditions:

C1: $\{\tilde{S}_i,\ \Delta_i,\ N_i^*(t),\ Y_i(t)dN_i^*(t),\ X_i(t),\ t \in [0,\ \tilde{S}_i]\}$, ($i$ = 1, ..., n) are independent and identically distributed;
C2: There exists a constant $\tau > 0$ such that $P(\tilde{S}_i \ge \tau) > 0$;
C3: $N_i^*(\tau \wedge \tilde{S}_i)$ (i = 1,..., n) are bounded by a constant;
C4: $X_i(\cdot)$, (i = 1, ..., n) have bounded total variations, i.e., there exists a fixed constant $K > 0$ such that $\|X_i(0)\| + \int_0^\tau \|dX_i(t)\| \le K$ for all i = 1, ..., n.
C5: $\Omega_\gamma = \int_0^\infty \left[\frac{r_{z1}^{(2)}(u;\gamma)}{r_{z1}^{(0)}(u;\gamma)} - \left(\frac{r_{z1}^{(1)}(u;\gamma)}{r_{z1}^{(0)}(u;\gamma)}\right)^{\otimes 2}\right] r_{z1}^{(0)}(u;\gamma) d\Lambda(u)$ is positive definite;
$\Omega_\xi = \int_0^\infty \left[\frac{r_{z2}^{(2)}(u;\xi)}{r_{z2}^{(0)}(u;\xi)} - \left(\frac{r_{z2}^{(1)}(u;\xi)}{r_{z2}^{(0)}(u;\xi)}\right)^{\otimes 2}\right] r_{z2}^{(0)}(u;\xi) d\Lambda_d(u)$ is positive definite;
$D = E\left[\int_0^\infty W(t)[X_1(t) - \bar{x}(t,\gamma)]^{\otimes 2} \delta_1^*(t) dN_1^*(t)\right]$ is positive definite.
C6: $W(\cdot)$ can be written as a difference of two monotone functions, each of which converges to a deterministic function as $n \to \infty$, and $w(\cdot)$ is the limit of $W(\cdot)$.

The assumptions (C1)–(C5) are similar to those of Andersen and Gill (1982, Theorem 4.1), Lin et al. (2000), Lin and Ying (2001) and Ghosh and Lin (2002). The assumption (C6) is similar to that in appendix A.1 of Lin and Ying (2001).

The following theorem gives the usual asymptotic properties of $\hat{\gamma}$ and $\hat{\beta}$.

Theorem 1 *Suppose that regularity conditions C1–C6 are satisfied.*

(i) *The estimator $\hat{\gamma}$ is strongly consistent and asymptotically zero-mean normal. Actually, $n^{1/2}(\hat{\gamma}-\gamma) = n^{-1/2}\sum_{i=1}^{n}\psi_i + o_P(1)$, where*

$$\psi_i = \Omega_\gamma^{-1}\left[\int_0^\infty\left(Z_{i1}(t) - \frac{r_{z1}^{(1)}(t;\gamma)}{r_{z1}^{(0)}(t;\gamma)}\right)d\mathcal{M}_i(t;\gamma,\xi) + \int_0^\infty\left\{A\left(Z_{i2}(t) - \frac{r_{z2}^{(1)}(t;\xi)}{r_{z2}^{(0)}(t;\xi)}\right) + \frac{q(t;\xi)}{r_{z2}^{(0)}(t;\xi)}\right\}dM_i^S(t;\xi)\right],$$

$$A = E\left[\int_0^\infty\left(Z_{11}(t) - \frac{r_{z1}^{(1)}(t;\gamma)}{r_{z1}^{(0)}(t;\gamma)}\right)g^{\mathrm{T}}(t,Z_{12};\xi)\Omega_\xi^{-1}d\mathcal{M}_1(t;\gamma,\xi)\right]$$

$$g(t,Z_{12};\xi) = \int_0^t\left[Z_{12}(u) - \frac{r_{z2}^{(1)}(u;\xi)}{r_{z2}^{(0)}(u;\xi)}\right]e^{\xi^{\mathrm{T}}Z_{12}}d\Lambda_d(u)$$

$$q(t;\xi) = E\left[\int_t^\infty\left(Z_{11}(u) - \frac{r_{z1}^{(1)}(u;\gamma)}{r_{z1}^{(0)}(u;\gamma)}\right)\delta_1^*(u)e^{\xi^{\mathrm{T}}Z_{12}(t)}d\mathcal{M}_1^*(u;\gamma,\xi)\right].$$

(ii) *The estimator $\hat{\beta}$ is strongly consistent and asymptotically zero-mean normal. In fact, $n^{1/2}(\hat{\beta}-\beta) = n^{-1/2}D^{-1}\sum_{i=1}^{n}\phi_i + o_P(1)$, where $\phi_i = -H\psi_i + \eta_i + \zeta_i$ with ψ_i being defined in (i), and*

$$H = E\left[\int_0^\infty w(t)\left(\frac{r_{xz}^{(2)}(t;\gamma)}{r^{(0)}(t;\gamma)} - \frac{r_x^{(1)}(t;\gamma)r_{z1}^{(1)}(t;\gamma)^{\mathrm{T}}}{r^{(0)}(t;\gamma)\ r^{(0)}(t;\gamma)}\right)(Y_1(t)-\bar{y}(t;\gamma))\delta_1^*(t)dN_1^*(t)\right]$$

$$\eta_i = \int_0^\infty w(t)(X_i(t)-\bar{x}(t;\gamma))\left[dM_i^*(t;\mathcal{A},\beta,\gamma,\xi) - (\bar{y}(t;\gamma)-\beta^T\bar{x}(t;\gamma))d\mathcal{M}_i^*(t;\gamma,\xi)\right]$$

$$\zeta_i = \int_0^\infty\left[B\left(Z_{i2}(t) - \frac{r_{z2}^{(1)}(t;\xi)}{r_{z2}^{(0)}(t;\xi)}\right) + \frac{q_w(t;\gamma,\xi)}{r_{z2}^{(0)}(t;\xi)}\right]dM_i^S(t;\xi)$$

$$B = E\left[\int_0^\infty w(t)(X_1(t)-\bar{x}(t;\gamma))g^T(t,Z_{12};\xi)\Omega_\xi^{-1} \times\left(dM_1^*(t;\mathcal{A},\beta,\gamma,\xi) - (\bar{y}(t;\gamma)-\beta^T\bar{x}(t;\gamma))d\mathcal{M}_1^*(t;\gamma,\xi)\right)\right]$$

$$q_w(t;\gamma,\xi) = E\left\{\int_t^\infty w(u)(X_1(u)-\bar{x}(u;\gamma))e^{\xi^T Z_{12}(t)} \times\left[dM_1^*(u;\mathcal{A},\beta,\gamma,\xi) - (\bar{y}(u;\gamma)-\beta^T\bar{x}(u;\gamma))d\mathcal{M}_1^*(u;\gamma,\xi)\right]\right\}.$$

