

Joint Model with Random Changepoint for Longitudinal Measures and Semi-competing Risks

Xuzhi Wang¹, Martin G. Larson¹, Yorghos Tripodis¹, Michael P. LaValley¹, Chunyu Liu¹

1. Department of Biostatistics, Boston University School of Public Health, Boston, MA, 02118,
United States

Corresponding author: Xuzhi Wang (xwang19@bu.edu)

Abstract

Dementia displays a gradual decline in cognitive abilities, often accompanied by an accelerated cognitive decline preceding diagnosis. Changepoint models are proposed to identify when cognitive decline accelerates and how it progresses. Joint models are developed to further account for dropout due to death or dementia. Cognitive decline in dementia patients may lead to complications that have an impact on their mortality. However, few joint models consider semi-competing risks (i.e., dementia and death) by distinguishing transitions between various health states, i.e., dementia without death, death after dementia, and death without dementia. We proposed a joint model that accounts for both changepoints and semi-competing risks by combining a multivariate random changepoint model for cognitive decline with an illness-death model that estimates health state transitions. We examined the proposed model with two types of random changepoints: one with a smooth change and another with an abrupt change. We also explored a shared random effect structure and a current value structure that connect both longitudinal and survival processes. Two types of cohorts, i.e., a disease cohort and a community cohort, were generated to evaluate the models. Simulation studies showed our proposed models could effectively characterize the influence of the longitudinal process on health state transitions. In addition, the choice of changepoint formulations, association structures, and cohort types impacted model performance. Real data application in the Framingham Heart Study indicated

significant associations between changepoints in cognitive trajectories and health states for dementia and death. Our method provides a flexible framework to integrate longitudinal trajectories with changepoints and semi-competing risks.

1. Introduction

Dementia refers to a category of brain disorders that are severe enough to interfere with daily activities. Alzheimer's disease (AD) is the most prevalent form of dementia. In 2014, approximately 5 million adults aged 65 and older were estimated to have dementia. This number is projected to rise to nearly 14 million by 2060 [1]. People with dementia have problems with memory, attention, communication, reasoning, or visual perception beyond typical age-related changes in vision. Previous studies have shown that cognitive decline can occur before all the criteria for dementia diagnosis are met with considerable individual differences [2]. During the preclinical phase, changes are typically gradual and often challenging to distinguish from the minor declines that accompany normal aging. However, as dementia progresses, cognitive impairments become more apparent, and the rate of functional decline accelerates [3]. Therefore, early detection of cognitive change over the course of aging has been increasingly important in the prevention and treatment of dementia. Various statistical models have been proposed to better understand the shape of cognitive decline over time. These models also aid in detecting when the cognitive trajectories of subjects who develop dementia diverge from normally aging individuals. Linear mixed effects models have been applied to examine the cognitive decline and the risk factors associated with it. However, most of these models assume a constant rate of decline over time, partly due to the lack of sufficient individual-level longitudinal data [4-6]. To account for the nonlinear trend in the cognitive decline preceding dementia onset, mixed effects models with polynomial functions and changepoint models have been developed. Compared to polynomial functions models, changepoint models allow researchers to identify when cognitive decline begins to accelerate [7]. In addition, changepoint models have the flexibility to capture variability in longitudinal cognitive measures as a function of age compared to polynomial

models [8]. As a result, studies using changepoint models are increasingly popular in aging research to estimate cognitive decline acceleration before dementia onset [9].

Changepoint models with various assumptions have been proposed and applied in several medical areas. Hall et al. (2000) proposed a piecewise linear mixed model with a fixed changepoint using a profile likelihood approach to analyze longitudinal cognitive function data [10]. Subsequently, the piecewise model was extended to include a random changepoint based on a Bayesian approach by allowing the changepoint to vary from person to person [7]. Since the piecewise model with an abrupt transition around the changepoint may not accurately reflect real-world situations, changepoint models that enable a more flexible and smooth transition around the changepoint have been proposed. For example, Chiu et al. proposed the bent-cable model that specifies a smooth transition zone by adding a quadratic transition function around the changepoint [11]. Similarly, Hout et al. (2011) developed a smooth polynomial model by using a polynomial function to achieve continuity in the regions around the changepoint [12]. Besides, to utilize the patterns of cognitive decline across different domains, bivariate changepoint models have also been proposed by introducing a correlation structure between two longitudinal measures [7]. Additionally, researchers explored other extensions to handle diverse situations, including the use of multiple random changepoints [13, 14].

In longitudinal cohort studies such as the Framingham Heart Study (FHS), both longitudinal repeated measures and time-to-event data are collected for each participant. For neurodegenerative disease progression, longitudinal measures include cognitive assessments such as neuropsychological (NP) tests and Mini-Mental State Examination (MMSE) scores, and time-to-event data include the date of death or dementia diagnosis [15, 16]. Typically, repeated measures and time-to-event data are modeled separately. For example, linear mixed effects

models or changepoint models are utilized to model cognitive decline, and proportional hazards models are used to model time to dementia or death. However, two issues may arise from the separate analysis: (i) bias can be caused in the longitudinal models because of informative dropout where the probability of dropout depends on the unobserved longitudinal measurements; (ii) proportional hazards models can lead to biased results if the observed longitudinal data with measurement errors are incorporated as time-varying covariates. To tackle these two issues, joint models have been proposed to simultaneously model both longitudinal and time-to-event data [17]. There are two scenarios where joint models are selected: (i) when our research focuses on the longitudinal process and we need to correct for informative dropout; (ii) when the time-to-event process is of interest and there is a need to account for the effect of longitudinal data as endogenous time-dependent covariates measured with error [17, 18]. Joint models with a changepoint in the longitudinal process handle informative dropout due to death or dementia and analyze the association between cognitive decline and dementia. Jacqmin-Gadda et al. (2006) proposed a joint model that incorporates a piecewise mixed model with a random changepoint for cognitive decline and a log-normal model for time to dementia. Both longitudinal and survival processes are linked by a random changepoint [19]. Based on the model proposed in Jacqmin-Gadda et al. (2006), Yu et al. (2010) considered two competing risks, i.e., dementia and dementia-free death. Specifically, the joint model integrated a Bayesian changepoint model for cognition with a mixture survival model for dementia and death [3]. Tapsoba et al. (2011) focused on the proportional hazards model in the time-to-event process of the joint model. Specifically, the joint model employs a random changepoint model for the longitudinal process and simultaneously treats the longitudinal process as a time-varying covariate measured with error [20]. Terrera et al. (2011) compared a series of random effects models in the joint model

framework including a linear model, a quadratic model, a fixed changepoint model, and a random changepoint model, to gain a better understanding of the association between cognitive functions and risk of death [21]. Furthermore, Wang (2021) proposed a joint model framework that includes a random changepoint model with a smooth transition and competing risks [22]. In this previous work, the researcher explored five different formulations of changepoints in longitudinal cognitive measures and addresses interval censoring for competing risks data. However, few joint models in the existing literature consider semi-competing risks (e.g., dementia and death) by distinguishing transitions between various health states, e.g., dementia without death, death following dementia, and death without dementia.

Semi-competing risks data involve two types of events: a non-terminal event and a terminal event. A terminal event (e.g., death) is an event that prevents any subsequent events of other types once a participant experiences it; a non-terminal event (e.g., dementia) is an event that does not prevent subsequent events. If the terminal event occurs first, it censors the non-terminal event; however, if the non-terminal event occurs first, the terminal event can still be observed [23, 24]. In the competing risks setting, participants can experience only one of the several potential terminal events. In contrast, semi-competing risks data provide more information because the terminal event can still be observed even after a non-terminal event has occurred. Therefore, semi-competing risks data allow for the characterization of the dependence between non-terminal and terminal events. Furthermore, semi-competing risks data enable modeling of how the non-terminal event influences the hazard of the terminal event, which can lead to increased prediction precision for the terminal event [25]. One of the widely used approaches in the current literature for the analysis of semi-competing risks data is the illness-death model [26, 27]. The illness-death model differentiates three types of hazards: hazard of

illness, hazard of death without illness, and hazard of death following illness. In the presence of semi-competing risks, participants can transition through a series of health states, e.g., the healthy state, the dementia state, and the death state. The hazard functions in the illness-death model act as forces transitioning individuals between health states. Joint models have been proposed to combine longitudinal models with illness-death models to investigate the effect of longitudinal measures on different transitions between health states. When incorporating changepoints in the longitudinal process, Dantan et al. (2010) proposed a joint multistate model that includes a changepoint and investigates the transition intensities between four health states in the survival process [28]. However, they only considered one longitudinal outcome and assumed identical transition intensities from three transient states to death. Rouanet et al. (2016) developed a joint model combining a mixed effects model, which includes either a quadratic trend or a class-specific fixed changepoint, with an illness-death model accounting for interval censoring [29]. However, this joint model uses a latent class approach without explicitly specifying the relationship between the changepoint and the semi-competing risks.

In this project, we aim to estimate the changepoint during disease progression for dementia while accounting for dropout due to semi-competing risks. Additionally, we explore the impact of the longitudinal process on each of the transitions between health states. Specifically, our proposed joint model is decomposed into two sub-models: a multivariate random changepoint model for the longitudinal data and an illness-death model that estimates the transitions between health states for the semi-competing risks data. We further evaluate two formulations of random changepoints, one characterized by an abrupt change and the other by a smooth change, alongside two types of association structures.

2. Methods

In this section, we propose a joint model for multiple longitudinal measures with changepoints and semi-competing risks. Specifically, the joint model is decomposed into two sub-models: a multivariate random changepoint model for the longitudinal data and an illness-death model that estimates the health transition states for the time-to-event data. Both sub-models are connected through a function of random effects. We consider two formulations of random changepoints in the longitudinal process and two types of association structures in the survival process. We then derive the likelihood function and describe the Bayesian estimation method for the joint model.

2.1 Longitudinal sub-model: random changepoint model

2.1.1 General framework for random changepoint model

In the general framework of changepoint models, we assume that for each participant i and each longitudinal measure k , the longitudinal trajectory is characterized by two linear phases connected by a transition zone. The random changepoint model can be categorized as different types based on the characteristics of the transition zone. Notably, when the transition zone takes the form of an abrupt change, we consider the model as a piecewise (Broken-Stick) model [7, 30]. Otherwise, models that apply smooth functions to the transition zone have been proposed, such as Bacon-Watts model, bent-cable model, and smooth polynomial model [12, 31, 32]. The general framework of changepoint models can be formulated as the following equations [22]

$$E(Y_i(t)) = \begin{cases} f_1(t), & t \text{ before } transition_i \\ f_2(t), & t \text{ during } transition_i \\ f_3(t), & t \text{ after } transition_i \end{cases}$$

where $Y_i(t)$ is the value of a longitudinal measure at time t for participant i , $transition_i$ denotes the period of the transition zone. $f_1(t)$ and $f_3(t)$ are two linear functions representing the two linear phases before and after the transition zone. $f_2(t)$ represents the trajectory function

during the transition zone. We consider two types of random changepoint models, each defined by a unique formulation of the transition zone.

2.1.2 Piecewise (Broken-Stick) model

Let $Y_{ki}(t_{ij})$ be the value of the k^{th} longitudinal measure at time t_{ij} for subject i , where $i = 1, 2, \dots, N$, $k = 1, 2, \dots, K$, and $j = 1, 2, \dots, n_{ki}$. N denotes the total number of participants, and K denotes the total number of longitudinal measures. n_{ki} represents the number of repeated measurements of the k^{th} longitudinal measure for participant i . Let Z_i be a baseline covariate for participant i . The piecewise model can be formulated as

$$\begin{aligned} Y_{ki}(t_{ij}) &= m_{ki}(t_{ij}) + \epsilon_{ki}(t_{ij}) \\ &= \beta_{k0i} + \beta_{k1i}t_{ij} + \beta_{k2i}(t_{ij} - \tau_{ki})I(t_{ij} - \tau_{ki}) + \epsilon_{ki}(t_{ij}) \end{aligned} \quad (1)$$

where $m_{ki}(t_{ij})$ denotes the true unobserved value of $Y_{ki}(t_{ij})$. The random error $\epsilon_{ki}(t_{ij})$ is distributed as $N(0, \sigma_k^2)$. We assume $\epsilon_{ki}(t_{ij})$ is independent of other parameters. $I(\cdot)$ denotes the indicator function such that $Y_{ki}(t_{ij}) = \beta_{k0i} + \beta_{k1i}t_{ij} + \epsilon_{ki}(t_{ij})$ when $t_{ij} \leq \tau_{ki}$, and $Y_{ki}(t_{ij}) = \beta_{k0i} + \beta_{k1i}t_{ij} + \beta_{k2i}(t_{ij} - \tau_{ki}) + \epsilon_{ki}(t_{ij})$ when $t_{ij} > \tau_{ki}$. τ_{ki} is the random changepoint for subject i and longitudinal measure k . β_{k0i} represents the subject-specific intercept, β_{k1i} is the subject-specific slope before the changepoint τ_{ki} , and β_{k2i} denotes the difference in slope before and after the changepoint τ_{ki} .

Each of the K random changepoints τ_{ki} included a population-level fixed effect, which may depend on a baseline covariate, and an individual-level random effect,

$$\tau_{ki} = \beta_{k\tau 0} + \beta_{k\tau 1}Z_i + b_{\tau ki} \quad (2)$$

where $\beta_{k\tau 0}$ is interpreted as the mean of random changepoint when $Z_i = 0$, and $\beta_{k\tau 1}$ is the effect of covariate Z_i on the mean of random changepoint. $b_{\tau ki}$ is the individual-level random effect.

We assume the vector of individual-level random effects in changepoints across K longitudinal

measures, denoted as $\mathbf{b}_{\tau i} = (b_{\tau 1i}, b_{\tau 2i}, \dots, b_{\tau Ki})'$, follows a multivariate normal distribution with mean zero and a $K \times K$ variance–covariance matrix Σ_{τ} :

$$\mathbf{b}_{\tau i} = (b_{\tau 1i}, b_{\tau 2i}, \dots, b_{\tau Ki})' \sim MVN(0, \Sigma_{\tau}) \quad (3)$$

Similarly, the subject-specific intercepts and slopes in each of the K longitudinal measures can be decomposed into a fixed part and a random part

$$\begin{pmatrix} \beta_{k0i} \\ \beta_{k1i} \\ \beta_{k2i} \end{pmatrix} = \begin{pmatrix} \beta_{k0} \\ \beta_{k1} \\ \beta_{k2} \end{pmatrix} + \begin{pmatrix} b_{k0i} \\ b_{k1i} \\ b_{k2i} \end{pmatrix} \quad (4)$$

where β_{k0} , β_{k1} , and β_{k2} represent the fixed intercept, slope before changepoint and slope difference before and after changepoint. b_{k0i} , b_{k1i} , and b_{k2i} denote the corresponding individual-level random effects associated with the fixed effects. Let $\mathbf{b}_{ki} = (b_{k0i}, b_{k1i}, b_{k2i})'$ be the vector of random intercepts and slopes for the k th longitudinal measure, then the vector of \mathbf{b}_{ki} across K longitudinal measures, denoted as $\mathbf{b}_i = (\mathbf{b}_{1i}, \mathbf{b}_{2i}, \dots, \mathbf{b}_{Ki})'$, follows a multivariate normal distribution with mean zero and a $3K \times 3K$ variance-covariance matrix Σ_b

$$\mathbf{b}_i = (\mathbf{b}_{1i}, \mathbf{b}_{2i}, \dots, \mathbf{b}_{Ki})' \sim MVN(0, \Sigma_b) \quad (5)$$

The multivariate piecewise changepoint model assumes that all K longitudinal measures are correlated via the random changepoints, as well as through the random intercepts and slopes. The random changepoints are further assumed to be independent of the random intercepts and slopes. The model gives us a straightforward interpretation for the slope parameters. However, it's important to note that the abrupt change in trajectory may not always align with real-world scenarios, where changes often occur more gradually, such as cognitive decline preceding dementia onset [14]. A visualization of the piecewise model is shown in **Figure 1**.