Proof of Theorem 1 is given in appendix. Theorem 1 shows that the asymptotic variance of $n^{1/2}(\hat{\beta} - \beta)$ is $D^{-1}VD^{-1}$, where V is the variance of ϕ_1. By evaluating empirical counterparts of Ω_γ, Ω_ξ, A, B, g, q, D, q_w and H, we can obtain an estimate $\hat{\phi}_i$ of ϕ_i, and $D^{-1}VD^{-1}$ can be consistently estimated by $\hat{D}^{-1}\hat{V}\hat{D}^{-1}$, where

$$\hat{D} = \frac{1}{n}\sum_{i=1}^{n}\int_0^\infty W(t)[X_i(t) - \hat{\bar{X}}^*(t,\hat{\gamma})]^{\otimes 2}\hat{\delta}_i^*(t)dN_i^*(t),$$

and $\hat{V} = \frac{1}{n}\sum_{i=1}^{n}\hat{\phi}_i^{\otimes 2}$.

As pointed out by Lin and Ying (2001), the $Y_i(t)$ may not be observed continuously. In that case, $\bar{Y}(t)$ can be approximated by $\bar{Y}^*(t)$ in which $Y_i(t)$ is replaced by an imputed value $Y_i^*(t)$ using some smoothing methods. Similarly, we can approximate $\bar{X}(t)$ by $\bar{X}^*(t)$ with imputed covariate processes.

5 Testing parametric assumption on $\alpha(\cdot)$

In practice, it is important to investigate the form of nonparametric component $\alpha(\cdot)$ since parametric forms are more easily interpretable. Moreover, if there is a strong evidence that $\alpha(t)$ is linear function of t, then one can use a marginal linear model or linear mixed effects model.

We develop a test that can be used to check hypothesis H_0: $\alpha(t) = \alpha_0(t;\theta)$, where $\alpha_0(\cdot;\theta)$ is a prespecified function depending on unknown parameter vector θ. Here, we construct a test by comparing two different estimators of $\mathcal{A}(\cdot)$: $\hat{\mathcal{A}}(\cdot;\hat{\beta},\hat{\xi},\hat{\gamma})$ and $\tilde{\mathcal{A}}(\cdot;\hat{\theta},\hat{\xi},\hat{\gamma})$, where $\hat{\mathcal{A}}(\cdot;\hat{\beta},\hat{\xi},\hat{\gamma})$ is the estimator of the $\mathcal{A}$ in Sect. 3,

$$\tilde{\mathcal{A}}(t;\theta,\hat{\xi},\hat{\gamma}) = \int_0^t \alpha_0(u;\theta)d\hat{\Lambda}(u;\hat{\xi},\hat{\gamma}), \quad (5.1)$$

and $\hat{\theta}$ is a consistent estimator of θ under the null hypothesis H_0. Such an estimator $\hat{\theta}$ can be obtained based on the usual least squares principle which leads to estimating equation $U_n(\theta;\hat{\beta}) = 0$, where

$$U_n(\theta;\hat{\beta}) = \sum_{i=1}^{n}\int_0^\infty \delta_i^*(t)\left[Y_i(t) - \hat{\beta}^{\mathrm{T}}X_i(t) - \alpha_0(t;\theta)\right]\frac{\partial\alpha_0(t;\theta)}{\partial\theta}dN_i^*(t). \quad (5.2)$$

Hereafter, we use $\hat{\theta}$ to denote the root of (5.2).

More notations are needed. When $n \to \infty$, let D_1 be the limit of $-(1/n)[(\partial U_n(\theta;\beta))/(\partial\theta)]$, D_2 be the limit of $(-1/n)[(\partial U_n(\theta;\beta))/(\partial\beta)]$, and $D_x(t)$ be the limit of $-\frac{1}{n}\sum_{i=1}^{n}\int_0^t \frac{\delta_i^*(s)X_i(s)dN_i^*(s)}{R_{z1}^{(0)}(s;\hat{\gamma})}$.

Theorem 2 *Under regularity conditions C1–C6, $n^{1/2}(\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma}))$ converges weakly to a zero-mean Gaussian process. In fact,*

$$n^{1/2}(\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})) = n^{-1/2}\sum_{i=1}^{n}\Gamma_i(t) + o_p(1),$$

where

$$\begin{aligned}
\Gamma_i(t) = & \int_0^t \frac{\epsilon_i(s)[dM_i^*(s;\mathcal{A},\beta,\gamma,\xi) - \alpha_0(s)d\mathcal{M}_i^*(s;\gamma,\xi)]}{r_{z1}^0(s;\gamma)} - H(t)\psi_i \\
& + \int_0^\infty C(t)\left[Z_{i2} - \frac{r_{z2}^{(1)}(s;\xi)}{r_{z2}^{(0)}(s;\xi)}\right]dM_i^S(s) + \frac{q_v(t,s)}{r_{z2}^{(0)}(s;\xi)}dM_i^S(s) \\
& + \left[D_x^T(t) + \int_0^t \left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T d\Lambda(s)D_1^{-1}D_2\right]\phi_i \\
& - \left[\int_0^t \left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T d\Lambda(s)\right]D_1^{-1}\int_0^\infty \left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T \\
& \times \left[dM_i^*(s;\mathcal{A},\beta,\gamma,\xi) - \alpha_0(s)d\mathcal{M}_i^*(s;\gamma,\xi)\right] \\
& - \left[\int_0^t \left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T d\Lambda(s)\right] \\
& \times D_1^{-1}\int_0^\infty \left\{K\left[Z_{i2} - \frac{r_{z2}^{(1)}(s;\xi)}{r_{z2}^{(0)}(s;\xi)}\right] + \frac{q_e(s)}{r_{z2}^{(0)}(s;\xi)}\right\}dM_i^S(s)
\end{aligned} \tag{5.3}$$

with

$$\begin{aligned}
H(t) = & -E\left[\int_0^t \frac{r_{z1}^1(s;\gamma)}{\left(r_{z1}^0(s;\gamma)\right)^2}\delta_1^*(s)\epsilon_1(s)dN_1^*(s)\right] \\
C(t) = & E\left[\int_0^t \frac{g(s;Z_{12};\xi)\left(dM_1^*(s;\mathcal{A},\beta,\gamma,\xi) - \alpha_0(s)d\mathcal{M}_1^*(s;\gamma,\xi)\right)}{r_{z1}^0(s;\gamma)}\Omega_\xi^{-1}\right] \\
q_v(t,s) = & E\left[\int_s^t \frac{dM_1^*(u;\mathcal{A},\beta,\gamma,\xi) - \alpha_0(u)d\mathcal{M}_1^*(u;\gamma,\xi)}{r_{z1}^0(u;\gamma)}\right] \\
K = & E\left[\int_0^\infty \left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T g(s;Z_{12};\xi)\epsilon_1(s)dN_1^*(s)\Omega_\xi^{-1}\right] \\
q_e(t) = & E\left[\int_t^\infty \left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T e^{\xi Z_{12}(t)}\epsilon_1(s)dN_1^*(s)\right].
\end{aligned}$$

A sketch of the proof for Theorem 2 is given in Appendix. The distributions of functionals of $n^{1/2}(\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma}))$ are difficult to evaluate analytically. However, representation expression (5.3) allows us to approximate them by a resampling technique (Lin et al. 1994; Lin and Ying 2001). Specifically, let $G_1,\ldots, G_n$ be a random sample of $N(0,1)$ distribution and $\hat{\Gamma}_i(\cdot)$ be the corresponding empirical quantity of $\Gamma_i(\cdot)$, then the conditional distribution of $n^{-1/2}\sum_{i=1}^{n}\hat{\Gamma}_i(t)G_i$ conditioning on observed data is asymptotically the same as the unconditional distribution of $n^{1/2}(\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})-\tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma}))$. Thus, the distribution of $n^{1/2}(\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})-\tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma}))$ can be approximated by a large number of realizations of $n^{-1/2}\sum_{i=1}^{n}\hat{\Gamma}_i(t)G_i$ by repeatedly generating $G_1,\ldots,G_n$.

Following test statistics can be used for checking H_0:

$$S_1=\sup_{t\geq 0} n^{1/2}|\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})-\tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})|,$$

$$S_2=\int_0^\infty n^{1/2}|\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})-\tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})|dt.$$

Let

$$S_1^*=\sup_{t\geq 0}|n^{-1/2}\sum_{i=1}^{n}\hat{\Gamma}_i(t)G_i|,$$

$$S_2^*=\int_0^\infty |n^{-1/2}\sum_{i=1}^{n}\hat{\Gamma}_i(t)G_i|dt.$$

Then, the critical values of the statistics S_1 and S_2 under H_0 can be obtained by the empirical distribution of S_1^* and S_2^*.

6 Numerical studies

6.1 Simulation studies

Simulation studies were conducted to assess the performance of the proposed method in practical settings. We considered the covariates (X_{1i}, X_{2i}), where X_{1i} was generated from Bernoulli(0.5), X_{2i} from $N(1, 0.5^2)$ and Z_i was generated from $I(X_{2i} \leq 1)N(3,1) + I(X_{2i} \geq 1)N(0.5, 0.5^2)$. The response process $Y_i(t)$ was generated from the model

$$Y_i(t)=\alpha_0(t)+X_{i1}\beta_1+X_{i2}\beta_2+\varepsilon_i(t),$$

where $\varepsilon_i(t) = \rho[Z_i - E(Z_i|X_i)] + \varepsilon_i^*(t)$. The $\varepsilon_i^*(t)$ was generated from Normal distribution with mean ϕ_i and variance 1, where $\phi_i \sim N(0,0.2^2)$, and thus the responses of the same subject are positively correlated. The baseline function $\alpha_0(t)$ was specified as constant, linear and nonlinear functions.

The counting process $N^*_i(\cdot)$ for the observation times was generated from a Poisson process with intensity rate $e^{\gamma X_{1i}}$. Time to the terminal event S_i was generated from the Cox proportional hazards model where the hazard function $\lambda(t|Z_i)=\lambda_0(t)e^{\xi Z_i}$. We also considered the independent censoring C_i and for simplicity, C_i was assumed to be of administrative type, i.e. $C_i = 4$. Therefore, $\tilde{S}_i=\min(S_i,C_i)$. The setting of $\lambda_0(t) = \exp(-3.5)$ and $(\gamma,\xi) = (1,0.5)$

yielded 30% of the average observed death proportion and about 6.2 average observation number per subject.

Recall that the estimation of the marginal association between the responses $Y_i(t)$ and $X_i(t)$ is of major interest. In the simulations, we also considered the original method of Lin and Ying (2001) that ignores the terminal event. Under each setting, 500 runs of simulations were conducted for sample size of 300. Table 1 summarizes the main results of the parameter estimates under various settings. It clearly shows that the proposed estimators are virtually unbiased and the coverage of corresponding confidence intervals are close to the nominal level. The performance of the proposed estimator becomes better for larger sample sizes. As anticipated, the Lin and Ying's method yields consistent estimators when the response process is independent of the terminal event, i.e. $\rho = 0$. However, when such independent assumption is violated, the Lin and Ying's approach leads to biases and yields improper coverage probabilities.

We further conducted more simulations to assess the power of our proposed test statistics S_1 and S_2. The forms for $\alpha_0(t;\theta)$ were chosen to be θ_0 (constant) and $\theta_0 + \theta_1 t$ (linear). Table 2 reports the type I error and power at nominal level 0.05 when sample size is 300. Each resampling was based on 1000 runs. Table 2 shows that the proposed testing procedure for the nonparametric component has type I errors close to the nominal level 0.05 under H_0. Moreover, the test statistics has a reasonable power to detect deviations from the null hypothesis.