2.1.3 Bent-Cable model

As an alternative to the piecewise model, the Bent-Cable model assumes a smooth transition between two linear phases [32]. Let $Y_{ki}(t_{ij})$ be the value of the k^{th} longitudinal measure at time t_{ij} for subject i , the model can be formulated as

$$\begin{aligned} Y_{ki}(t_{ij}) &= m_{ki}(t_{ij}) + \epsilon_{ki}(t_{ij}) \\ &= \beta_{k0i} + \beta_{k1i}t_{ij} + \beta_{k2i}q(t_{ij}, \tau_{ki}, \gamma_k) + \epsilon_{ki}(t_{ij}) \end{aligned} \quad (6)$$

where the smooth transition function $q(t, \tau_{ki}, \gamma_k)$ is expressed as

$$\begin{aligned} q(t_{ij}, \tau_{ki}, \gamma_k) &= \frac{(t_{ij} - \tau_{ki} + \gamma)^2}{4\gamma} I(\tau_{ki} - \gamma_k < t_{ij} \leq \tau_{ki} + \gamma_k) \\ &\quad + (t_{ij} - \tau_{ki}) I(t > \tau_{ki} + \gamma_k) \end{aligned} \quad (7)$$

with $I(\cdot)$ as an indicator function such that equation (6) can be rewritten as

$$\begin{aligned} &Y_{ki}(t_{ij}) \\ &= \begin{cases} \beta_{k0i} + \beta_{k1i}t_{ij} + \epsilon_{ki}(t_{ij}), & t_{ij} \leq \tau_{ki} - \gamma_k \\ \beta_{k0i} + \beta_{k1i}t_{ij} + \beta_{k2i} \frac{(t_{ij} - \tau_{ki} + \gamma)^2}{4\gamma} + \epsilon_{ki}(t_{ij}), & \tau_{ki} - \gamma_k < t_{ij} \leq \tau_{ki} + \gamma_k \\ \beta_{k0i} + \beta_{k1i}t_{ij} + \beta_{k2i}(t_{ij} - \tau_{ki}) + \epsilon_{ki}(t_{ij}), & t_{ij} > \tau_{ki} + \gamma_k \end{cases} \end{aligned} \quad (8)$$

In the smooth transition function (7), γ_k represents the half width of the transition zone around the changepoint τ_{ki} . Using this transition function, we define the transition zone as the interval $[\tau_{ki} - \gamma_k, \tau_{ki} + \gamma_k]$. All the other parameters share the same definitions and interpretations as those in the piecewise model. Specifically, β_{k0i} , β_{k1i} , and β_{k2i} are subject-specific intercept, slope before the changepoint, and slope difference before and after the changepoint. τ_{ki} is the subject-specific random changepoint, and $\epsilon_{ki}(t)$ is the random error term. The random effects follow the same distributions as defined in the piecewise model.

When γ_k approaches the value of 0, the smooth transition zone shifts to an abrupt change at the changepoint, which closely approximates the Bent-Cable model to a piecewise model. The

Bent-Cable model adds more flexibility to the piecewise model by defining a smooth transition zone. Both the Bent-Cable model and the piecewise model share the same interpretation for the intercept, slopes and random changepoint, which facilitates the comparison of these two models. **Figure 1** illustrates the visualization of both models, clearly depicting their respective features and differences.

2.2 Survival sub-model: illness-death model

During follow-up, a participant may undergo one or more of three health states (**Figure 2**): healthy state (state 0), dementia state (state 1), and death state (state 2). For example, dementia is a non-terminal event and death is a terminal event. Since the occurrence of the non-terminal event dementia is subject to the terminal event death, they are considered as semi-competing risks [24]. Let T_{1i} and T_{2i} be the times to the non-terminal and terminal events. In the survival part of our joint model framework, we use an illness-death model [26, 27]. The model is designed to capture the dynamics of transitioning between the three distinct health states via three hazard functions (**Figure 2**): $h_1(t_1)$ denotes the hazard of the non-terminal event dementia at time t_1 given that neither dementia nor death occurred before time t_1 , $h_2(t_2)$ represents the hazard of the terminal event death at time t_2 given that neither events occurred before time t_2 , and $h_3(t_2|t_1)$ is the hazard of death at time t_2 conditional on the occurrence of dementia at time t_1 . Specifically, we have

$$\begin{aligned}
 & h_1(t_1|X_{1i}, b_i, \tau_i) \\
 &= \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} P(T_1 \in [t_1, t_1 + \Delta) | T_1 \geq t_1, T_2 \geq t_1, X_{1i}, W_1(\mathbf{b}_i, \boldsymbol{\tau}_i, t_1)), t_1 > 0
 \end{aligned} \tag{9}$$

$$\begin{aligned}
 & h_2(t_2|X_{2i}, b_i, \tau_i) \\
 &= \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} P(T_2 \in [t_2, t_2 + \Delta) | T_1 \geq t_2, T_2 \geq t_2, X_{2i}, W_2(\mathbf{b}_i, \boldsymbol{\tau}_i, t_2)), t_2 > 0
 \end{aligned} \tag{10}$$

$$\begin{aligned}
 & h_3(t_2|t_1, X_{3i}, b_i, \tau_i) \\
 & = \lim_{\Delta \rightarrow 0} \frac{1}{\Delta} P(T_2 \in [t_2, t_2 + \Delta) | T_1 = t_1, T_2 \geq t_2, X_{3i}, W_3(\mathbf{b}_i, \boldsymbol{\tau}_i, t_2)), t_2 > t_1 > 0
 \end{aligned} \tag{11}$$

where X_{gi} is a transition-specific baseline covariate associated with each of the three hazards, $g=1, 2, 3$. W_g is a transition-specific function of individual-level random effects (i.e., random intercepts and slopes \mathbf{b}_i and random changepoints $\boldsymbol{\tau}_i$) from the longitudinal process, which defines the association structure between the longitudinal sub-model and the survival sub-model in our proposed joint model framework. Besides, W_g also captures the dependence structure between T_1 and T_2 . Different association structures are described in more details in the next section. To examine the effects of covariates and functions of random effects on the three hazards, we explicitly specified three parametric proportional hazards models:

$$h_1(t_1|X_{1i}, b_i, \tau_i) = h_{01}(t_1) \exp(\alpha_{10}X_{1i} + W_1(\mathbf{b}_i, \boldsymbol{\tau}_i, t_1)^T \boldsymbol{\alpha}_1), t_1 > 0 \tag{12}$$

$$h_2(t_2|X_{2i}, b_i, \tau_i) = h_{02}(t_2) \exp(\alpha_{20}X_{2i} + W_2(\mathbf{b}_i, \boldsymbol{\tau}_i, t_2)^T \boldsymbol{\alpha}_2), t_2 > 0 \tag{13}$$

$$h_3(t_2|t_1, X_{3i}, b_i, \tau_i) = h_{03}(t_2|t_1) \exp(\alpha_{30}X_{3i} + W_3(\mathbf{b}_i, \boldsymbol{\tau}_i, t_2)^T \boldsymbol{\alpha}_3), t_2 > t_1 > 0 \tag{14}$$

in which h_{0g} is the transition-specific Weibull baseline hazards, $g=1,2,3$. γ_g is the transition-specific log-hazard ratio regression parameters associated with the baseline covariates X_{gi} . $\boldsymbol{\alpha}_g$ is a transition-specific vector of association coefficients for the dependence structure between two processes, which quantifies the impact of the dynamics of the longitudinal measures on the hazards. We further assume a semi-Markov model for the conditional baseline hazard function in Equation (14) such that $h_{03}(t_2|t_1) = h_{03}(t_2 - t_1)$ [24]. The semi-Markov model assumes that the conditional baseline hazard for death at time t_2 is dependent on the time from dementia to death given the onset of dementia at time t_1 , which is also called the ‘sojourn time’. Another parametrization for the conditional baseline hazard is the Markov model, which sets

$h_{03}(t_2|t_1) = h_{03}(t_2)$. The Markov model posits that the hazard for death after dementia onset is independent of the time of dementia t_1 . The semi-Markov model offers an advantage over the Markov model by incorporating t_1 into its hazard function. This inclusion may allow the semi-Markov model to more effectively capture the relationship between the onset of dementia and the risk of death [33].

2.3 Association structures

To characterize the relationship between the longitudinal process and survival process, the association function W_g for $g \in \{1,2,3\}$, can be formulated in various ways, which allows for a more tailored and flexible interpretation for the interplay between the two processes. We consider the following two different association structures in our joint model framework: the current value association structure and the shared random effect association structure [17, 18].

2.3.1 Current value association structure

Under the current value parametrization, the association function is equal to the vector of true unobserved values of the longitudinal measures, i.e.,

$$W_g(\mathbf{b}_i, \boldsymbol{\tau}_i, t) = \mathbf{m}_i(t) = (m_{1i}(t), m_{2i}(t), \dots, m_{Ki}(t))^T \quad (15)$$

Equations (12), (13), and (14) can be rewritten as

$$h_1(t_1|X_{1i}, b_i, \tau_i) = h_{01}(t_1) \exp(\alpha_{10}X_{1i} + \sum_{k=1}^K \alpha_{1k}m_{ki}(t_1)), t_1 > 0 \quad (16)$$

$$h_2(t_2|X_{2i}, b_i, \tau_i) = h_{02}(t_2) \exp(\alpha_{20}X_{2i} + \sum_{k=1}^K \alpha_{2k}m_{ki}(t_2)), t_2 > 0 \quad (17)$$

$$h_3(t_2|t_1, X_{3i}, b_i, \tau_i) = h_{03}(t_2 - t_1) \exp(\alpha_{30}X_{3i} + \sum_{k=1}^K \alpha_{3k}m_{ki}(t_2)), t_2 > t_1 > 0 \quad (18)$$

where α_{gk} ($g = 1,2,3, k = 1,2, \dots, K$) is a measure of the strength of the association between the true unobserved value of the k^{th} longitudinal measure and the hazards for an event at the same time. Take the transition from the healthy state to the dementia state (Equation (16)) as an example, each one-unit increase on the current value of $m_{1i}(t_1)$ at time t_1 is associated with an

$\exp(\alpha_{11})$ -fold increase in the hazards for dementia at the same time t_1 , on condition that neither of dementia nor death occurred before time t_1 . It's important to note that although we utilize a semi-Markov model for conditional hazard function, we assume the true value of the longitudinal measure at time t_2 , denoted as $m_{ki}(t_2)$, is associated with $h_3(t_2|t_1)$ due to the lack of interpretability for $m_{ki}(t_2 - t_1)$. This particular association structure is often employed in scenarios where the primary focus is on survival time and the longitudinal process is considered a time-dependent covariate, which is measured with error [17].

2.3.2 Shared random effect association structure

In real-world scenarios, it's not always realistic to assume a current value association structure. For example, the hazard for dementia may depend on subject-specific deviations from the average intercept, slope, or changepoint of the longitudinal sub-model. Therefore, the shared random effect association structure posits that the hazard for an event is dependent on the random intercepts, slopes and changepoints of the longitudinal sub-model. We have

$$W_g(\mathbf{b}_i, \boldsymbol{\tau}_i, t) = (\tau_{1i}, \mathbf{b}_{1i}, \tau_{2i}, \mathbf{b}_{2i}, \dots, \tau_{Ki}, \mathbf{b}_{Ki})^T \quad (19)$$

Equations (12), (13), and (14) can be rewritten as

$$h_1(t_1|X_{1i}, b_i, \tau_i) = h_{01}(t_1) \exp(\alpha_{10}X_{1i} + \sum_{k=1}^K \zeta_{1k} \tau_{ki} + \sum_{k=1}^K \boldsymbol{\alpha}_{1k} \mathbf{b}_{ki}), t_1 > 0 \quad (20)$$

$$h_2(t_2|X_{2i}, b_i, \tau_i) = h_{02}(t_2) \exp(\alpha_{20}X_{2i} + \sum_{k=1}^K \zeta_{2k} \tau_{ki} + \sum_{k=1}^K \boldsymbol{\alpha}_{2k} \mathbf{b}_{ki}), t_2 > 0 \quad (21)$$

$$h_3(t_2|t_1, X_{3i}, b_i, \tau_i) = h_{03}(t_2 - t_1) \exp(\alpha_{30}X_{3i} + \sum_{k=1}^K \zeta_{3k} \tau_{ki} + \sum_{k=1}^K \boldsymbol{\alpha}_{3k} \mathbf{b}_{ki}), t_2 > t_1 \quad (22)$$

> 0

where τ_{ki} and \mathbf{b}_{ki} are two parameters shared with the longitudinal process: τ_{ki} is the random changepoint and \mathbf{b}_{ki} is the vector of random intercept and slope. ζ_{gk} and $\boldsymbol{\alpha}_{gk}$ represent the association strength parameters between random effects and the hazard during each of the transitions between health states. Let's consider the transition from the healthy state to the

dementia state as exemplified in Equation (20) and assume $\zeta_{1k} < 0$ and $\alpha_{1k} < 0$. The shared random effect association structure suggests that individuals who encounter a changepoint at a later time, or exhibit a higher baseline value, or demonstrate a steeper slope in the k th longitudinal measure, are at a lower risk of developing dementia at time t_1 relative to the average levels observed in the population. Typically, this structure is utilized in scenarios where the primary focus is the longitudinal process with informative dropout, or in situations where both the longitudinal and survival processes are of equal interest [17].

2.4 The joint model framework

If we assume there are two longitudinal measures ($K = 2$), the structure of our proposed joint model framework is displayed in **Figure 3**. The random effects including the random changepoints (τ_{1i} and τ_{2i}), intercepts and slopes (\mathbf{b}_{1i} and \mathbf{b}_{2i}) are crucial in constructing the joint model in the following three aspects. First, the correlation between the longitudinal measures is established through these random effects. Second, the longitudinal process and the survival process are linked by the random effects. Specifically, the two processes are linked directly in the shared random effect model through the random effects, while in the current value model, the random effects have an indirect effect on the hazards via the true values of longitudinal measures $m_{1i}(t)$ and $m_{2i}(t)$. Third, the dependence structure between the non-terminal and the terminal events are captured by the random effects.

2.5 Estimation of the joint model

2.5.1 Likelihood

The following data are observed for each participant i . Let $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{iK})$ be the vector of K longitudinal measures. Let C_i be the right-censoring time for times to events. T_{1i} and T_{2i} are the times to the non-terminal and terminal events. $L_{1i} = \min \{C_i, T_{1i}, T_{2i}\}$, $L_{2i} = \min \{C_i, T_{2i}\}$,

$\Delta_{1i} = I\{T_{1i} \leq \min(T_{2i}, C_i)\}$ denotes the indicator of the non-terminal event, and $\Delta_{2i} = I\{T_{2i} \leq C_i\}$ denotes the indicator of the terminal event. Let $\mathbf{D}_i = \{L_{1i}, \Delta_{1i}, L_{2i}, \Delta_{2i}\}$ be the set of observed time to event data. Let $\boldsymbol{\psi}$ be the full parameter vector. $\boldsymbol{\tau}_i$ and \mathbf{b}_i are vectors of random changepoints and random intercepts and slopes, respectively. The likelihood is constructed under the assumption that both the longitudinal and survival processes are assumed independent given the random effects. Therefore, the likelihood conditional on the random effects is as follows:

$$\begin{aligned} & \prod_{i=1}^N p(\mathbf{Y}_i, \mathbf{D}_i | \boldsymbol{\tau}_i, \mathbf{b}_i, \boldsymbol{\psi}) \\ &= \prod_{i=1}^N p(\mathbf{Y}_i | \boldsymbol{\tau}_i, \mathbf{b}_i, \boldsymbol{\psi}) p(\mathbf{D}_i | \boldsymbol{\tau}_i, \mathbf{b}_i, \boldsymbol{\psi}) \end{aligned} \quad (23)$$

For the longitudinal part, we have

$$\begin{aligned} & p(\mathbf{Y}_i | \boldsymbol{\tau}_i, \mathbf{b}_i, \boldsymbol{\psi}) \\ &= \prod_{j=1}^{n_{ki}} \prod_{k=1}^K p(Y_{ki}(t_{ij}) | \tau_{ki}, b_{ki}, \psi_k) \\ &= \prod_{k=1}^K \frac{1}{(2\pi\sigma_k^2)^{\frac{n_{ki}}{2}}} \exp\left(-\frac{\sum_{j=1}^{n_{ki}} (Y_{ki}(t_{ij}) - m_{ik}(t_{ij}))^2}{2\sigma_k^2}\right) \end{aligned} \quad (24)$$

For the survival part, we construct the following likelihood

$$\begin{aligned} & p(\mathbf{D}_i | \boldsymbol{\tau}_i, \mathbf{b}_i, \boldsymbol{\psi}) \\ &= h_1(L_{1i})^{\Delta_{1i}} h_2(L_{1i})^{(1-\Delta_{1i})\Delta_{2i}} h_3(L_{2i} - L_{1i})^{\Delta_{1i}\Delta_{2i}} \\ & \quad \exp(-H_1(L_{1i}) - H_2(L_{1i}) - \Delta_{1i}H_3(L_{2i} - L_{1i})) \end{aligned} \quad (25)$$

where $h_g(t)$ is the hazard function for each of the three transitions in **Figure 2**, $g \in \{1,2,3\}$.