Table 1 Summary of estimates of regression coefficients in simulations

$\alpha_0(t)$	ρ	Para	Proposed method				Lin and Ying's method			
			Bias	SSE	SE	CP	Bias	SSE	SE	CP
1	0	β_1	–0.00	0.06	0.06	0.96	–0.00	0.06	0.06	0.95
		β_2	–0.00	0.06	0.06	0.94	0.00	0.05	0.06	0.94
	4	β_1	0.00	0.44	0.48	0.92	–0.00	0.39	0.41	0.94
		β_2	–0.01	0.46	0.46	0.94	–0.38	0.43	0.44	0.84
2	0	β_1	–0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.95
		β_2	–0.00	0.05	0.05	0.96	0.00	0.05	0.05	0.94
	4	β_1	–0.00	0.44	0.48	0.92	–0.00	0.39	0.41	0.94
		β_2	–0.01	0.46	0.46	0.94	–0.38	0.43	0.44	0.84
3	0	β_1	–0.00	0.06	0.06	0.96	–0.00	0.06	0.06	0.95
		β_2	–0.00	0.06	0.05	0.94	0.00	0.05	0.06	0.93
	4	β_1	–0.00	0.44	0.48	0.93	–0.00	0.39	0.42	0.93
		β_2	–0.01	0.46	0.46	0.94	–0.38	0.43	0.44	0.84
4	0	β_1	–0.00	0.06	0.06	0.96	0.00	0.06	0.06	0.95
		β_2	–0.00	0.06	0.06	0.94	0.00	0.06	0.06	0.94
	4	β_1	–0.00	0.44	0.48	0.92	–0.00	0.39	0.42	0.93
		β_2	–0.01	0.46	0.46	0.94	–0.38	0.43	0.44	0.84

Note: 1: $\alpha_0(t)=\theta_0$ (θ_0=1); 2: $\alpha_0(t)=\theta_0+\theta_1 t$ (θ_0=1, θ_1=0.4); 3: $\alpha_0(t) = \theta_0 + \theta_1 \sin(\theta_2 t)$ ($\theta_0 = 1$, $\theta_1 = 1.2$, $\theta_2 = 1$) and 4: $\alpha_0(t) = \theta_0 + \theta_1 \log(\theta_2 t)$ ($\theta_0 = 1$, $\theta_1 = 0.5$, $\theta_2 = 1$); SE is the sampling standard error; SEE is the average of the standard error estimators; CP is the empirical coverage probability of the 95% confidence intervals.

Table 2 Simulation results for hypothesis testing

ρ	Test	H_0	H_a^1	H_a^2	H_a^3
0	S_1	0.044	1	1	1
	S_2	0.044	1	1	1
4	S_1	0.046	0.968	0.974	0.954
	S_2	0.050	0.978	0.972	0.970

Note: H_0: $\alpha_0(t)=\theta_0$ (θ_0=1); H_a^1: $\alpha_0(t)=\theta_0+\theta_1 t$ (θ_0=1, θ_1=0.4); H_a^2 : $\alpha_0(t) = \theta_0 + \theta_1 \sin(\theta_2 t)$ (θ_0 = 1, θ_1 = 1.2, θ_2 = 1) and H_a^3: $\alpha_0(t)=\theta_0 + \theta_1 \log(\theta_2 t)$ (θ_0 = 1, θ_1 = 0.5, θ_2 = 1)

6.2 Real example

We now apply the proposed model to analyze the dataset from the ''Predictor's Cohort'' study. The primary interest is to determine factors associated with daily activity measurement of AD patients. The daily activity was assessed by the frequency of 15 activities judged to be within the capacity of patients with dementia receiving supervision and aid. This included 5 outdoor activities and 10 indoor activities. The 5 outside activities consisted of going outside, going to movies and other forms of entertainment, going to church or synagogue or religious events, going shopping, and going for a ride in a car. The 10 indoor activities included contact with pet, getting together with family, talking to family or friends over the phone, reading or being read to, listening to radio or watching TV, exercising, playing games or puzzles, doing handicrafts, tending to plants or a garden, completing an additional unspecified task judged difficult by caregiver. The outcome measurement was the sum of activity scores.

The time-independent covariates were gender, education level and baseline age, while time-dependent covariates were mMMS test score and residence status. The education level was the number of years of education. The mMMS test score was evaluated at patient's each visit. The range of mMMS test score was 0 to 57. The residence status included home, retirement home, nursing home, hospital, rehabilitation center and others. An AD patient's residence status might vary according to different stage of disease at different time.

The original dataset consists of a total of 252 patients. However, the investigation of HRQoL began late in the study, when only 123 subjects were still alive. Consequently, our analysis is based on the data from the 123 patients. Among the 123 patients, 74 were female and 49 were male. The patient's age ranged from 49 to 83 with a median of 73, education level varied from 1 to 20 with a median of 12. Although patients were supposed to be evaluated every 6 months, the actual observation time deviated substantially from the schedule. There were 70 patients died before the end of study and average number of observations per subject was 5. The Fig. 1 shows the survival curve of the 123 patients.

The survival time was modeled by Cox proportional hazards model (2.3) with all time-independent covariates, the estimates were –0.36 for gender, 0.04 for age and 0.02 for education level, respectively. The observation times were modeled by model (2.2) with covariates gender, age and years of education.

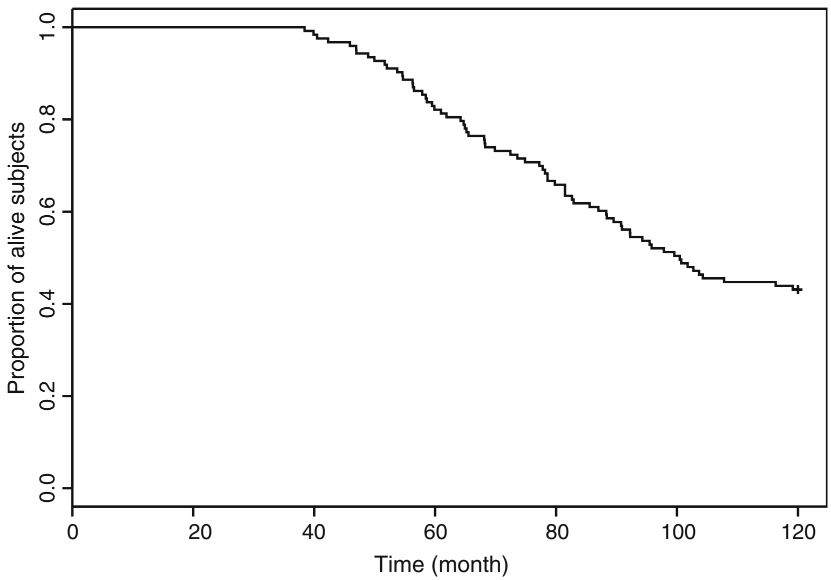

Fig. 1 The Kaplan–Meier survival curve for the 123 AD patients

The coefficient estimates obtained from the estimating Eq. (3.9) were –0.023 for gender, 0.011 for age and 0.002 for education level, respectively. With these estimates and the weight function $W(\cdot) \equiv 1$, we fit following model:

$$Y(t) = \alpha(t) + \beta_1 * mMMS(t) + \beta_2 * Gender + \beta_3 * Age + \beta_4 * Education + \sum_{i=5}^{9} \beta_i * \xi_i(t) + \epsilon(t).$$

where *Gender* takes value 1 for female, 0 for male; $\xi_5(t)$ takes value 1 if patient's residence is home at time t, 0 otherwise; $\xi_6(t)$ takes value 1 if patient's residence is retirement home at time t, 0 otherwise; $\xi_7(t)$ takes value 1 if patient's residence is nursing home at time t, 0 otherwise; $\xi_8(t)$ takes value