$H_g(t) = \int_0^t h_g(s) ds$ represents the cumulative hazard function. Note that we use Gauss-Kronrod

quadrature with 15 nodes for numerical approximation of integration when the survival sub-

model is a current value model. Specifically, $H_g(t) = \int_0^t h_g(s) ds \approx \frac{t}{2} \sum_{q=1}^Q w_q h(\frac{t(1+s_q)}{2})$, in which w_q is the standardized weight for quadrature node q with $q \in \{1, 2, \dots, Q\}$. We set $Q = 15$. s_q is the location for quadrature node q . The specific individual likelihood functions in the survival part for four specific scenarios (**Figure 4**) are given as follows [27].

Scenario 1. The participant i develop dementia at time L_{1i} and is censored at time C_i prior to death:

$$p(\mathbf{D}_i | \tau_i, \mathbf{b}_i, \boldsymbol{\psi}) = h_1(L_{1i}) \exp(-H_1(L_{1i}) - H_2(L_{1i}) - H_3(C_i - L_{1i})) \quad (26)$$

Scenario 2. The participant i is dead at time L_{2i} following the onset of dementia at time L_{1i} :

$$p(\mathbf{D}_i | \tau_i, \mathbf{b}_i, \boldsymbol{\psi}) = h_1(L_{1i}) h_3(L_{2i} - L_{1i}) \exp(-H_1(L_{1i}) - H_2(L_{1i}) - H_3(L_{2i} - L_{1i})) \quad (27)$$

Scenario 3. The participant i is dead at time L_{1i} without developing dementia:

$$p(\mathbf{D}_i | \tau_i, \mathbf{b}_i, \boldsymbol{\psi}) = h_2(L_{1i}) \exp(-H_1(L_{1i}) - H_2(L_{1i})) \quad (28)$$

Scenario 4. The participant i is censored at time C_i prior to dementia or death:

$$p(\mathbf{D}_i | \tau_i, \mathbf{b}_i, \boldsymbol{\psi}) = \exp(-H_1(C_i) - H_2(C_i)) \quad (29)$$

2.5.2 Bayesian Inference

We use a Bayesian approach based on Markov Chain Monte Carlo (MCMC) methods for the parameter estimation of the joint model since the model involves a large number of parameters with a complicated joint likelihood. For the Bayesian approach, the random effects, including random changepoints, intercepts and slopes, are treated as model parameters. The joint posterior probability distribution is analogous to:

$$p(\boldsymbol{\psi}, \mathbf{b}, \boldsymbol{\tau}) \propto \prod_{i=1}^N p(\mathbf{Y}_i | \tau_i, \mathbf{b}_i, \boldsymbol{\psi}) p(\mathbf{D}_i | \tau_i, \mathbf{b}_i, \boldsymbol{\psi}) p(\mathbf{b}_i | \boldsymbol{\psi}) p(\tau_i | \boldsymbol{\psi}) p(\boldsymbol{\psi}) \quad (30)$$

where $p(\boldsymbol{\psi})$ is the joint prior distribution for the full parameter vector $\boldsymbol{\psi}$. $p(\mathbf{b}_i|\boldsymbol{\psi})$ is the joint distribution for random intercepts and slopes across longitudinal measures, as is defined in Equation (5). $p(\boldsymbol{\tau}_i|\boldsymbol{\psi})$ denotes the joint distribution for random changepoints, as is defined in Equation (3).

We use weakly informative priors, $p(\boldsymbol{\psi})$, on all the model parameters to ensure that the estimates reflect what is observed in the data. The fixed intercepts and slopes in the longitudinal sub-model, i.e., $\beta_{k0}, \beta_{k1}, \beta_{k2}$, are assigned weakly informative normal priors with a large variance. Similarly, the fixed effects in the survival sub-model, denoted as γ_g, ζ_g , and $\boldsymbol{\alpha}_g$ for $g \in \{1,2,3\}$, are imposed with weakly informative normal priors. The fixed effects in the changepoints are assigned bounded uniform priors to prevent estimation of changepoint values from being biologically implausible: $\beta_{k\tau 0} \sim \text{uniform}(a_0, b_0)$, $\beta_{k\tau 1} \sim \text{uniform}(a_1, b_1)$, $\gamma \sim \text{uniform}(a_\gamma, b_\gamma)$. The specification of the parameter values in the priors depends on the data. For the residual standard deviations of the longitudinal measures σ_k , we use a half Cauchy distribution, i.e., $\sigma_k \sim \text{half-Cauchy}(0, 2.5)$.

We decompose each of the random effects variance-covariance matrix, i.e., Σ_τ for random changepoints and Σ_b for random intercepts and slopes, into a covariance matrix and separate standard deviation terms using Cholesky decomposition. This decomposition allows us to assign priors to the correlation matrix and covariance matrix separately [34]. We specify $\Sigma_\tau = \Lambda_\tau \Omega_\tau \Lambda_\tau$ for random changepoints, where Λ_τ is a diagonal matrix containing standard deviation terms for each random changepoint along its diagonal. Ω_τ is a correlation matrix with 1's on the diagonal and correlation terms representing relationships between random changepoints on the off-diagonal. Similarly, we express random intercepts and slopes as $\Sigma_b = \Lambda_b \Omega_b \Lambda_b$, where Λ_b and Ω_b are the diagonal matrix for standard deviation terms and the correlation matrix, respectively. We

employ the “Lewandowski-Kurowicka-Joe (LKJ) priors” [35], i.e., the parameterization of the LKJ correlation matrix density in terms of its Cholesky factor, for estimating the correlation matrices Ω_b and Ω_τ . Additionally, we impose half-Cauchy priors on the standard deviation terms in Λ_b and Λ_τ .

The model parameters are estimated using the R interface of Stan with the R package rstan [36]. Stan utilizes a No-U-Turn Sampler version of Hamiltonian Monte Carlo algorithm for generating the Markov Chain Monte Carlo (MCMC) samples [37]. Compared to other MCMC algorithms such as Gibbs sampling, Stan offers faster convergence and lower autocorrelation between MCMC samples. To assess model convergence, we use the potential scale reduction statistic \hat{R} calculated by Stan [38]. \hat{R} values < 1.1 for all parameters indicate successful model convergence.

Model comparison is evaluated using the Widely Applicable Information Criterion (WAIC) [39], also known as the Watanabe-Akaike for Bayesian model selection. WAIC is a useful tool for estimating pointwise out-of-sample prediction accuracy within the Bayesian framework. It is calculated based on the log-likelihood evaluated at the posterior simulations of the parameter values. Compared to another commonly used information criterion for Bayesian methods, the deviance information criterion (DIC), WAIC has been shown to be more stable and can be considered as an improvement over DIC [40]. A smaller value of WAIC indicates a better model fit.

3. Simulation study design

Our proposed models were evaluated in the simulation study.

3.1 Data generation

We simulated two correlated longitudinal measures ($K = 2$) for 1000 subjects and two semi-competing risks (i.e., dementia and death) in each simulated dataset. We chose the bent-cable model to simulate the two longitudinal measures. Compared to the piecewise model with an abrupt change at the changepoint, the bent-cable model characterized by a smooth change aligns more closely with real-world scenarios in practice. Moreover, the bent-cable model can be approximated to a piecewise model when the transition zone is small enough. Therefore, the bent-cable model exhibits more flexibility, making it a more comprehensive choice for modeling longitudinal trajectories with changepoints. The two longitudinal measures were simulated from the following models:

$$\begin{aligned} Y_{1i}(t_{ij}) &= m_{1i}(t_{ij}) + \epsilon_{1i}(t_{ij}) \\ &= (\beta_{10} + b_{10i}) + \beta_{11}t_{ij} + (\beta_{12} + b_{12i})q(t_{ij}, \tau_{1i}, \gamma_1) + \epsilon_{1i}(t_{ij}) \end{aligned} \quad (31)$$

$$\begin{aligned} Y_{2i}(t_{ij}) &= m_{2i}(t_{ij}) + \epsilon_{2i}(t_{ij}) \\ &= (\beta_{20} + b_{20i}) + \beta_{21}t_{ij} + (\beta_{22} + b_{22i})q(t_{ij}, \tau_{2i}, \gamma_2) + \epsilon_{2i}(t_{ij}) \end{aligned} \quad (32)$$

where the forms of $q(t_{ij}, \tau_{1i}, \gamma_1)$ and $q(t_{ij}, \tau_{2i}, \gamma_2)$ were described in Equation (7) from the Methods section. We assumed that $Y_{1i}(t_{ij})$ had a narrow transition zone around the changepoint, mimicking a longitudinal trajectory with a sharp change, while $Y_{2i}(t_{ij})$ was postulated to exhibit a smooth change. The omission of the random slope before the changepoint was due to the flat and homogeneous trajectory pattern observed in real data before the changepoint. Considering the complexity of our proposed joint model, we further assumed that the correlation between the two longitudinal measures were fully captured by the random changepoints. The random intercepts and slopes had the following distribution

$$(b_{10i}, b_{12i}, b_{20i}, b_{22i})^T \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{10}^2 & \rho_{b1}\sigma_{10}\sigma_{12} & 0 & 0 \\ \rho_{b1}\sigma_{10}\sigma_{12} & \sigma_{12}^2 & 0 & 0 \\ 0 & 0 & \sigma_{20}^2 & \rho_{b2}\sigma_{20}\sigma_{22} \\ 0 & 0 & \rho_{b2}\sigma_{20}\sigma_{22} & \sigma_{22}^2 \end{pmatrix} \right) \quad (33)$$

The random changepoints were

$$(\tau_{1i}, \tau_{2i})^T \sim MVN \left(\begin{pmatrix} \beta_{1\tau 0} + \beta_{1\tau 1}Z_i \\ \beta_{2\tau 0} + \beta_{2\tau 1}Z_i \end{pmatrix}, \begin{pmatrix} \sigma_{\tau 1}^2 & \rho_{\tau}\sigma_{\tau 1}\sigma_{\tau 2} \\ \rho_{\tau}\sigma_{\tau 1}\sigma_{\tau 2} & \sigma_{\tau 2}^2 \end{pmatrix} \right) \quad (34)$$

The times to semi-competing risks were generated under two scenarios where two different association structures (i.e., current value and shared random effects) between the longitudinal and survival processes were assumed. For the scenario where the shared random effects association structure was adopted, the time-to-event endpoints were generated based on the following survival sub-model:

$$h_1(t_1|X_i, b_i, \tau_i) = a_1 b_1 t_1^{b_1-1} \exp(\alpha_1 X_i + \zeta_1 \tau_{1i}), t_1 > 0 \quad (35)$$

$$h_2(t_2|X_i, b_i, \tau_i) = a_2 b_2 t_2^{b_2-1} \exp(\alpha_2 X_i + \zeta_2 \tau_{1i}), t_2 > 0 \quad (36)$$

$$h_3(t_2|t_1, X_i, b_i, \tau_i) = a_3 b_3 (t_2 - t_1)^{b_3-1} \exp(\alpha_3 X_i + \zeta_3 \tau_{1i}), t_2 > t_1 > 0 \quad (37)$$

Random intercepts and slopes were excluded due to their insignificant effects on the survival part observed in the real data application and the extended simulation time with increased model complexity. Additionally, only the changepoint in the first longitudinal measure, as defined by τ_{1i} , was included in the survival part to avoid multicollinearity. This choice was made under the assumption that the two changepoints were highly correlated. The specification of shared random effects models with correlated changepoints using competing risks data for dementia and death has been investigated in Wang (2021) [22]. X_i was a binary baseline covariate that had an impact on the hazards for semi-competing risks. We followed the cause-specific hazard (CSH) approach

which was described in Beyersmann et al. (2009) [41] and Bender et al. (2005) [42] to generate the time-to-event endpoints. Specifically, this approach is divided into the following several steps:

Step 1. Let $S_1(t) = \exp(-H_1(t) - H_2(t))$, where $H_1(t)$ and $H_2(t)$ are the cumulative hazard functions for equations (35) and (36). Then we simulated a variable $u_i \sim \text{Uniform}(0,1)$. Next, we obtained T_i^* by solving for the equation $u_i = S_1(t)$. T_i^* was the time to the first event, i.e., dementia or death without dementia.

Step 2. We ran a Bernoulli experiment for T_i^* to decide which event occurred at time T_i^* .

Specifically, u_i^* was generated from $\text{Uniform}(0,1)$. If $u_i^* \leq \frac{h_1(T_i^*)}{h_1(T_i^*) + h_2(T_i^*)}$, T_i^* was denoted as the time to dementia, i.e., $T_i^* = T_{1i}$. Otherwise, T_i^* was considered as the time to death without dementia, i.e., $T_i^* = T_{2i}$.

Step 3. If $T_i^* = T_{1i}$, which indicated a participant developed dementia at time T_{1i} , we further simulated T_{12i} based on the hazard function $h_3(t)$. We then defined C_i as the time to the maximum time of follow-up for each participant. We set the observed outcome information as (time to event 1, time from event 1 to event 2, indicator for dementia, indicator for death), which can be demonstrated in the following scenarios:

- $(T_{1i}, T_{12i}, 1, 1)$, if $T_{1i} + T_{12i} \leq C_i$
- $(T_{1i}, C_i - T_{1i}, 1, 0)$, if $T_{1i} + T_{12i} > C_i$ & $T_{1i} \leq C_i$
- $(T_{2i}, 0, 0, 1)$, if $T_{2i} \leq C_i$
- $(C_i, 0, 0, 0)$, if $T_{1i} > C_i$ or $T_{2i} > C_i$

(38)

In the context where we assumed a current value association structure, the time-to-event endpoints were derived using the survival sub-model outlined below:

$$h_1(t_1|X_i, b_i, \tau_i) = a_1 b_1 t_1^{b_1-1} \exp(\alpha_1 X_i + \zeta_1 m_{1i}(t_1)), t_1 > 0 \quad (39)$$

$$h_2(t_2|X_i, b_i, \tau_i) = a_2 b_2 t_2^{b_2-1} \exp(\alpha_2 X_i + \zeta_2 m_{1i}(t_2)), t_2 > 0 \quad (40)$$

$$h_3(t_2|t_1, X_i, b_i, \tau_i) = a_3 b_3 (t_2 - t_1)^{b_3-1} \exp(\alpha_3 X_i + \zeta_3 m_{1i}(t_2)), t_2 > t_1 > 0 \quad (41)$$

Like the shared random effects association structure, we only included the true unobserved value for the first longitudinal measure at the same time of the event, denoted as $m_{1i}(t_1)$ and $m_{1i}(t_2)$, to avoid multicollinearity between the two highly correlated longitudinal measures. We followed analogous procedures to those outlined in the shared random effects association structure to generate two semi-competing risks.