Table 3 Regression analysis of "Predictor's Cohort Study"

Parameter	Proposed method			Lin and Ying's method		
	Estimate	SE	95% CI	Estimate	SE	95% CI
mMMS	0.127	0.022	(0.085, 0.170)	0.132	0.021	(0.091, 0.172)
Gender	–0.672	0.661	(–1.968, 0.625)	–0.645	0.632	(–1.883, 0.593)
Age	–0.124	0.038	(–0.198, –0.050)	–0.130	0.035	(–0.199, –0.061)
Education	–0.059	0.083	(–0.221, 0.103)	–0.071	0.076	(–0.221, 0.079)
β_5	–1.579	1.619	(–4.792, 1.595)	–2.005	1.476	(–4.897, 0.887)
β_6	–1.695	2.189	(–5.985, 2.595)	–2.154	2.062	(–6.197, 1.888)
β_7	–5.322	1.665	(–8.586, –2.058)	–5.679	1.512	(–8.641, –2.716)
β_8	–4.439	4.405	(–13.074, 4.195)	–5.534	4.116	(–13.601, 2.533)
β_9	–7.269	1.771	(–10.741, –3.798)	–7.578	1.631	(–10.775, –4.381)

1 if patient's residence is hospital at time t, 0 otherwise; $\xi_9(t)$ takes value 1 if patient's residence is rehabilitation at time t, 0 otherwise. The first 3 columns of Table 3 present the analysis results.

The results clearly show that mMMS score was an significant predictor of daily activity. The higher mMMS score is, the more daily activity is. Age also significantly affected the daily activity. The patients in nursing home and rehabilitation center had less daily activity compared to those in other residence. Gender and education level were not significantly related to the daily activity.

If $\alpha(t;\theta)$ is constant, then $\hat{\Lambda}(t;\hat{\xi},\hat{\gamma})$ and $\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})$ is approximately linearly related. The plot of cumulative number of observation $\hat{\Lambda}(t;\hat{\xi},\hat{\gamma})$ against $\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})$ indicates a trend that the function $\alpha(t;\theta)$ is a constant. Next, we test the hypothesis $H_0{:}\alpha(t)=\theta$ with the statistics $S_1 = \sup_{t\geq 0} n^{1/2}|\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})|$ and $S_2 = n^{1/2}\int_0^\infty |\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})|\, dt$. Under H_0, the estimate of θ is 17.5, the test statistic S_1 = 1.114 with p-value being 0.873, and test statistic S_2 = 11.240 with p-value being 0.965 based on the values of S_1^* and S_2^* in 1000 resampling runs. So both tests show that the null hypothesis that baseline is a constant cannot be rejected.

For comparison, we also fitted data with the approach of Lin and Ying (2001) without modeling the terminal events. With the model (2.2), the estimating equation (3.4) yielded coefficient estimates –0.049 for gender, 0.010 for age and 0.000 for education level, respectively. The last 3 columns of Table 3 present the analysis results.

The results show that the two approaches yield qualitatively similar conclusion. The point estimates of parameters from both methods are similar, however, the estimates of their standard errors with the approach of Lin and Ying are smaller than those with the proposed approach.

7 Discussion

This paper considers the analysis of HRQoL data with semiparametric regression model. The model allows one to examine time-dependent covariate effect as well as the change of outcome variable due to factors that are not included in the model. The examination of covariate effect is done without direct estimation of the unknown baseline function $\alpha(\cdot)$. The model is easy to interpret and exhibits robust properties. Two statistics for testing specific parametric form of the nonparametric component are considered. However, other test statistics can also be constructed, such as von-Mises types of statistics.

The method is applied to the "Predictor's Cohort" study. The analysis using the new approach shows that the cognitive status and age are important predictors for the HRQoL for patients with AD. Older AD patients with poor cognitive status would have worse HRQoL. The gender and education level are not significantly related to HRQoL. Interestingly, the residence of patients is related to HRQoL of AD patients.

An important feature is the joint modeling of survival time and longitudinal measurements. We have used the Cox proportional hazards model for the survival time. Other competing models, such as the accelerated failure time model and the transformation models, may be used as well. In this connection, it is also important to be able to perform model checking by using, for example, residual analysis.

Acknowledgement The authors would like to thank the Associate Editor and two referees for their insightful and constructive comments. This research was supported in part by grants from the National Institutes of Health, the National Science Foundation and the New York City Council Speaker's Fund for Public Health Research.

Appendix

Proof of Theorem 1 (i) Recall that the estimating Eq. (3.9) is used to estimate γ, for convenience, denote

$$U_{n0}(\gamma,\hat{\xi}) = \sum_{i=1}^{n}\int_0^{\infty}\Big(Z_{i1}(t) - \hat{\bar{Z}}_1^*(t;\gamma)\Big)\hat{\delta}_i^*(t)dN_i^*(t).$$

The consistency of $\hat{\gamma}$ follows from the almost identical arguments in Appendix A.1 of Lin et al. (2000). Thus, we omit the details.