3.2 Simulation scenarios

Our simulated dataset mimicked the Offspring cohort in the FHS to allow for reasonable generalization. Note that the offspring cohort in FHS represented a community-based group, predominantly consisting of relatively healthy individuals who entered the cohort at middle-age. Additionally, this cohort was characterized by a low dementia onset rate of less than 10%. Therefore, by varying the changepoints and the rate of dementia onset, we simulated a disease cohort with earlier occurrence of changepoints and a higher dementia rate, along with a community cohort with a similarly low rate of dementia that was observed in the FHS offspring cohort. Within each cohort, two association structures were considered due to the uncertainty about the true association structure in real-world scenarios. The simulation parameters were chosen based on our real data analysis. The main parameter settings in all four scenarios were summarized in **Table 1**. For each scenario, 200 datasets with 1000 subjects were generated and analyzed.

Scenario 1: disease cohort with shared random effects association structure.

We set total subjects $n=1000$ with the baseline time variable t_{i0} generated from a uniform distribution $U(-5, 5)$ to represent the baseline age subtracted by 60. Each participant was followed every two years until the 20th year. The maximum number of follow-up visits was 11. The longitudinal measures were observed until the occurrence of either dementia or death or at the 20th year of the follow-up period. The intercept, slope before the changepoint, and slope difference before and after the changepoint for the first longitudinal measure Y_1 , denoted by $(\beta_{10}, \beta_{11}, \beta_{12})^T$, were $(-0.18, -0.02, -0.15)^T$. The longitudinal parameters for the second longitudinal measure Y_2 were $(\beta_{20}, \beta_{21}, \beta_{22})^T = (-0.03, -0.02, -0.2)^T$. The half widths of the transition zone for two longitudinal measures were $(\gamma_1, \gamma_2)^T = (2, 4)^T$, such that Y_1 had a narrower transition zone while the transition zone for Y_2 was wider. Mean changepoint intercepts, mean changepoint slopes, and standard deviations of changepoints were set as follows: $\beta_{1\tau_0} = 15$, $\beta_{2\tau_0} = 12$, $\beta_{1\tau_1} = -3$, $\beta_{2\tau_1} = -5$, $\sigma_{\tau_1} = 3$, $\sigma_{\tau_2} = 3$. These parameters were chosen to reflect the mean changepoints observed during the follow-up period for each participant. Z_i was a binary variable generated from a Bernoulli distribution with probability of 0.2. Two changepoints were assumed to be correlated with a correlation coefficient $\rho_\tau = 0.8$. The association parameters for the first random changepoint τ_{1i} in the illness-death model were $(\zeta_1, \zeta_2, \zeta_3)^T = (-0.3, -0.1, -0.1)^T$. The baseline covariate X_i was generated from Bernoulli (0.5) with associated coefficients $(\alpha_1, \alpha_2, \alpha_3)^T = (-0.6, -0.6, -0.1)^T$. The event rates were around 30%, 15%, and 20%, respectively, for dementia, death without dementia, and death following dementia. All the other parameters were presented in the Supplemental materials.

Scenario 2: disease cohort with current value association structure.

The parameters in the longitudinal sub-model were the same as those described in Scenario 1. In the survival sub-model, the association parameters for the true observed value of Y_1 was set as

$(\zeta_1, \zeta_2, \zeta_3)^T = (-2.5, -0.6, -0.5)^T$. The coefficients for the baseline covariate X_i were

$(\alpha_1, \alpha_2, \alpha_3)^T = (-0.2, -0.4, -0.2)^T$. The event rates were similar to those in Scenario 1.

Scenario 3: community cohort with shared random effects association structure.

By varying the location and variation of the changepoints along with the time interval between two adjacent visits, we aligned our data more closely with a community cohort, i.e., the FHS offspring cohort. Specifically, participants were followed every 5 years up to a maximum duration of 20 years. We set $\beta_{1\tau_0} = 25$ and $\beta_{2\tau_0} = 27$ to reflect the scenario where the mean locations for two random changepoints were beyond the average follow-up time of 20 years. We also assigned larger variation of two changepoints by assigning $\sigma_{\tau_1} = 8$ and $\sigma_{\tau_2} = 8$. Other parameters were identical to those outlined in Scenario 1. In this scenario, the approximate event rates were 10% for dementia, 10% for death without dementia, and 6% for death following dementia.

Scenario 4: community cohort with current value association structure.

The longitudinal parameters resembled those outlined in Scenario 3 to reflect a community-based cohort, and we adopted the same survival parameter settings from Scenario 2 to capture the relationship between the true unobserved value for Y_1 and the semi-competing risks. Slightly higher event rates were observed compared to Scenario 3: 15% for dementia, 10% for death without dementia, and 10% for death following dementia.

3.3 Analysis models

The selection of an appropriate model is based not only on model criterion values but also on the underlying rationale. Specifically, the piecewise model may be preferred due to its straightforward interpretation of parameters and lower computational burden compared to the bent-cable model. The current value association structure could be selected if the focus is on

understanding the survival process, while the shared random effect association structure is chosen when the focus is correcting longitudinal bias. Therefore, we compared four joint models with different changepoint formulations and association structures using the simulated data. We calculated and compared the percent bias and 95% coverage probabilities for each parameter, along with the WAIC for each model. Below we describe the four joint models to analyze the four simulated scenarios.

Model 1: bent-cable model + shared random effects association structure (BC+RE).

In the longitudinal sub-model of Model 1, we chose the bent-cable model outlined in Equations (31) and (32) for two longitudinal measures. In the survival sub-model of Model 1, the shared random effects association structure defined in Equations (35)-(37) was incorporated. We utilized commonly used weakly informative priors for Bayesian estimation to reflect what was observed in the data. For the fixed intercept and slopes (i.e., $\beta_{10}, \beta_{11}, \beta_{12}, \beta_{20}, \beta_{21},$ and β_{22}), we used normal distributions with zero mean and a large variance of 100, denoted as $N(0, 100)$. For the fixed parameters associated with the transition zone, we utilized uniform distributions to exclude biologically implausible changepoint values. Specifically, the priors were set as follows: $\beta_{1\tau_0}$ and $\beta_{2\tau_0} \sim \text{Uniform}(10, 20)$ for the disease cohort scenarios, $\beta_{1\tau_0}$ and $\beta_{2\tau_0} \sim \text{Uniform}(20, 30)$ for the community cohort scenarios, $\beta_{1\tau_1}$ and $\beta_{2\tau_1} \sim \text{Uniform}(-10, 0)$, γ_1 and $\gamma_2 \sim \text{Uniform}(0, 6)$. As elaborated in the Methods section, each of the random effects variance-covariance matrices, defined as Σ_b and Σ_τ , can be decomposed into a correlation matrix (Ω_b and Ω_τ) and separate standard deviation terms. For the standard deviations of the random intercepts and slopes (i.e., $\sigma_{10}, \sigma_{12}, \sigma_{20}, \sigma_{22}$), we used half-Cauchy distributions with location 0 and scale 2.5. For the standard deviations of the random changepoints (i.e., σ_{τ_1} and σ_{τ_2}), we chose $\text{Uniform}(0, 6)$

(disease cohort) or $Uniform(0,10)$ (community cohort) to provide a reasonable range for the variability of changepoints.

For each of the correlation matrices, we used LKJ correlation priors parameterized in terms of its Cholesky factor with shape parameters equal to 1 or 2. For each of the association parameters in the survival sub-model (i.e., $\alpha_g, \zeta_g; g = 1,2,3$), we employed normal distributions with zero mean and a variance of 100.

Model 2: piecewise model + shared random effects association structure (PW+RE).

Compared to Model 1, Model 2 maintained the same shared random effects association structure but incorporated the following piecewise model in the longitudinal process:

$$\begin{aligned} Y_{1i}(t_{ij}) &= m_{1i}(t_{ij}) + \epsilon_{1i}(t_{ij}) \\ &= (\beta_{10} + b_{10i}) + \beta_{11}t_{ij} + (\beta_{12} + b_{12i})(t_{ij} - \tau_{1i})I(t_{ij} - \tau_{1i}) + \epsilon_{1i}(t_{ij}) \end{aligned} \quad (42)$$

$$\begin{aligned} Y_{2i}(t_{ij}) &= m_{2i}(t_{ij}) + \epsilon_{2i}(t_{ij}) \\ &= (\beta_{20} + b_{20i}) + \beta_{21}t_{ij} + (\beta_{22} + b_{22i})(t_{ij} - \tau_{2i})I(t_{ij} - \tau_{2i}) + \epsilon_{2i}(t_{ij}) \end{aligned} \quad (43)$$

We adopted the same priors specified in Model 1 except for γ_1 and γ_2 , which were absent from Model 2.

Model 3: bent-cable model + current association structure (BC+CV).

In Model 3, we substituted the survival process in Model 1 with the current value association structure specified in Equations (39)-(41). Like Model 1 and Model 2, the association coefficients in the survival part were assigned normal priors with a mean of 0 and a variance of 100.

Model 4: piecewise model + current value association structure (PW+CV).

To construct Model 4, we integrated the piecewise model specified in Equations (42) and (43) into the joint model framework along with the current value association structure. The prior settings were the same as Model 3.

In Bayesian estimation, we obtained 6000 MCMC sample iterations from each of the 3 parallel chains with the first 4000 iterations as a warm-up phase. Convergence was evaluated through the potential scale reduction statistic.

4. Simulation results

We generated our data in four scenarios described in section 3.2 and analyzed the data using four joint models proposed in section 3.3. Then we evaluated the results based on posterior mean, percent bias, 95% coverage probability, and WAIC. Numerical results were presented in **Tables 2-6**. The comparison of point estimates for mean changepoints, denoted as $\beta_{\tau_{10}}$ and $\beta_{\tau_{20}}$, as well as the association parameter ζ_3 in four models was illustrated in **Figures 5-8**. Our primary interests were the slope parameters ($\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$) along with changepoint parameters ($\beta_{1\tau_0}, \beta_{1\tau_1}, \beta_{2\tau_0}, \beta_{2\tau_1}, \sigma_{\tau_1}, \sigma_{\tau_2}, \rho_{\tau}$) in the longitudinal sub-model, as well as the association parameters ($\zeta_1, \zeta_2, \zeta_3$) in the survival sub-model.

Scenario 1: disease cohort with shared random effects association structure (Table 2 and Figure 5)

The simulation results showed that the bent-cable (BC) model with the shared random effects (RE) association structure (BC+RE) generally performed well, with low percent bias $< 5\%$ for most of the parameters and 95% coverage probability rates around the nominal rate. This finding was expected since BC+RE was the generating model. Note that the only parameter with a percent bias $> 5\%$ was ζ_3 (percent bias = 7.26%), which can be attributed to the relatively small effect size combined with the relatively low event rate of death following dementia.

Compared to BC+RE, the piecewise (PW) model with the shared random effects (RE) association structure (PW+RE) produced biased estimates in slopes for the longitudinal process. More specifically, the slope differences before and after the changepoints (β_{12} and β_{22}) were underestimated (percent bias: -4.56% and -11.65%, respectively) in terms of their absolute values with low coverage probabilities (72% and 0%), suggesting more gradual declines after the changepoints were estimated using PW+RE. Y_2 which exhibited a broader transition zone showed greater deviations from the true values compared to Y_1 which contained a narrower transition zone. For changepoint parameters, PW+RE yielded similar estimates to BC+RE with low bias, although the occurrence of the estimated changepoints was slightly earlier compared to the true values, especially for Y_2 (12 in BC+RE vs 11.55 in PW+RE). The estimated association parameters were similar in both models.

Substituting the shared random effects association structure in BC+RE with the current value (CV) association structure (BC+CV) led to an overestimated absolute value of the slope differences for Y_1 (β_{12}) (percent bias: 6.62%), indicating a steeper estimated decrease after the changepoint for the longitudinal measure with a more abrupt transition Y_1 . In addition, the width of the narrow transition zone γ_1 for Y_1 was overestimated by about 21%. The estimated changepoint locations for both longitudinal measures were close to the true values with the slightly later occurrence of changepoints estimated in BC+CV (15.47 vs 15 for Y_1 ; 12.13 vs 12 for Y_2). Since the incorporation of the current value association structure resulted in different parameter interpretations for ζ_1 , ζ_2 , and ζ_3 compared to the shared random effect structure, the bias and coverage probabilities for the survival sub-model in BC+CV were not available. However, the negative associations between the longitudinal and survival processes, introduced by the shared random effects in the simulated data, were still evident through the posterior means

in the three hazard functions within BC+CV. Specifically, a one-unit increase in the true unobserved value for the first longitudinal measure at the time of t_1 , denoted by $m_{1i}(t_1)$, was significantly associated with a 49% decrease (HR=0.51, 95% CI: 0.42 to 0.61) in the hazard for dementia at time t_1 . The estimated values for ζ_2 and ζ_3 demonstrated negative associations between the two processes although they were not significant due to small effect sizes and low event rates.

In contrast to BC+CV, replacing the bent-cable model with the more parsimonious piecewise model (PW+CV) produced underestimated absolute values of the slope after the changepoint (percent bias: -1.49% for β_{12} and -11.32% for β_{22}) and a shift to the left in the locations of the changepoints, especially for Y_2 (11.64 vs 12). All the other parameter estimates remained similar to those in BC+CV.

Scenario 2: disease cohort with current value association structure (Table 3 and Figure 6)

The joint model consisting of the bent-cable model and the current value association structure (BC+CV), whose structure was consistent with the generating model, had the best performance. BC+CV showed small percent bias of $< 5\%$ for all parameters and appropriate coverage probabilities. Similar to what we observed in Scenario 1, replacing the bent-cable model with the piecewise model (PW+CV) induced a more gradual decline following the changepoints and a left shift of the changepoint locations for the longitudinal measures, especially for Y_2 with a wide transition zone. In addition, PW+CV slightly biased the estimated association coefficients for the hazards of semi-competing risks (e.g., 7.04% bias for ζ_3 in PW+CV vs 4.51% bias for ζ_3 in BC+CV), while other association parameters were similar.

Mis-specifying the association structure with the shared random effects association structure (BC+RE) caused an underestimation of the absolute slope difference β_{12} (percent bias:

-5.62%) and the transition width γ_1 (percent bias: -9.03%) for Y_1 . Despite the unbiased estimation for the mean changepoints $\beta_{1\tau_0}$ and $\beta_{2\tau_0}$, the standard deviation σ_{τ_1} for the first changepoint τ_{1i} was overestimated by around 9% and the correlation between the two changepoints ρ_{τ} was underestimated by about 9%. Regardless of the different parametrization in the survival sub-model, the negative associations between the longitudinal and survival processes were evident through the effects of the shared random changepoint. Note that in both Scenarios 1 and 2, although some of the estimates in the mis-specified models deviated in terms of bias and coverage probabilities, the estimates remained acceptable. These estimates reflected the true longitudinal trajectories as well as the associations between the two processes.

According to **Table 6**, the WAIC values showed that the correctly specified joint model with a smooth transition produced the smallest WAIC in the first two scenarios, indicating the best predictive accuracy for model fit.

Scenario 3: community cohort with shared random effects association structure (Table 4 and Figure 7).

In this scenario, the gap between consecutive visits increased to 5 years, and the mean changepoints occurred beyond the average follow-up duration. This resulted in only a portion of participants experiencing the observed changepoints. Additionally, the delayed occurrence of changepoints resulted in reduced event rates, as these changepoints were associated with semi-competing risks. The correctly specified model BC+RE produced overestimated absolute slope difference β_{12} (percent bias: 14%) and transition width γ_1 (percent bias: 48%) for Y_1 , indicating the narrow transition zone cannot be estimated properly due to the large time intervals between visits. In addition, the association estimates in the hazard function for death following dementia were not accurate (percent bias: 16% for ζ_3), but the coverage probabilities were acceptable at

around 95%. Like the previous scenarios, PW+RE yielded an underestimated absolute slope difference (percent bias: -18.38%) and changepoint occurrence (percent bias: -3.61%) for Y_2 , as well as similar estimated survival parameters compared to BC+RE.