Now we sketch the proof of $n^{1/2}(\hat{\gamma}-\gamma) = n^{-1/2}\sum_{i=1}^{n}\psi_i + o_P(1)$. The Taylor series expansion of $U_{n0}(\hat{\gamma},\hat{\xi})$ at $(\gamma,\hat{\xi})$ and the law of large numbers lead to

$$n^{1/2}(\hat{\gamma}-\gamma) = \Omega_{\gamma}^{-1}n^{-1/2}U_{n0}(\gamma,\hat{\xi}) + o_P(1). \tag{A.1}$$

Note that

$$\begin{aligned}U_{n0}(\gamma,\hat{\xi}) =& \sum_{i=1}^{n}\int_0^{\infty}\left(Z_{i1}(t) - \frac{R_{z1}^{(1)}(t;\gamma)}{R_{z1}^{(0)}(t;\gamma)}\right)\delta_i^*(t)d\mathcal{M}_i(t)\\ &+\sum_{i=1}^{n}\int_0^{\infty}\left(Z_{i1}(t) - \frac{R_{z1}^{(1)}(t;\gamma)}{R_{z1}^{(0)}(t;\gamma)}\right)\Big(\hat{\delta}_i^*(t)-\delta_i^*(t)\Big)d\mathcal{M}_i(t) + o_P(1).\end{aligned} \tag{A.2}$$

Following Lin et al. (1994),

$$\begin{aligned}n^{1/2}(\hat{\delta}_i^*(t)-\delta_i^*(t)) =& \delta_i^*(t)n^{-1/2}\sum_{k=1}^{n}\Bigg[\int_0^t\frac{e^{\xi^{\mathrm{T}}Z_{i2}(u)}dM_k^S(u;\xi)}{R_{z2}^{(0)}(u;\xi)}\\ &+\int_0^t e^{\xi^{\mathrm{T}}Z_{i2}(u)}\left(Z_{i2}(u) - \frac{R_{z2}^{(1)}(u;\xi)}{R_{z2}^{(0)}(u;\xi)}\right)^{\mathrm{T}}d\Lambda_d(u)\Omega_{\xi}^{-1}\\ &\times\int_0^{\infty}\left(Z_{k2}(u) - \frac{R_{z2}^{(1)}(u;\xi)}{R_{z2}^{(0)}(u;\xi)}\right)dM_k^S(u;\xi)\Bigg] + o_P(1).\end{aligned} \tag{A.3}$$

Plugging (A.3) and (A.2) into (A.1) and interchanging integrals, we get $n^{1/2}(\hat{\gamma}-\gamma)=n^{-1/2}\sum_{i=1}^{n}\psi_i+o_P(1)$.

(ii) Since $\hat{\gamma}$, $\hat{\xi}$ are consistent, the consistency of $\hat{\beta}$ follows from the expression (3.12) by applying law of large numbers.

Let

$$U_{n1}(\beta,\gamma,\hat{\xi})=\sum_{i=1}^{n}\int_0^{\infty}\Big\{W(t)\hat{\delta}_i^*(t)\Big[X_i(t)-\hat{\bar{X}}^*(t;\gamma)\Big]\\ \times\Big[Y_i(t)-\hat{\bar{Y}}^*(t;\gamma)-\beta^{\mathrm{T}}(X_i(t)-\hat{\bar{X}}^*(t;\gamma))\Big]\Big\}dN_i^*(t),$$

$$U_{n1}(\beta,\gamma,\xi)=\sum_{i=1}^{n}\int_0^{\infty}\big\{W(t)\delta_i^*(t)\big[X_i(t)-\bar{X}^*(t;\gamma)\big]\\ \times\big[Y_i(t)-\bar{Y}^*(t;\gamma)-\beta^{\mathrm{T}}(X_i(t)-\bar{X}^*(t;\gamma))\big]\big\}dN_i^*(t).$$

Clearly, $-\frac{1}{n}\frac{\partial U_{n1}(\beta,\gamma,\hat{\xi})}{\partial\beta}$ converges to D as $n\to\infty$. As in Lin and Ying (2001), $-\frac{1}{n}\frac{\partial U_{n1}(\beta,\gamma,\hat{\xi})}{\partial\gamma}$ converges in probability to H. Thus, the Taylor series expansion of $U_{n1}(\hat{\beta},\hat{\gamma},\hat{\xi})$, and the law of large numbers lead to

$$n^{1/2}(\hat{\beta}-\beta)=n^{-1/2}D^{-1}\Big[U_{n1}(\beta,\gamma,\hat{\xi})-H(\hat{\gamma}-\gamma)\Big]+o_P(1). \tag{A.4}$$

Now

$$n^{-1/2}\Big[U_{n1}(\beta,\gamma,\hat{\xi})-U_{n1}(\beta,\gamma,\xi)\Big]=n^{-1/2}\sum_{i=1}^{n}\int_0^{\infty}W(t)\Big(\hat{\delta}_i^*(t)-\delta_i^*(t)\Big)\big[X_i(t)-\bar{X}^*(t;\gamma)\big]\\ \times\big[Y_i(t)-\bar{Y}^*(t;\gamma)-\beta^{\mathrm{T}}(X_i(t)-\bar{X}^*(t;\gamma))\big]dN_i^*(t)\\ +o_P(1). \tag{A.5}$$

Plugging (A.3) into (A.5) and interchanging integrals with simplification, we get

$$n^{-1/2}U_{n1}(\beta,\gamma,\hat{\xi})=n^{-1/2}\sum_{i=1}^{n}(\eta_i+\zeta_i)+o_P(1). \tag{A.6}$$

Plugging (A.6) and the asymptotic expression of $n^{1/2}(\hat{\gamma}-\gamma)$ in (i) into (A.4), the asymptotic expression of $\hat{\beta}$ follows.

Proof of theorem 2 Note that $n^{1/2}[\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma})-\tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})]$ can be decomposed as follows:

$$n^{1/2}[\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})] = n^{1/2}(I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7), \quad \text{(A.7)}$$

where

$$\begin{aligned}
I_1 &= \hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \hat{\mathcal{A}}(t;\beta,\hat{\xi},\hat{\gamma}),\\
I_2 &= \hat{\mathcal{A}}(t;\beta,\hat{\xi},\hat{\gamma}) - \hat{\mathcal{A}}(t;\beta,\hat{\xi},\gamma),\\
I_3 &= \hat{\mathcal{A}}(t;\beta,\hat{\xi},\gamma) - \hat{\mathcal{A}}(t;\beta,\xi,\gamma),\\
I_4 &= \hat{\mathcal{A}}(t;\beta,\xi,\gamma) - \int_0^t \alpha_0(s,\theta)d\hat{\Lambda}(s;\xi,\gamma),\\
I_5 &= \int_0^t \alpha_0(s,\theta)d\hat{\Lambda}(s;\xi,\gamma) - \int_0^t \alpha_0(s,\theta)d\hat{\Lambda}(s;\hat{\xi},\gamma),\\
I_6 &= \int_0^t \alpha_0(s,\theta)d\hat{\Lambda}(s;\hat{\xi},\gamma) - \tilde{\mathcal{A}}(t;\theta,\hat{\xi},\hat{\gamma}),\\
I_7 &= \tilde{\mathcal{A}}(t;\theta,\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})].
\end{aligned}$$