The misspecification of the survival process using a current value structure (BC+CV) resulted in a greater overestimated absolute slope difference and transition widths compared to BC+RE, especially for Y_1 with a narrow transition zone (percent bias: 24.75% for β_{11} and 65.25% for γ_1). Overestimated changepoint locations were also observed in BC+CV (percent bias: 12.72% for $\beta_{\tau_{10}}$ and 5.31% for $\beta_{\tau_{20}}$). Like the previous scenarios, BC+CV and PW+CV were able to capture the associations between the longitudinal and survival processes.

Scenario 4: community cohort with current value association structure (Table 5 and Figure 8).

Most of the conclusions drawn in Scenario 2 also applied to Scenario 4, but larger bias was observed on average for all parameters due to the community cohort settings. It was worth noting that even the correctly specified model BC+CV still introduced bias in slopes, changepoints, and association parameters (e.g., percent bias: 13.87% for β_{12} , 58.59% for γ_1 , 4.31% for $\beta_{1\tau_0}$, and 11.23% for ζ_3).

The predictive accuracy results for Scenarios 3 and 4 in **Table 6** revealed that joint models combining piecewise models with correctly specified association structures (PW+RE in Scenario 3 and PW+CV in Scenario 4) achieved the lowest WAIC values. These models slightly outperformed the bent-cable models (BC+RE and BC+CV), which were used to generate the data (4108 for PW+RE versus 4113 for BC+RE in Scenario 3; 5509 for PW+CV versus 5522 for BC+CV in Scenario 4).

Summary

For the disease cohort with relatively frequent longitudinal visits and high event rates, the joint models that included the bent-cable model, which allowed for a smooth change, along with correctly specified association structures, showed the best performance regarding bias, coverage probability, and predictive accuracy. These findings demonstrated the flexibility of the bent-cable model in handling changepoints in longitudinal measures that exhibited either abrupt or smooth change with sufficient data. Parameter estimates from other joint models may deviate from the true values, but most bias and coverage probability values were within an acceptable range. When evaluating the impact of mis-specifying the longitudinal model with the survival model fixed (BC+RE vs PW+RE and BC+CV vs PW+CV), the piecewise model that assumed an abrupt change produced a more gradual change in cognitive decline after the changepoint and earlier detection of the changepoint compared to the bent-cable model. Differences in the estimates were more pronounced for the cognitive measure with a wider transition zone. The estimated survival parameters demonstrated robustness, maintaining stability across different selections of longitudinal models. When assessing the impact of mis-specifying the survival model with the longitudinal model fixed (BC+RE vs BC+CV and PW+RE vs PW+CV), the current value association structure resulted in a more drastic decline after the changepoint, a wider transition width, and a later occurrence of the changepoint, especially for the cognitive measure with a sharp change. The associations between the longitudinal and survival processes were captured despite the change in the association structure.

For the community cohort with relatively wide visit intervals and small event rates, more bias was observed compared to the disease cohort overall. The joint models that incorporated the bent-cable model and correctly specified associations produced the best performance for the cognitive measure with a wide and smooth change. In contrast, the piecewise model coupled

with correctly specified associations showed optimal results for the cognitive measure with a sharp change. Additionally, both piecewise and bent-cable models showed comparable predictive accuracy. These results indicated that the bent-cable model failed to improve upon the more parsimonious piecewise model in the presence of insufficient data, especially for the measure with an abrupt transition. Most of the other findings in the disease cohort were also applicable to the community cohort, but the differences among models became more obvious.

5. Real data application

In this section, we used longitudinal data from neuropsychological tests (NP tests) along with dementia and death data in the Framingham Heart Study (FHS) Offspring cohort to demonstrate our methods. Specifically, we fitted our proposed joint models to the FHS data and compared the estimated changepoints and slopes. The impact of the longitudinal process on the transitions between health states was also evaluated under different association structures.

5.1 Data description

The FHS is a multigenerational cohort study initiated in 1948 by enrolling 5209 residents from Framingham, Massachusetts in the Original cohort [43]. In 1971, 5214 participants who were offspring of the Original cohort and the spouses of these offspring were included in the Offspring cohort [44]. The participants in the Offspring cohort have undergone up to 10 examinations, which are scheduled every four to six years. Beginning in 1999, the surviving participants in the Offspring cohort were invited to join a secondary study, where they underwent a battery of NP tests every five or six years [15]. Participants identified as having potential cognitive impairment were invited to undergo additional, annual neurologic and NP tests. A dementia review panel determined whether the participants had dementia, as well as the dementia type and the date of

onset by reviewing every case of possible cognitive decline based on the participants' cognitive information such as neurologic and NP tests, medical records, and neuroimaging studies [15, 16].

NP tests are used to measure participants' cognitive changes over time. The NP test measures cover four domains in the FHS: memory, attention and executive function, visuo-perceptual, and language [45]. In this research project, we mainly focus on memory and language domains. The memory domain includes the following tests: Wechsler Memory Scale (WMS) Logical Memory – Immediate Recall, WMS Logical Memory – Delayed Recall, WMS Visual Reproductions – Immediate Recall, WMS Visual Reproductions – Delayed Recall, WMS Paired Associates – Immediate Recall, and WMS Paired Associates – Delayed Recall. The language domain includes Boston Naming Test 30 item version, Wide Range Achievement Test-3 (WRAT-3) Reading subtest, and Wechsler Adult Intelligence Scale (WAIS) Similarities subtest. For each participant at each visit, we calculated a memory score as the average of the Z scores for the NP tests in the memory domain and a language score as the average of the Z scores for the NP tests in the attention and executive function domain (**Figure 9**).

Of the 5124 Offspring cohort participants, 2992 had at least one NP test in the memory domain, while 2971 had at least one NP test in the language domain. We excluded participants with missing education information from our analysis: 471 out of 2992 participants for the memory domain, and 466 out of 2971 participants for the language domain. In addition, we restricted the data to participants who had at least three visits with NP tests in both memory and language domains, leaving 1128 participants. Of the 1128 participants, we excluded 11 participants who had dementia onset before the age of 60 or at baseline. Our final sample included 1117 participants. Of the 1117 participants in our final sample, 86 were diagnosed with

dementia during follow-up, 66 died without the onset of dementia, and 165 died following the diagnosis of dementia.

5.2 Model specification

In the longitudinal part of the joint model, we fitted two types of bivariate random changepoint models for the z-scores in the memory and language domains. The timescale was age centered at 60 years, and covariates included sex and education years. To reduce model complexity and facilitate model convergence, we only included random effects in the intercept term and the changepoint term. We encountered convergence problems in some of the models as we included more random effects and covariates, which prevented us from conducting proper model comparisons. These issues were also reported in previous Bayesian joint models [3, 21]. We assumed that two cognitive measures, i.e., memory z-scores and language z-scores, were correlated through their random changepoints. The piecewise model and the bent-cable model were shown as follows:

Piecewise model

$$\begin{aligned} \text{Memory}_i(t_{ij}) &= m_{1i}(t_{ij}) + \epsilon_{1i}(t_{ij}) \\ &= (\beta_{100} + \beta_{101}sex_i + \beta_{102}edu_i + b_{10i}) + \beta_{11}t_{ij} + \beta_{12}(t_{ij} \\ &\quad - \tau_{1i})I(t_{ij} - \tau_{1i}) + \epsilon_{1i}(t_{ij}) \end{aligned} \quad (44)$$

$$\begin{aligned} \text{Language}_i(t_{ij}) &= m_{2i}(t_{ij}) + \epsilon_{2i}(t_{ij}) \\ &= (\beta_{200} + \beta_{201}sex_i + \beta_{202}edu_i + b_{20i}) + \beta_{21}t_{ij} + \beta_{22}(t_{ij} \\ &\quad - \tau_{2i})I(t_{ij} - \tau_{2i}) + \epsilon_{2i}(t_{ij}) \end{aligned} \quad (45)$$

Bent-cable model

$$\text{Memory}_i(t_{ij}) = m_{1i}(t_{ij}) + \epsilon_{1i}(t_{ij}) \quad (46)$$

$$\begin{aligned}
 &= (\beta_{100} + \beta_{101}sex_i + \beta_{102}edu_i + b_{10i}) + \beta_{11}t_{ij} + \beta_{12}q(t_{ij}, \tau_{1i}, \gamma_1) \\
 &+ \epsilon_{1i}(t_{ij}) \\
 \text{Language}_i(t_{ij}) &= m_{2i}(t_{ij}) + \epsilon_{2i}(t_{ij}) \\
 &= (\beta_{200} + \beta_{201}sex_i + \beta_{202}edu_i + b_{20i}) + \beta_{21}t_{ij} \\
 &+ \beta_{22}q(t_{ij}, \tau_{2i}, \gamma_2) + \epsilon_{2i}(t_{ij})
 \end{aligned} \tag{47}$$

where β_{100} and β_{200} are intercepts for memory and language z-scores, and β_{101} , β_{102} , β_{201} , and β_{202} are coefficients of sex and education for two longitudinal measures. τ_{1i} and τ_{2i} are random changepoints for memory and language z-scores with the following distributions:

$$\begin{pmatrix} \tau_{1i} \\ \tau_{2i} \end{pmatrix} \sim N \left(\begin{pmatrix} \mu_{\tau 1} \\ \mu_{\tau 2} \end{pmatrix}, \begin{pmatrix} \sigma_{\tau 1}^2 & \rho_{\tau} \sigma_{\tau 1} \sigma_{\tau 2} \\ \rho_{\tau} \sigma_{\tau 1} \sigma_{\tau 2} & \sigma_{\tau 2}^2 \end{pmatrix} \right). b_{10i} \text{ and } b_{20i} \text{ are random intercepts for memory and}$$

language z-scores that follow normal distributions: $\begin{pmatrix} b_{10i} \\ b_{20i} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{b 1}^2 & 0 \\ 0 & \sigma_{b 2}^2 \end{pmatrix} \right)$. $q(\cdot)$ is the

function for smooth transition defined in Equation (7). The definitions for other parameters are described in the previous sections.

The random changepoint models were then combined with either of the two types of illness-death models to create our proposed joint models. We specified three hazard functions h_1 , h_2 , and h_3 for time to dementia, time to death without dementia, and time to death following dementia, respectively. Time in the illness-death models is the age at diagnosis of dementia or at death subtracted by 60 years in order to be consistent with the timescale in the longitudinal part.

We considered two association structures as follows:

Shared random effects model

$$h_1(t_1|X_i, b_i, \tau_i) = b_1 t_1^{b_1-1} \exp(\alpha_{10} + \alpha_{11}sex_i + \alpha_{12}edu_i + \zeta_{11}(\tau_{1i} - 20) + \zeta_{12}b_{10i}) \tag{48}$$

$$h_2(t_2|X_i, b_i, \tau_i) = b_2 t_2^{b_2-1} \exp(\alpha_{20} + \alpha_{21}sex_i + \alpha_{22}edu_i + \zeta_{21}(\tau_{1i} - 20) + \zeta_{22}b_{10i}) \tag{49}$$

$$h_3(t_2|t_1, X_i, b_i, \tau_i) \tag{50}$$
$$= b_3(t_2 - t_1)^{b_3-1} \exp(\alpha_{30} + \alpha_{31}sex_i + \alpha_{32}edu_i + \zeta_{31}(\tau_{1i} - 20) + \zeta_{32}b_{10i})$$

Current value model

$$h_1(t_1|X_i, b_i, \tau_i) = b_1 t^{b_1-1} \exp(\alpha_{10} + \alpha_{11}sex_i + \alpha_{12}edu_i + \zeta_{11}m_{1i}(t_1)) \tag{51}$$

$$h_2(t_2|X_i, b_i, \tau_i) = b_2 t^{b_2-1} \exp(\alpha_{20} + \alpha_{21}sex_i + \alpha_{22}edu_i + \zeta_{21}m_{1i}(t_2)) \tag{52}$$

$$h_3(t_2|t_1, X_i, b_i, \tau_i) = b_3(t_2 - t_1)^{b_3-1} \exp(\alpha_{30} + \alpha_{31}sex_i + \alpha_{32}edu_i + \zeta_{31}m_{1i}(t_2)) \tag{53}$$

Note that in both models, sex and education were included as covariates. The specifications of the association structures resembled those in our simulations, in which we only considered the changepoint or true current value for the first longitudinal cognitive measure to prevent multicollinearity. In the shared random effect model, we incorporated both random changepoint and random intercept as the link between the longitudinal and survival processes. For better model convergence, we further centered the time at changepoint since age 60 by subtracting 20, denoted as $(\tau_{1i} - 20)$.

We adopted Bayesian methods for estimation, incorporating weakly informative priors similar to those used in our simulation studies. Convergence was assessed using the potential scale reduction statistic \hat{R} [38]. Models were compared based on their WAIC values.

5.3 Results

The parameter estimates of the four proposed joint models, including posterior means and 95% credible intervals (CI), were displayed in **Table 7** and **Figures 10 and 11**. Based on the WAIC values of these four models, the models with smooth transition periods (BC+RE and BC+CV) resulted in a better model fit. Comparing BC+RE and BC+CV, the former yielded the smallest WAIC value, indicating a superior fit to the FHS Offspring cohort data. Therefore, we mainly referred to the results in BC+RE model for the following parameter interpretations.

Parameter interpretations in the longitudinal process

The slope parameters suggested a flat trajectory and then a drastic downward decline in cognitive functions for both memory and language domains. Specifically, the rate of decline in cognitive functions increased by 17 times for memory and 19 times for language after changepoints. There was a significant difference between men and women in baseline cognitive scores, with women showing higher cognitive function in both domains ($\beta_{101} = 0.28$, 95% CI: 0.22 to 0.34; $\beta_{201} = 0.1$, 95% CI: 0.04 to 0.16). The length of education was positively associated with baseline cognition ($\beta_{102} = 0.08$, 95% CI: 0.07 to 0.1; $\beta_{202} = 0.14$, 95% CI: 0.12 to 0.15). The smooth transition parameters demonstrated a 12-year transition period for the memory domain and a 14-year transition period for the language domain. On average, the changepoint was at the age of 86 ($\mu_{\tau 1} = 26$) for the memory domain while the changepoint for the language domain occurred later at the age of 90 ($\mu_{\tau 1} = 30$). The random changepoints for both cognitive domains showed a strong correlation, with a correlation coefficient ρ_{τ} of 0.9.

Parameter interpretations in the survival process

For the survival process from the healthy state to the dementia state, sex, education, and the random changepoint in the memory cognitive domain were significantly associated with the hazard for dementia. Specifically, women had a smaller risk of developing dementia adjusting for other covariates compared to men (HR=0.55, 95% CI: 0.31 to 0.93). A longer duration of education corresponded to a smaller risk for dementia (HR=0.87, 95% CI: 0.78 to 0.96). An increase of one year in the changepoint for the memory cognitive domain was associated with a 26% lower hazard (HR=0.74, 95% CI: 0.69 to 0.79) for the development of dementia. Of note, the effect of the random intercept of the memory cognitive domain was negative but not significant (HR=0.73, 95% CI: 0.36 to 1.52).

For the survival process from the healthy state to the death state, women had a lower risk of death compared to men (HR=0.55, 95% CI: 0.4 to 0.77) while education duration displayed non-significant association for participants who never developed dementia. Participants with later changepoint occurrence for the memory cognitive domain were less likely to die compared to those with earlier changepoint occurrence among the participants who had not experienced dementia (HR=0.92, 95% CI: 0.89 to 0.96).

For the transition from the dementia state to the death state, women displayed a lower hazard for death conditional on the occurrence of dementia, but this effect was not significant (HR=0.92, 95% CI: 0.53 to 1.58). Unlike the previous hazard function for death without dementia, we observed a borderline significant deleterious effect of education (HR=1.13, 95% CI: 1.01 to 1.26) and the occurrence of changepoint (HR=1.06, 95% CI: 1.02 to 1.11) on death following dementia. It appeared that having a longer period of education and a later occurrence of changepoint increased the risk for death for the participants with dementia.

Comparison of parameters among models

We compared some of the parameters of our main interest in different models, e.g., the slope differences before and after the changepoint, mean changepoints, and association parameters between two processes (**Figures 10 and 11**). Most of the findings in the simulation section applied to our real data. Overall, the estimation of longitudinal parameters, including the slope differences and mean changepoints, was influenced by the selection of changepoint formulations and association structures. The survival parameters showed stableness in different changepoint formulations. In the current value models (**Figure 11 Panel B**), the association between the true unobserved longitudinal measures and the hazard for death after dementia was not significant while the directions of associations were in alignment with the shared random effect models.

6. Discussion

In this paper, we proposed a joint model framework that incorporated a random changepoint model for longitudinal cognitive decline and an illness-death model for dementia and death. Our main objective was to model the longitudinal cognitive trajectories while investigating the effect of the longitudinal process on the transitions between health states. Within this framework, we further assessed two changepoint formulations coupled with two association structures between longitudinal and survival processes. To evaluate the robustness and accuracy of our proposed models, we performed simulation studies in the settings of both disease and community cohorts, with different longitudinal measures and joint model association structures. Our simulations confirmed that, the joint model, which combined the bent-cable model with a smooth transition and correctly specified association structures, displayed the best performance in the disease cohort. Other joint models also performed well with bias within a reasonable range. The estimation of longitudinal parameters tended to be influenced by the choice of both changepoint formulations and association structures. In contrast, survival parameter estimates proved robust against different changepoint formulations and allowed for effective characterization of the relationship between longitudinal and survival processes even with misspecifications. Compared to the disease cohort, the estimates for the community cohort yielded a larger bias. The bent-cable model failed to offer advantages over the simpler piecewise model using the data with less frequent longitudinal visits and low event rates in the community cohort. While most findings from the disease cohort also applied to the community cohort, the distinctions between models became more pronounced in the community cohort.

The application using NP test scores and time to dementia and death in the FHS Offspring cohort confirmed our simulation study findings regarding model comparison. The joint

model that included the bent-cable model and the shared random effect association structure displayed the best model fit based on WAIC values. This result suggested that the FHS data may resemble a structure where longitudinal cognitive decline displayed a smooth change over time, linked with times to dementia and death by the random changepoint. Additionally, we showed that sex, education, and the timing of changepoints were associated with transitions between different health states. Importantly, we found that the well-educated individuals tended to have a higher risk of death once they developed dementia compared to those who were less educated. It was reported that cognitive reserve for individuals with higher education contributes to more severe underlying brain lesions at the onset of dementia, resulting in a shorter subsequent survival period [46, 47]. Experiencing the cognitive changepoint at a later time led to a higher risk for death following dementia. Those with later changepoint occurrence might develop dementia at an older age, which could lead to a faster progression to death after the onset of dementia. The effect of changepoint occurrence might be mitigated if we further included the age at dementia onset as one of the explanatory variables or used the Markovian model.

The findings in our study are consistent with prior research. First, previous joint models, which combined changepoints in cognitive decline with times to dementia or death, demonstrated that changepoints estimated from cognitive scores were associated with the risk for dementia or death [19, 22, 28]. These associations correspond with the results in our study. Second, as for model comparison, Yang et al. (2013) compared three bivariate changepoint models for cognitive scores and BMI [14]. They suggested that the estimated changepoint of cognitive scores in the most flexible smooth polynomial model was years later than the other two simplified models including the broken-stick model. The smooth polynomial model, which included an additional smooth interval, allowed the identification of changepoints in a later time

window. Wang (2021) compared various formulations of changepoints in random changepoint models within the joint model framework [22]. This work suggested that both the bent-cable and smooth polynomial models could effectively characterize the changepoint data, having either abrupt or gradual transitions. This finding aligns with what we found in our disease cohort simulation.

Our study has some limitations. First, it is assumed that all participants will eventually experience a changepoint in their cognitive scores and develop dementia if they are followed long enough. To eliminate this assumption, the cure rate model that allows a fraction of participants to have a null risk of developing events of interest has been suggested in previous joint model literature [3, 48]. Second, a participant may develop dementia between his/her dementia diagnosis date and the preceding visit. Our method considers the time to dementia onset as the time at the dementia diagnosis date, potentially introducing bias due to interval censoring. Previous studies have suggested that using inappropriate models to handle time-to-event data with interval-censoring may cause an underestimation of dementia incidence [16, 46]. Third, the FHS Offspring cohort is a community-based cohort, where the participants underwent NP tests every five or six years, on average, with a low rate of dementia onset. Our simulations and previous research implied that the estimation of joint models can be improved if observations are recorded at more frequent intervals [3, 29]. Additionally, an FHS participant may be required to undergo additional NP testing annually on condition that his/her test score drops significantly or falls below an education-adjusted cutoff, making the visitation process informative. Although this study design contributes to more frequent observations, failing to account for the informative visitation process may result in biased longitudinal estimates [49].

Plourde (2020) proposed a joint model by treating the visitation process as repeated events using the FHS data [49].

Our study possesses some strengths. First, our proposed joint models combine changepoint estimation with illness-death models, distinguishing different health state transitions for semi-competing risks. This represents an improvement over previous joint models incorporating changepoints and semi-competing risks. Previous studies using a joint model with changepoints failed to consider each of the distinct transitions between health states. For example, Wang (2021) considered disease onset and disease-free death but ignored death following disease [22]. Dantan et al. (2011) failed to distinguish the transitions from different non-terminal states to the death state [28]. Second, we assessed the impact of different selections of changepoint formulations and association structures between two processes on parameter estimations in both sub-models. Previous studies neglected to consider different association structures in the survival process. Wang (2021) compared five changepoint models but assumed a shared random effect structure in all joint models [22]. Rouanet et al. (2016) took a latent class approach without explicitly specifying the association between longitudinal and survival processes [29]. Third, we evaluated a bivariate random changepoint model within the joint model framework, which can be easily extended to a multivariate case. The bivariate/multivariate approach accommodates different sources of variation and correlation among longitudinal outcomes. Additionally, it allows researchers to compare the temporal orders of changepoints in different longitudinal measures.

Our method can be extended in the following ways. First, our proposed joint model can be extended to account for interval censoring in the survival sub-model. Approaches to handle interval censoring have been used in the illness-death models and joint models. Joly et al. (2002)

proposed an illness-death model for interval-censored data using a penalized likelihood approach [50]. Rouanet et al. (2016) handled both interval censoring and semi-competing risks in a joint model framework, but the correlation between cognitive measures and times to dementia or death is captured by latent classes. Second, a participant may experience mild cognitive impairment (MCI) before progressing to dementia or death. Such disease progression can also be defined as a multistate process. To incorporate additional non-terminal health states such as MCI, the illness-death model can be extended to a multistate model within our joint model framework [51]. Some of the previous joint model studies included more health states compared with our study. Dantan et al. (2010) included a latent health state before dementia onset, referred to as the pre-diagnosis stage, which is closely related to MCI. Ferrer et al. (2016) proposed a joint model that combines a linear mixed effects model with a multistate model and focused on the impact of longitudinal dynamics on the transitions between clinical states [52]. Furthermore, personalized dynamic prediction using joint models has gained increasing research interest [53, 54]. In dynamic prediction, prediction results based on fitted joint models can be updated utilizing new information recorded over time. One of our next steps is to perform dynamic predictions for future longitudinal trajectories as well as risks for dementia and death using our fitted joint models.

In conclusion, we proposed a joint model that integrates a multivariate random changepoint model for longitudinal cognitive decline with an illness-death model for semi-competing risks. Our proposed model provided a flexible framework for estimating longitudinal trajectories with changepoints and for characterizing the influence of longitudinal measures on transitions between health states. We demonstrated that the selection of changepoint formulations and association structures influences model performance for different cohort structures. Future

extensions include addressing interval censoring and incorporating multistate models in our current joint model framework.

References

1. CDC. Available from: <https://www.cdc.gov/aging/dementia/index.html#:~:text=Dementia%20is%20not%20a%20specific,a%20part%20of%20normal%20aging>.
2. Masur, D.M., et al., *Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons*. *Neurology*, 1994. **44**(8): p. 1427-32.
3. Yu, B. and P. Ghosh, *Joint modeling for cognitive trajectory and risk of dementia in the presence of death*. *Biometrics*, 2010. **66**(1): p. 294-300.
4. Korten, A.E., et al., *A prospective study of cognitive function in the elderly*. *Psychol Med*, 1997. **27**(4): p. 919-30.
5. Backman, L., et al., *Rate of cognitive decline in preclinical Alzheimer's disease: the role of comorbidity*. *J Gerontol B Psychol Sci Soc Sci*, 2003. **58**(4): p. P228-36.
6. Dik, M.G., et al., *Insulin-like growth factor I (IGF-I) and cognitive decline in older persons*. *Neurobiol Aging*, 2003. **24**(4): p. 573-81.
7. Hall, C.B., et al., *Estimation of bivariate measurements having different change points, with application to cognitive ageing*. *Stat Med*, 2001. **20**(24): p. 3695-714.
8. Dominicus, A., et al., *A random change point model for assessing variability in repeated measures of cognitive function*. *Stat Med*, 2008. **27**(27): p. 5786-98.
9. Karr, J.E., et al., *When does cognitive decline begin? A systematic review of change point studies on accelerated decline in cognitive and neurological outcomes preceding mild cognitive impairment, dementia, and death*. *Psychol Aging*, 2018. **33**(2): p. 195-218.
10. Hall, C.B., et al., *A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease*. *Stat Med*, 2000. **19**(11-12): p. 1555-66.
11. Chiu, G., Lockhart, R., and Routledge, R., *Bent-cable regression theory and applications*. *Journal of the American statistical association*. **101**: p. 542-553.
12. van den Hout, A., G. Muniz-Terrera, and F.E. Matthews, *Smooth random change point models*. *Stat Med*, 2011. **30**(6): p. 599-610.
13. Huang, Y., et al., *Piecewise mixed-effects models with skew distributions for evaluating viral load changes: A Bayesian approach*. *Stat Methods Med Res*, 2015. **24**(6): p. 730-46.
14. Yang, L. and S. Gao, *Bivariate random change point models for longitudinal outcomes*. *Stat Med*, 2013. **32**(6): p. 1038-53.
15. Satizabal, C.L., et al., *Incidence of Dementia over Three Decades in the Framingham Heart Study*. *N Engl J Med*, 2016. **374**(6): p. 523-32.
16. Binder, N., J. Balmford, and M. Schumacher, *A multi-state model based reanalysis of the Framingham Heart Study: Is dementia incidence really declining?* *Eur J Epidemiol*, 2019. **34**(11): p. 1075-1083.
17. Furgal, A.K.C., A. Sen, and J.M.G. Taylor, *Review and Comparison of Computational Approaches for Joint Longitudinal and Time-to-Event Models*. *Int Stat Rev*, 2019. **87**(2): p. 393-418.
18. Hickey, G.L., et al., *Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues*. *BMC Med Res Methodol*, 2016. **16**(1): p. 117.

19. Jacqmin-Gadda, H., D. Commenges, and J.F. Dartigues, *Random change point model for joint modeling of cognitive decline and dementia*. Biometrics, 2006. **62**(1): p. 254-60.
20. Tapsoba Jde, D., S.M. Lee, and C.Y. Wang, *Joint modeling of survival time and longitudinal data with subject-specific changepoints in the covariates*. Stat Med, 2011. **30**(3): p. 232-49.
21. Terrera, G.M., et al., *Joint Modeling of Longitudinal Change and Survival: An Investigation of the Association Between Change in Memory Scores and Death*. GeroPsych (Bern), 2011. **24**(4): p. 177-185.
22. Wang, Y., *Bayesian Joint Modeling of Longitudinal and Time-to-Event Data with Application to Alzheimer's Disease*. 2021, University of Illinois at Chicago.
23. Fine, J.P., H. Jiang, and R. Chappell, *On Semi-Competing Risks Data*. Biometrika, 2001. **88**(4): p. 907-919.
24. Haneuse, S. and K.H. Lee, *Semi-Competing Risks Data Analysis: Accounting for Death as a Competing Risk When the Outcome of Interest Is Nonterminal*. Circ Cardiovasc Qual Outcomes, 2016. **9**(3): p. 322-31.
25. Fu, H., et al., *Joint modeling of progression-free survival and overall survival by a Bayesian normal induced copula estimation model*. Stat Med, 2013. **32**(2): p. 240-54.
26. Xu, J., J.D. Kalbfleisch, and B. Tai, *Statistical analysis of illness-death processes and semicompeting risks data*. Biometrics, 2010. **66**(3): p. 716-25.
27. Lee, K.H., et al., *Bayesian Semi-parametric Analysis of Semi-competing Risks Data: Investigating Hospital Readmission after a Pancreatic Cancer Diagnosis*. J R Stat Soc Ser C Appl Stat, 2015. **64**(2): p. 253-273.
28. Dantan, E., et al., *Joint model with latent state for longitudinal and multistate data*. Biostatistics, 2011. **12**(4): p. 723-36.
29. Rouanet, A., et al., *Joint latent class model for longitudinal data and interval-censored semi-competing events: Application to dementia*. Biometrics, 2016. **72**(4): p. 1123-1135.
30. Kiuchi, A.S., et al., *Change points in the series of T4 counts prior to AIDS*. Biometrics, 1995. **51**(1): p. 236-48.
31. Bacon, D.W. and D.G. Watts, *Estimating the Transition between Two Intersecting Straight Lines*. Biometrika, 1971. **58**(3): p. 525-534.
32. Chiu, G., R. Lockhart, and R. Routledge, *Bent-Cable Regression Theory and Applications*. Journal of the American Statistical Association, 2006. **101**(474): p. 542-553.
33. Reeder, H.T., J. Lu, and S. Haneuse, *Penalized estimation of frailty-based illness-death models for semi-competing risks*. Biometrics, 2023. **79**(3): p. 1657-1669.
34. Brilleman, S.L., et al., *Bayesian Piecewise Linear Mixed Models With a Random Change Point: An Application to BMI Rebound in Childhood*. Epidemiology, 2017. **28**(6): p. 827-833.
35. Lewandowski, D., D. Kurowicka, and H. Joe, *Generating random correlation matrices based on vines and extended onion method*. Journal of Multivariate Analysis, 2009. **100**(9): p. 1989-2001.
36. Stan-Development-Team, *Stan: A C++ Library for Probability and Sampling*. 2014, Stan Development Team.

37. Hoffman, M.D. and A. Gelman, *The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo*. Journal of Machine Learning Research, 2014. **15**: p. 1593-1623.
38. Gelman, A., et al., *Bayesian Data Analysis*. 3 ed. 2013: Chapman and Hall/CRC.
39. Watanabe, S., *Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory*. Journal of Machine Learning Research, 2010. **11**.
40. Aki Vehtari, A.G., Jonah Gabry *Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC*. Statistics and Computing, 2017. **27**: p. 1413-1432.
41. Beyersmann, J., et al., *Simulating competing risks data in survival analysis*. Stat Med, 2009. **28**(6): p. 956-71.
42. Bender, R., T. Augustin, and M. Blettner, *Generating survival times to simulate Cox proportional hazards models*. Stat Med, 2005. **24**(11): p. 1713-23.
43. Dawber, T.R., G.F. Meadors, and F.E. Moore, Jr., *Epidemiological approaches to heart disease: the Framingham Study*. Am J Public Health Nations Health, 1951. **41**(3): p. 279-81.
44. Feinleib, M., et al., *The Framingham Offspring Study. Design and preliminary data*. Prev Med, 1975. **4**(4): p. 518-25.
45. Yang, J., et al., *Establishing cognitive baseline in three generations: Framingham Heart Study*. Alzheimers Dement (Amst), 2023. **15**(1): p. e12416.
46. Leffondre, K., et al., *Interval-censored time-to-event and competing risk with death: is the illness-death model more accurate than the Cox model?* Int J Epidemiol, 2013. **42**(4): p. 1177-86.
47. Stern, Y., et al., *Increased risk of mortality in Alzheimer's disease patients with more advanced educational and occupational attainment*. Ann Neurol, 1995. **37**(5): p. 590-5.
48. Law, N.J., J.M. Taylor, and H. Sandler, *The joint modeling of a longitudinal disease progression marker and the failure time process in the presence of cure*. Biostatistics, 2002. **3**(4): p. 547-63.
49. PLOURDE, K., *Joint models for estimating determinants of cognitive decline in the presence of selection bias after enrollment*. 2020, Boston University.
50. Joly, P., et al., *A penalized likelihood approach for an illness-death model with interval-censored data: application to age-specific incidence of dementia*. Biostatistics, 2002. **3**(3): p. 433-43.
51. Putter, H., M. Fiocco, and R.B. Geskus, *Tutorial in biostatistics: competing risks and multi-state models*. Stat Med, 2007. **26**(11): p. 2389-430.
52. Ferrer, L., et al., *Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer*. Stat Med, 2016. **35**(22): p. 3933-48.
53. Wang, J., S. Luo, and L. Li, *Dynamic Prediction for Multiple Repeated Measures and Event Time Data: An Application to Parkinson's Disease*. Ann Appl Stat, 2017. **11**(3): p. 1787-1809.
54. Rizopoulos, D., *Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data*. Biometrics, 2011. **67**(3): p. 819-29.