With the expression (3.13) of $\hat{\mathcal{A}}(t;\beta,\xi,\gamma)$, after some algebra, we get

$$I_1 = \sum_{i=1}^n \int_0^t \frac{X_i^{\mathrm{T}}(s)(\beta-\hat{\beta})\hat{\delta}_i^*(s)dN_i^*(s)}{\sum_{j=1}^n \hat{\delta}_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}} = D_x(t)(\hat{\beta}-\beta) + o_P(\|\hat{\beta}-\beta\|),$$

$$I_2 = \sum_{i=1}^n \int_0^t \frac{(Y_i(s)-\beta^{\mathrm{T}}X_i(s))\sum_{j=1}^n \hat{\delta}_j^*(s)(e^{\gamma^{\mathrm{T}}Z_{j1}(s)} - e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)})\hat{\delta}_i^*(s)dN_i^*(s)}{\sum_{j=1}^n \hat{\delta}_j^*(s)e^{\gamma^{\mathrm{T}}Z_{j1}(s)}\sum_{k=1}^n \hat{\delta}_k^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{k1}(s)}}.$$

Moreover, by the expression (3.11) of $\hat{\Lambda}$ and Taylor series expansions,

$$I_3 = \sum_{i=1}^n \int_0^t \left(Y_i(s)-\beta^{\mathrm{T}}X_i(s)\right)\left[\frac{\hat{\delta}_i^*(t)}{\sum_{j=1}^n \hat{\delta}_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}} - \frac{\delta_i^*(t)}{\sum_{j=1}^n \delta_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}}\right]dN_i^*(t),$$

$$I_4 = \sum_{i=1}^n \int_0^t \frac{\delta_i^*(s)\epsilon_i(s)dN_i^*(s)}{\sum_{j=1}^n \delta_j^*(s)e^{\gamma^{\mathrm{T}}Z_{j1}(s)}} = \sum_{i=1}^n \int_0^t \frac{\delta_i^*(s)[dM_i(s)-\alpha_0(s)d\mathcal{M}_i(s)]}{\sum_{j=1}^n \delta_j^*(s)e^{\gamma^{\mathrm{T}}Z_{j1}(s)}},$$

$$I_5 = \sum_{i=1}^n \int_0^t \alpha_0(s,\theta)\left[\frac{\delta_i^*(t)}{\sum_{j=1}^n \delta_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}} - \frac{\hat{\delta}_i^*(t)}{\sum_{j=1}^n \hat{\delta}_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}}\right]dN_i^*(t),$$

$$I_6 = \sum_{i=1}^n \int_0^t \alpha_0(s,\theta)\hat{\delta}_i^*(s)\left[\frac{1}{\sum_{j=1}^n \hat{\delta}_j^*(s)e^{\gamma^{\mathrm{T}}Z_{j1}(s)}} - \frac{1}{\sum_{j=1}^n \hat{\delta}_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}}\right]dN_i^*(t),$$

$$I_7 = \sum_{i=1}^n \int_0^t \frac{[\alpha_0(s,\theta)-\alpha_0(s,\hat{\theta})]\hat{\delta}_i^*(s)dN_i^*(s)}{\sum_{j=1}^n \hat{\delta}_j^*(s)e^{\hat{\gamma}^{\mathrm{T}}Z_{j1}(s)}}.$$

The consistency of $\hat{\gamma}$ and $\hat{\xi}$ and Taylor series expansion, we can get

$$I_7 = \left[-\sum_{i=1}^{n} \int_0^t \frac{[\frac{\partial \alpha_0(s,\theta)}{\partial \theta}\hat{\delta}_i^*(s) dN_i^*(s)}{\sum_{j=1}^n \hat{\delta}_j^*(s) e^{\hat{\gamma}^{\mathrm{T}} Z_{j1}(s)}} \right] (\hat{\theta} - \theta) + o_P(\|\hat{\theta} - \theta\|),$$

$$I_2 + I_6 = \left[-\int_0^t \frac{\hat{\delta}_i^*(s) \epsilon_i(s) \sum_{j=1}^n \hat{\delta}_j^*(s) e^{\gamma^{\mathrm{T}} Z_{j1}(s)} Z_{j1}(s) dN_i^*(s)}{\left(\sum_{j=1}^n \hat{\delta}_j^*(s) e^{\gamma^{\mathrm{T}} Z_{j1}(s)} \right)^2} \right] (\hat{\gamma} - \gamma) + o_P(\|\hat{\gamma} - \gamma\|)$$
$$= H(t)(\hat{\gamma} - \gamma) + o_P(\|\hat{\gamma} - \gamma\|).$$

In addition,

$$I_3 + I_5 = \sum_{i=1}^{n} \int_0^t \epsilon_i(t) \left[\frac{\delta_i^*(t)}{\sum_{j=1}^n \delta_j^*(s) e^{\hat{\gamma}^{\mathrm{T}} Z_{j1}(s)}} - \frac{\hat{\delta}_i^*(t)}{\sum_{j=1}^n \hat{\delta}_j^*(s) e^{\hat{\gamma}^{\mathrm{T}} Z_{j1}(s)}} \right] dN_i^*(t)$$
$$= \sum_{i=1}^{n} \int_0^t \frac{[\hat{\delta}_i^*(s) - \delta_i^*(s)][dM_i(s) - \alpha_0(s) d\mathcal{M}_i(s)]}{n r_{z2}^{(0)}(s)} + o_P(1).$$

Recall that D_1 is the limit of $(-1/n)[(\partial U_n(\theta;\beta))/(\partial\theta)]$, D_2 is the limit of $(-1/n)[(\partial U_n(\theta;\beta))/(\partial\beta)]$, $D_x(t)$ is the limit of $-\frac{1}{n}\sum_{i=1}^n \int_0^t \frac{X_i(s)\delta_i^*(s) dN_i^*(s)}{R_{z1}(s;\hat{\gamma})}$. With Taylor series expansion, notice that

$$0 = n^{-1/2} U_n(\hat{\theta}; \hat{\beta}, \hat{\xi}) = n^{-1/2} \Big(U_n(\theta; \beta, \hat{\xi}) - D_1 n^{1/2}(\hat{\theta} - \theta) - D_2 n^{1/2}(\hat{\beta} - \beta) \Big) + o_P(1).$$

which leads to,

$$n^{1/2}(\hat{\theta} - \theta) = D_1^{-1} \Big(n^{-1/2} U_n(\theta; \beta, \hat{\xi}) - D_2 n^{1/2}(\hat{\beta} - \beta) \Big) + o_P(1)$$
$$= n^{-1/2} D_1^{-1} \Big(U_n(\theta; \beta, \xi) + U_n(\theta; \beta, \hat{\xi}) - U_n(\theta; \beta, \xi) \Big)$$
$$- D_1^{-1} D_2 n^{1/2}(\hat{\beta} - \beta) + o_P(1).$$