Tables and Figures

Table 1. Main parameter settings for all four scenarios in the simulation study.

Parameters	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Longitudinal				
Interval between visits	2 years	2 years	5 years	5 years
β_{10}	-0.18	-0.18	-0.18	-0.18
β_{11}	-0.02	-0.02	-0.02	-0.02
β_{12}	-0.15	-0.15	-0.15	-0.15
γ_1	2	2	2	2
σ_{10}	0.4	0.4	0.4	0.4
σ_{12}	0.05	0.05	0.05	0.05
ρ_{b1}	0.3	0.3	0.3	0.3
β_{20}	-0.03	-0.03	-0.03	-0.03
β_{21}	-0.02	-0.02	-0.02	-0.02
β_{22}	-0.2	-0.2	-0.2	-0.2
γ_2	4	4	4	4
σ_{20}	0.5	0.5	0.5	0.5
σ_{22}	0.05	0.05	0.05	0.05
ρ_{b2}	0.3	0.3	0.3	0.3
Changepoint				
$\beta_{1\tau 0}$	12	12	25	25
$\beta_{1\tau 1}$	-3	-3	-3	-3
$\sigma_{\tau 1}$	3	3	8	8
$\beta_{2\tau 0}$	15	15	27	27
$\beta_{2\tau 1}$	-5	-5	-5	-5
$\sigma_{\tau 2}$	3	3	8	8
ρ_{τ}	0.8	0.8	0.8	0.8
Survival				
α_1	-0.6	-0.2	-0.6	-0.2
ζ_1	-0.3	-2.5	-0.3	-2.5
α_2	-0.6	-0.4	-0.6	-0.4
ζ_2	-0.1	-0.6	-0.1	-0.6
α_3	-0.1	-0.2	-0.1	-0.2
ζ_3	-0.1	-0.5	-0.1	-0.5

Scenario 1, disease cohort with shared random effect association structure; Scenario 2, disease cohort with current value association structure; Scenario 3, community cohort with shared random effect association structure; Scenario 4, community cohort with current value association structure.

Table 2. Simulation results for Scenario 1 (disease cohort with shared random effect association structure).

Parameter	True	BC+RE			PW+RE			BC+CV			PW+CV		
		Mean	PB	CP	Mean	PB	CP	Mean	PB	CP	Mean	PB	CP
Longitudinal													
β_{10}	-0.18	-0.18	-0.24	0.95	-0.18	-0.99	0.95	-0.18	-0.39	0.96	-0.18	-1.38	0.96
β_{11}	-0.02	-0.02	0.02	0.94	-0.02	3.02	0.89	-0.02	0.29	0.95	-0.02	4.08	0.81
β_{12}	-0.15	-0.15	1.81	0.94	-0.14	-4.56	0.72	-0.16	6.62	0.79	-0.15	-1.49	0.89
β_{20}	-0.03	-0.03	4.67	0.91	-0.03	-2.73	0.93	-0.03	3.23	0.90	-0.03	-3.12	0.92
β_{21}	-0.02	-0.02	-0.42	0.95	-0.02	22.45	0.00	-0.02	-0.09	0.95	-0.02	23.36	0.00
β_{22}	-0.20	-0.20	0.43	0.96	-0.18	-11.65	0.00	-0.20	1.15	0.92	-0.18	-11.32	0.00
γ_1	2.00	2.03	1.58	0.94	-	-	-	2.43	21.40	0.87	-	-	-
γ_2	4.00	4.04	0.99	0.96	-	-	-	4.12	3.03	0.93	-	-	-
Changepoint													
$\beta_{1\tau 0}$	15.00	15.09	0.60	0.92	14.80	-1.33	0.82	15.47	3.13	0.63	15.11	0.75	0.91
$\beta_{2\tau 0}$	12.00	12.02	0.21	0.97	11.55	-3.78	0.08	12.13	1.06	0.85	11.64	-3.03	0.26
$\beta_{1\tau 1}$	-3.00	-2.99	-0.44	0.95	-2.98	-0.80	0.95	-3.07	2.34	0.91	-3.08	2.63	0.92
$\beta_{2\tau 1}$	-5.00	-5.01	0.18	0.94	-4.87	-2.64	0.90	-5.05	0.95	0.93	-4.91	-1.80	0.93
$\sigma_{\tau 1}$	3.00	3.03	1.07	0.96	3.02	0.59	0.97	3.11	3.63	0.90	3.10	3.43	0.91
$\sigma_{\tau 2}$	3.00	3.00	0.12	0.93	2.97	-1.01	0.93	3.03	1.01	0.93	3.00	-0.10	0.92
ρ_{τ}	0.80	0.80	-0.62	0.97	0.80	-0.35	0.97	0.79	-0.64	0.97	0.79	-1.27	0.98
Survival													
α_1	-0.60	-0.59	-1.50	0.94	-0.59	-1.62	0.95	-0.52	-	-	-0.51	-	-
ζ_1	-0.30	-0.30	0.12	0.95	-0.31	3.79	0.96	-0.68	-	-	-0.68	-	-
α_2	-0.60	-0.58	-2.92	0.93	-0.58	-2.70	0.93	-0.56	-	-	-0.56	-	-
ζ_2	-0.10	-0.10	0.12	0.95	-0.11	9.96	0.93	-0.20	-	-	-0.21	-	-
α_3	-0.10	-0.10	1.48	0.93	-0.10	-1.94	0.93	-0.08	-	-	-0.08	-	-
ζ_3	-0.10	-0.11	7.26	0.95	-0.11	7.48	0.93	-0.14	-	-	-0.14	-	-

True, parameter true value; Mean, posterior mean; PB, percent bias; CP: 95% credible interval; BC, bent-cable model; RE, shared random effect association structure; PW, piecewise model; CV, current value association structure.

Table 3. Simulation results for Scenario 2 (disease cohort with current value association structure).

Parameter	True	BC+RE			PW+RE			BC+CV			PW+CV		
		Mean	PB	CP	Mean	PB	CP	Mean	PB	CP	Mean	PB	CP
Longitudinal													
β_{10}	-0.18	-0.18	-0.61	0.96	-0.18	-1.21	0.95	-0.18	-0.41	0.97	-0.18	-1.17	0.97
β_{11}	-0.02	-0.02	-1.28	0.94	-0.02	1.21	0.94	-0.02	0.15	0.95	-0.02	3.02	0.83
β_{12}	-0.15	-0.14	-5.62	0.71	-0.13	-10.09	0.13	-0.15	1.73	0.95	-0.14	-4.00	0.79
β_{20}	-0.03	-0.03	4.05	0.92	-0.03	-2.61	0.93	-0.03	3.70	0.91	-0.03	-2.10	0.92
β_{21}	-0.02	-0.02	-0.40	0.96	-0.02	22.40	0.00	-0.02	-0.36	0.96	-0.02	22.72	0.00
β_{22}	-0.20	-0.20	0.04	0.96	-0.18	-11.52	0.00	-0.20	0.33	0.95	-0.18	-11.41	0.00
γ_1	2.00	1.82	-9.03	0.97	-	-	-	2.01	0.67	0.96	-	-	-
γ_2	4.00	4.00	-0.11	0.97	-	-	-	4.04	0.94	0.95	-	-	-
Changepoint													
$\beta_{1\tau_0}$	15.00	14.90	-0.65	0.93	14.68	-2.13	0.67	15.08	0.56	0.95	14.85	-0.98	0.85
$\beta_{2\tau_0}$	12.00	11.94	-0.49	0.92	11.48	-4.31	0.03	12.02	0.20	0.95	11.57	-3.62	0.10
$\beta_{1\tau_1}$	-3.00	-2.83	-5.64	0.94	-2.84	-5.49	0.93	-2.98	-0.60	0.95	-2.99	-0.46	0.95
$\beta_{2\tau_1}$	-5.00	-4.96	-0.70	0.91	-4.82	-3.56	0.90	-5.01	0.24	0.91	-4.87	-2.66	0.92
σ_{τ_1}	3.00	3.28	9.42	0.52	3.26	8.55	0.55	3.05	1.60	0.93	3.06	2.08	0.94
σ_{τ_2}	3.00	3.00	-0.15	0.97	2.96	-1.41	0.93	3.00	0.14	0.97	2.97	-1.13	0.93
ρ_{τ}	0.80	0.73	-9.37	0.40	0.74	-8.08	0.51	0.79	-0.88	0.95	0.79	-1.50	0.95
Survival													
α_1	-0.20	-0.14	-	-	-0.14	-	-	-0.20	1.01	0.93	-0.20	1.12	0.94
ζ_1	-2.50	-0.28	-	-	-0.28	-	-	-2.54	1.43	0.92	-2.55	2.06	0.92
α_2	-0.40	-0.38	-	-	-0.38	-	-	-0.39	-1.77	0.94	-0.39	-1.67	0.94
ζ_2	-0.60	-0.05	-	-	-0.06	-	-	-0.58	-3.85	0.94	-0.58	-3.19	0.93
α_3	-0.20	-0.16	-	-	-0.16	-	-	-0.20	-1.03	0.97	-0.20	-1.42	0.97
ζ_3	-0.50	-0.05	-	-	-0.05	-	-	-0.52	4.51	0.97	-0.54	7.04	0.95

True, parameter true value; Mean, posterior mean; PB, percent bias; CP: 95% credible interval; BC, bent-cable model; RE, shared random effect association structure; PW, piecewise model; CV, current value association structure.

Table 4. Simulation results for Scenario 3 (community cohort with shared random effect association structure).

Parameter	True	BC+RE			PW+RE			BC+CV			PW+CV		
		Mean	PB	CP	Mean	PB	CP	Mean	PB	CP	Mean	PB	CP
Longitudinal													
β_{10}	-0.18	-0.18	0.80	0.94	-0.18	0.47	0.94	-0.18	0.61	0.95	-0.18	0.27	0.95
β_{11}	-0.02	-0.02	-0.21	0.93	-0.02	0.89	0.90	-0.02	1.63	0.91	-0.02	3.06	0.83
β_{12}	-0.15	-0.17	14.28	0.94	-0.15	-2.45	0.96	-0.19	24.75	0.74	-0.16	7.28	0.89
β_{20}	-0.03	-0.03	2.17	0.94	-0.03	-1.06	0.95	-0.03	2.32	0.94	-0.03	-1.30	0.95
β_{21}	-0.02	-0.02	0.10	0.97	-0.02	2.04	0.85	-0.02	0.70	0.96	-0.02	2.58	0.83
β_{22}	-0.20	-0.21	2.86	0.95	-0.16	-18.38	0.20	-0.20	1.12	0.96	-0.16	-18.54	0.17
γ_1	2.00	2.96	48.10	0.98	-	-	-	3.31	65.25	0.98	-	-	-
γ_2	4.00	4.16	3.96	0.96	-	-	-	4.10	2.42	0.96	-	-	-
Changepoint													
$\beta_{1\tau_0}$	25.00	25.81	3.23	0.92	25.10	0.41	0.97	28.18	12.72	0.15	27.72	10.87	0.33
$\beta_{2\tau_0}$	27.00	27.40	1.49	0.96	26.03	-3.61	0.82	28.43	5.31	0.67	27.46	1.70	0.92
$\beta_{1\tau_1}$	-3.00	-2.94	-1.88	0.97	-3.01	0.42	0.97	-3.40	13.17	0.96	-3.55	18.43	0.95
$\beta_{2\tau_1}$	-5.00	-4.98	-0.45	0.93	-5.04	0.76	0.94	-5.23	4.56	0.93	-5.46	9.10	0.93
σ_{τ_1}	8.00	8.13	1.64	0.96	8.18	2.20	0.96	8.57	7.14	0.88	8.71	8.84	0.83
σ_{τ_2}	8.00	8.15	1.93	0.93	8.06	0.70	0.93	8.40	4.98	0.86	8.45	5.61	0.89
ρ_{τ}	0.80	0.79	-1.45	0.97	0.80	0.44	0.96	0.80	-0.26	0.97	0.81	1.41	0.92
Survival													
α_1	-0.60	-0.66	9.62	0.94	-0.68	13.01	0.92	-0.50	-	-	-0.49	-	-
ζ_1	-0.30	-0.32	6.37	0.93	-0.32	7.04	0.92	-1.53	-	-	-1.52	-	-
α_2	-0.60	-0.61	2.03	0.97	-0.62	3.22	0.97	-0.57	-	-	-0.57	-	-
ζ_2	-0.10	-0.10	-4.67	0.92	-0.09	-7.12	0.94	-0.32	-	-	-0.32	-	-
α_3	-0.10	-0.15	50.11	0.97	-0.16	57.49	0.97	-0.08	-	-	-0.07	-	-
ζ_3	-0.10	-0.12	16.22	0.96	-0.12	15.78	0.94	-0.29	-	-	-0.28	-	-

True, parameter true value; Mean, posterior mean; PB, percent bias; CP: 95% credible interval; BC, bent-cable model; RE, shared random effect association structure; PW, piecewise model; CV, current value association structure.

Table 5. Simulation results for Scenario 4 (community cohort with current value association structure).