Use the asymptotic representation (A.3) of $\hat{\delta}_i^*(t) - \delta_i^*(t)$ and the fact that

$$U_n(\theta; \beta, \xi) = \sum_{i=1}^{n} \int_0^\infty \frac{\partial \alpha_0(t, \theta)}{\partial \theta} [dM_i(t) - \alpha_0(t) d\mathcal{M}_i(t)]$$

$$U_n(\theta;\beta,\hat{\xi}) - U_n(\theta;\beta,\xi) = \sum_{i=1}^{n} \int_0^\infty \left[K(Z_{i2} - \bar{Z}_{i2}(t)) + \frac{q_e(t)}{r_{z2}^{(0)}(t)} \right] dM_i^S(t) + o_P(n\|\hat{\xi} - \xi\|),$$

we can get $n^{1/2}(\hat{\mathcal{A}}(t;\hat{\beta},\hat{\xi},\hat{\gamma}) - \tilde{\mathcal{A}}(t;\hat{\theta},\hat{\xi},\hat{\gamma})) = n^{-1/2}\sum_{i=1}^n \Gamma_i(t) + o_p(1)$, where

$$
\begin{aligned}
\Gamma_i(t)=&\int_0^t \frac{\epsilon_i(s)[dM_i^*(s;\mathcal{A},\beta,\gamma,\xi)-\alpha_0(s)d\mathcal{M}_i^*(s;\gamma,\xi)]}{r_{z1}^0(s;\gamma)}-H(t)\psi_i \\
&+\int_0^\infty C(t)\left[Z_{i2}-\frac{r_{z2}^{(1)}(s;\xi)}{r_{z2}^{(0)}(s;\xi)}\right]dM_i^S(s)+\frac{q_v(t,s)}{r_{z2}^{(0)}(s;\xi)}dM_i^S(s) \\
&+\left[D_x^T(t)+\int_0^t\left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T d\Lambda(s)D_1^{-1}D_2\right]\phi_i \\
&-\left[\int_0^t\left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T d\Lambda(s)\right]D_1^{-1}\int_0^\infty\left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T \\
&\times\left[dM_i^*(s;\mathcal{A},\beta,\gamma,\xi)-\alpha_0(s)d\mathcal{M}_i^*(s;\gamma,\xi)\right] \\
&-\left[\int_0^t\left(\frac{\partial\alpha_0(s,\theta)}{\partial\theta}\right)^T d\Lambda(s)\right]D_1^{-1}\int_0^\infty\left\{K\left[Z_{i2}-\frac{r_{z2}^{(1)}(s;\xi)}{r_{z2}^{(0)}(s;\xi)}\right]+\frac{q_e(s)}{r_{z2}^{(0)}(s;\xi)}\right\}dM_i^S(s).
\end{aligned}
$$

This completes the proof. □

References

Albert SM, Castillo-Castaneda CD, Jacobs DM, Sano M, Bell K, Merchant C, Small S, Stern Y (1999) Proxy-reported quality of life in Alzheimer's patients: comparison of clinical and population-based samples. J Mental Health Aging 5:49–58.

Albert SM, Castillo-Castaneda CD, Sano M, Jacobs DM, Marder K, Bell K, Bylsma F, Lafleche G, Brandt J, Albert M, Stern Y (1996) Quality of life in patients with Alzheimer's disease as reported by patient proxies. J Am Geriatr Soc 44:1342–1347.

Albert SM, Jacobs DM, Sano M, Marder K, Bell K, Devanand D, Brandt J, Albert M, Stern Y (2000) Longitudinal study of quality of life in people with advanced Alzheimer's disease. Am J Geriatr Psychiatr 9:160–168.

Andersen PK, Gill RD (1982) Cox's regression model for counting processes: a large sample study. Ann Stat 10:1100–1120.

Cole BF, Gelber RD, Goldhirsch A (1993) Cox regression models for quality adjusted survival analysis. Stat Med 12:975–987.

Cole BF, Gelber RD, Gelber S, Mukhopadhyay P (2004) A quality-adjusted survival (Q-TWiST) model for evaluating treatments for advanced stage cancer. J Biopharmaceut Stat 14(1): 111–124.

Cox DR (1972) Regression models and life-tables (with discussion). J Roy Stat Soc Ser B 34:187–220.

Dupuy J, Mesbah M (2002) Joint modeling of event time and nonignorable missing longitudinal data. Lifetime Data Anal 8:99–115.

Fairclough DL (2002) Design and analysis of quality of life studies in clinical trials. Chapman & Hall/CRC, New York, NY.

Fayers PM, Machin D (2000) Quality of life: assessment, analysis and interpretation. John Wiley & Sons, New York, NY.

Ghosh D, Lin DY (2002) Marginal regression models for recurrent and terminal events. Stat Sinica 12:663–688.

Hoover DR, Rice JA, Wu CO, Yang LP (1998) Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. Biometrika 85:809–822.

Lawless JF, Nadeau C (1995) Some simple robust methods for the analysis of recurrent events. Technometrics 37:158–168.

Lin DY, Fleming TR, Wei LJ (1994) Confidence bands for survival curves under the proportional hazards model. Biometrika 81:73–81.
Lin DY, Wei LJ, Yang I, Ying Z (2000) Semiparametric regression for the mean and rate functions of recurrent events. J Roy Stat Soc Ser B 62:711–730.
Lin DY, Ying Z (2001) Semiparametric and nonparametric regression analysis of longitudinal data (with discussion). J Am Stat Assoc 96:103–126.
Moyeed RA, Diggle PJ (1994) Rates of convergence in semiparametric modeling of longitudinal data. Aust J Stat 36:75–93.
Pepe MS, Cai J (1993) Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. J Am Stat Assoc 88:811–820.
Robins J, Rotnitzky A (1992) Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell N, Dietz K, Farewell V (eds) AIDS epidemiology-methodological issues. Birkhauser, Boston, pp 297–331.
Stern Y, Sano M, Paulson J et al. (1987) Modified mini-mental state exam: validity and reliability (abstract). Neurology 37(suppl1): 179.
Zeger SL, Diggle PJ (1994) Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. Biometrics 50:689–699.
Zhao H, Tsiatis AA (1997) A consistent estimator for the distribution of quality adjusted survival time. Biometrika 84:339–348.