Parameter	True	BC+RE			PW+RE			BC+CV			PW+CV		
		Mean	PB	CP	Mean	PB	CP	Mean	PB	CP	Mean	PB	CP
Longitudinal													
β_{10}	-0.18	-0.18	-1.03	0.95	-0.18	-1.41	0.94	-0.18	0.89	0.94	-0.18	0.50	0.93
β_{11}	-0.02	-0.02	-6.43	0.71	-0.02	-4.28	0.83	-0.02	-0.17	0.91	-0.02	1.03	0.89
β_{12}	-0.15	-0.13	-11.91	0.77	-0.12	-23.02	0.38	-0.17	13.87	0.89	-0.15	-2.90	0.90
β_{20}	-0.03	-0.03	0.20	0.95	-0.03	-1.45	0.95	-0.03	0.62	0.94	-0.03	-2.70	0.93
β_{21}	-0.02	-0.02	0.00	0.95	-0.02	2.10	0.88	-0.02	0.33	0.93	-0.02	2.27	0.87
β_{22}	-0.20	-0.21	2.83	0.91	-0.17	-16.63	0.27	-0.20	2.08	0.94	-0.17	-17.02	0.23
γ_1	2.00	3.89	94.63	0.95	-	-	-	3.17	58.59	0.98	-	-	-
γ_2	4.00	4.04	0.94	0.97	-	-	-	4.04	0.99	0.96	-	-	-
Changepoint													
$\beta_{1\tau_0}$	25.00	24.10	-3.59	0.89	23.47	-6.13	0.70	26.08	4.31	0.93	25.34	1.35	0.94
$\beta_{2\tau_0}$	27.00	26.21	-2.94	0.92	24.92	-7.70	0.45	27.26	0.97	0.99	26.01	-3.66	0.84
$\beta_{1\tau_1}$	-3.00	-2.41	-19.52	0.91	-2.46	-18.14	0.93	-3.15	5.15	0.95	-3.09	3.06	0.97
$\beta_{2\tau_1}$	-5.00	-4.61	-7.83	0.93	-4.65	-7.08	0.95	-5.02	0.47	0.94	-4.99	-0.21	0.95
σ_{τ_1}	8.00	9.17	14.59	0.59	9.18	14.76	0.59	8.26	3.31	0.92	8.28	3.46	0.91
σ_{τ_2}	8.00	7.91	-1.15	0.95	7.81	-2.33	0.93	8.09	1.17	0.94	7.98	-0.30	0.96
ρ_{τ}	0.80	0.74	-7.90	0.72	0.76	-5.41	0.87	0.79	-1.84	0.95	0.80	-0.26	0.96
Survival													
α_1	-0.20	-0.15	-	-	-0.15	-	-	-0.19	-5.34	0.93	-0.19	-5.88	0.92
ζ_1	-2.50	-0.21	-	-	-0.20	-	-	-2.55	1.87	0.92	-2.55	2.07	0.90
α_2	-0.40	-0.38	-	-	-0.39	-	-	-0.39	-2.22	0.96	-0.40	-0.53	0.96
ζ_2	-0.60	-0.01	-	-	-0.01	-	-	-0.59	-2.29	0.92	-0.58	-2.98	0.92
α_3	-0.20	-0.19	-	-	-0.19	-	-	-0.22	12.19	0.91	-0.22	9.02	0.90
ζ_3	-0.50	-0.03	-	-	-0.03	-	-	-0.56	11.23	0.95	-0.57	14.47	0.95

True, parameter true value; Mean, posterior mean; PB, percent bias; CP: 95% credible interval; BC, bent-cable model; RE, shared random effect association structure; PW, piecewise model; CV, current value association structure.

Table 6. Mean WAIC values for each model in each scenario.

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
BC+RE	7616	7904	4113	5559
PW+RE	7716	8004	4108	5557
BC+CV	7777	7536	4363	5522
PW+CV	7880	7639	4367	5509

BC, bent-cable model; RE, shared random effect association structure; PW, piecewise model; CV, current value association structure.

Table 7. Joint model results for the Framingham Heart Study Offspring cohort data.

Param	BC+RE (WAIC=6136)			PW+RE (WAIC=6193)			BC+CV (WAIC=6214)			PW+CV (WAIC=6283)		
	Mean	Lower	Upper	Mean	Lower	Upper	Mean	Lower	Upper	Mean	Lower	Upper
Longitudinal												
β_{100}	-0.18	-0.27	-0.08	-0.18	-0.27	-0.08	-0.18	-0.28	-0.08	-0.19	-0.28	-0.09
β_{101}	0.28	0.22	0.34	0.28	0.22	0.33	0.28	0.22	0.34	0.28	0.22	0.34
β_{102}	0.08	0.07	0.10	0.08	0.07	0.10	0.08	0.07	0.10	0.08	0.07	0.10
β_{11}	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
β_{12}	-0.17	-0.20	-0.15	-0.15	-0.17	-0.14	-0.17	-0.20	-0.15	-0.16	-0.18	-0.14
β_{200}	-0.04	-0.14	0.06	-0.03	-0.13	0.06	-0.04	-0.14	0.07	-0.03	-0.14	0.06
β_{201}	0.10	0.04	0.16	0.10	0.04	0.16	0.10	0.04	0.16	0.10	0.04	0.16
β_{202}	0.14	0.12	0.15	0.14	0.12	0.15	0.14	0.12	0.15	0.14	0.12	0.15
β_{21}	-0.01	-0.01	0.00	-0.01	-0.01	0.00	-0.01	-0.01	0.00	-0.01	-0.01	0.00
β_{22}	-0.19	-0.24	-0.15	-0.12	-0.15	-0.10	-0.19	-0.25	-0.14	-0.13	-0.15	-0.10
γ_1	6.18	3.21	7.91	-	-	-	4.94	2.28	7.57	-	-	-
γ_2	7.44	6.17	7.98	-	-	-	7.39	5.98	7.98	-	-	-
σ_{b1}	0.44	0.41	0.46	0.44	0.41	0.46	0.44	0.42	0.46	0.44	0.42	0.46
σ_{b2}	0.49	0.46	0.51	0.48	0.46	0.51	0.49	0.47	0.51	0.48	0.46	0.51
Changepoint												
μ_{τ_1}	25.82	23.99	27.96	25.43	23.61	27.59	27.64	25.40	30.17	27.51	25.21	30.05
μ_{τ_2}	30.00	27.80	32.34	27.90	25.70	30.36	31.39	28.91	34.11	29.38	26.94	32.21
σ_{τ_1}	9.54	8.45	10.86	9.72	8.52	11.18	10.46	9.16	11.98	10.65	9.31	12.23
σ_{τ_2}	8.97	7.74	10.45	9.63	8.19	11.35	9.54	8.17	11.25	10.20	8.61	12.24
ρ_{τ}	0.90	0.85	0.94	0.90	0.84	0.94	0.90	0.85	0.95	0.90	0.84	0.95
Survival												
b_1	6.59	5.36	7.98	6.44	5.25	7.82	1.97	1.41	2.57	1.98	1.42	2.60
α_{10}	-22.32	-26.80	-18.32	-22.05	-26.51	-18.09	-10.15	-12.17	-8.28	-10.17	-12.22	-8.29
α_{11}	-0.60	-1.16	-0.07	-0.60	-1.15	-0.06	-0.01	-0.53	0.49	-0.01	-0.55	0.54
α_{12}	-0.14	-0.25	-0.04	-0.14	-0.25	-0.04	0.04	-0.06	0.14	0.04	-0.06	0.14
ζ_{11}	-0.30	-0.37	-0.23	-0.29	-0.36	-0.23	-2.53	-3.10	-2.02	-2.55	-3.17	-2.01
ζ_{12}	-0.32	-1.01	0.42	-0.42	-1.18	0.33	-	-	-	-	-	-
b_2	4.43	3.85	5.05	4.36	3.79	5.00	3.53	3.07	3.98	3.55	3.11	3.99
α_{20}	-14.42	-16.38	-12.56	-14.26	-16.27	-12.48	-12.40	-13.90	-10.91	-12.45	-13.93	-11.00
α_{21}	-0.59	-0.92	-0.26	-0.58	-0.91	-0.25	-0.42	-0.73	-0.12	-0.43	-0.75	-0.12
α_{22}	-0.01	-0.07	0.05	-0.01	-0.08	0.05	0.04	-0.02	0.10	0.04	-0.02	0.10
ζ_{21}	-0.08	-0.12	-0.04	-0.07	-0.12	-0.03	-0.62	-0.91	-0.32	-0.60	-0.89	-0.30
ζ_{22}	-0.16	-0.56	0.25	-0.20	-0.59	0.18	-	-	-	-	-	-
b_3	1.58	1.27	1.91	1.58	1.27	1.93	1.59	1.22	2.00	1.60	1.23	2.01
α_{30}	-2.15	-3.12	-1.23	-2.11	-3.13	-1.19	-1.85	-2.82	-0.91	-1.83	-2.84	-0.87
α_{31}	-0.08	-0.63	0.46	-0.08	-0.62	0.46	-0.21	-0.73	0.33	-0.20	-0.72	0.33
α_{32}	0.12	0.01	0.23	0.12	0.01	0.24	0.07	-0.04	0.17	0.06	-0.04	0.17
ζ_{31}	0.06	0.02	0.10	0.06	0.02	0.10	0.29	-0.11	0.69	0.31	-0.12	0.74
ζ_{32}	0.53	-0.19	1.24	0.47	-0.32	1.27	-	-	-	-	-	-

Mean, posterior mean; Lower, lower bound of 95% credible interval; Upper: upper bound of 95% credible interval; BC, bent-cable model; RE, shared random effect association structure; PW, piecewise model; CV, current value association structure.

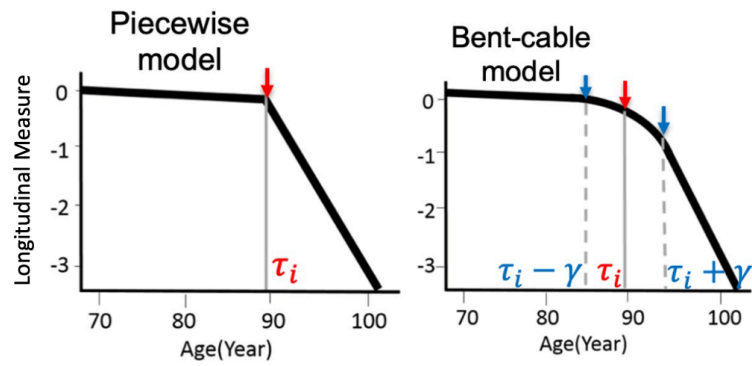


Figure 1. The visualization of the piecewise model (left) and the Bent-Cable model (right). The y-axis represents the value of a longitudinal measure, and the x-axis represents age. τ_i is the changepoint for participant i . $[\tau_i - \gamma, \tau_i + \gamma]$ denotes the smooth transition zone for the Bent-Cable model.

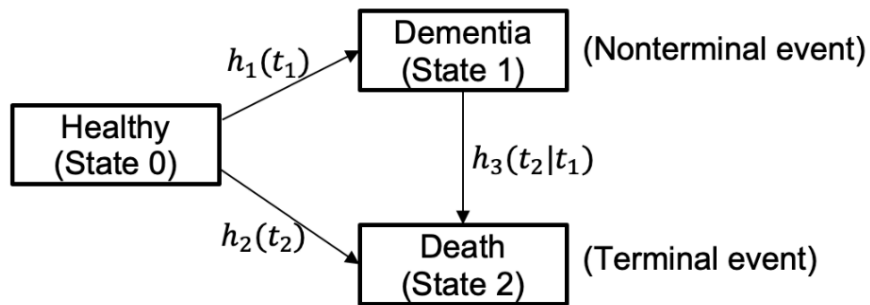


Figure 2. Three health states with semi-competing risks. $h_1(t_1)$ is the hazard function for dementia at time t_1 , $h_2(t_2)$ is the hazard function for death at time t_2 , and $h_3(t_2|t_1)$ denotes the hazard function for death at time t_2 given that the participant develop dementia at time t_1 .

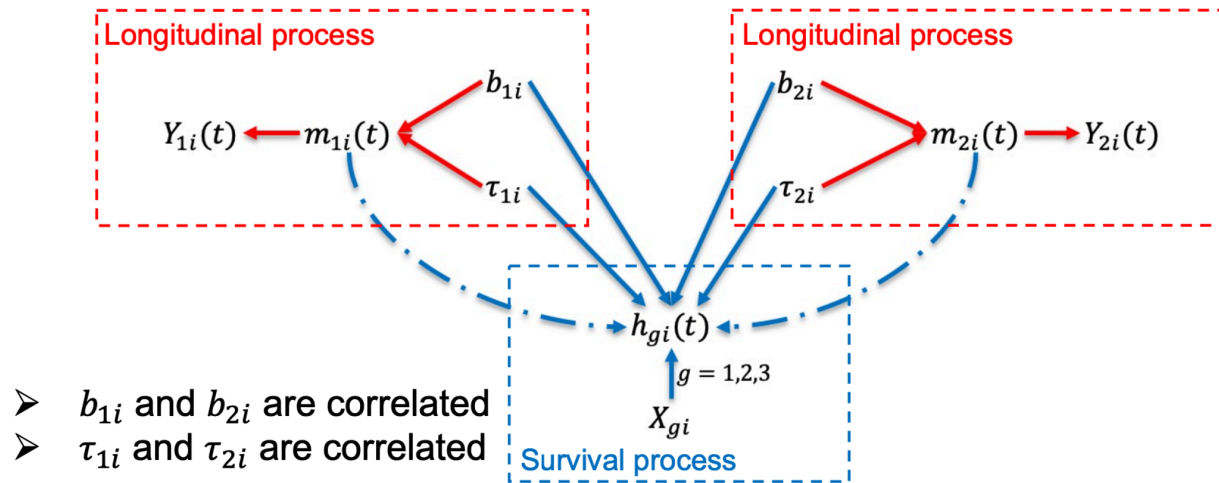


Figure 3. The joint model framework. Two red boxes represent the longitudinal process for two longitudinal measures, and the blue box is the survival process. Y_{1i} and Y_{2i} are two longitudinal measures, $m_{1i}(t)$ and $m_{2i}(t)$ are the true unobserved values for two longitudinal measures, τ_{1i} and τ_{2i} are two changepoints, b_{1i} and b_{2i} are random intercepts and slopes for the first and second longitudinal measures. $h_{gi}(t)$ denotes the hazard function for each health state transition, and X_g is a baseline covariate associated with each hazard function. Both longitudinal and survival processes are linked by either random effects (solid blue lines for the shared random effect model) or the true unobserved values of longitudinal measures (dashed blue lines for the current value model).

Case 1: dementia and censored before death



Case 2: death following dementia



Case 3: death without dementia



Case 4: censored before dementia or death

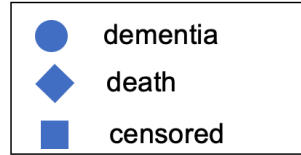


Figure 4. Four scenarios for dementia and death. The circle symbol represents the occurrence of dementia, the diamond symbol denotes the occurrence of death, and the square means the participant is censored.

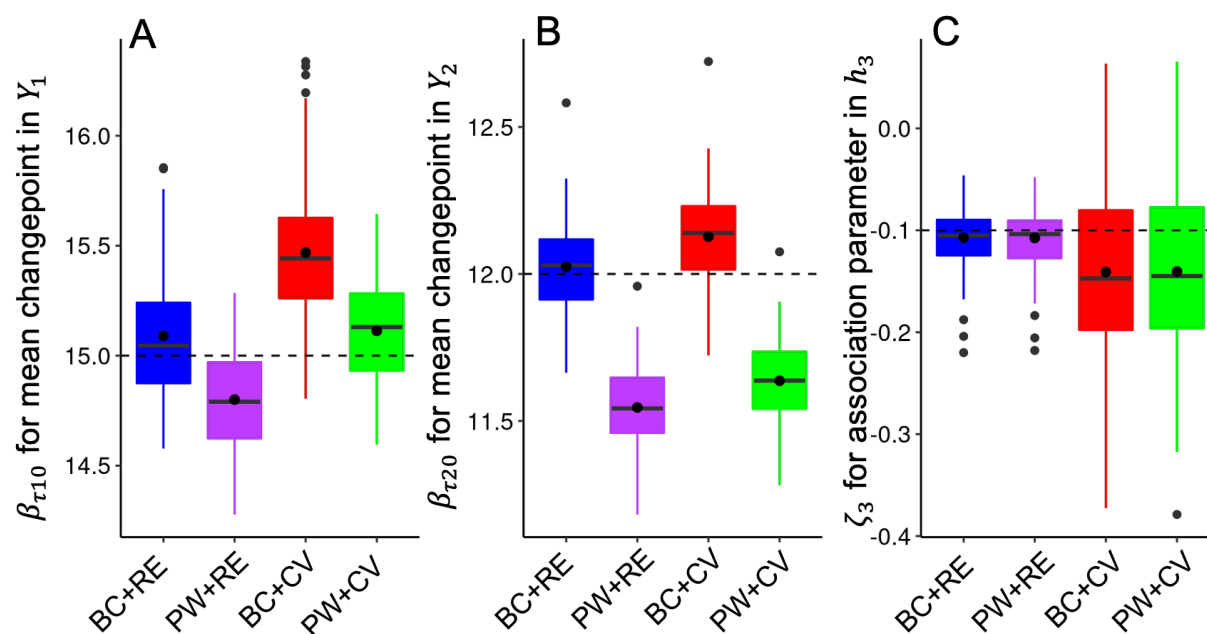


Figure 5. Distributions of posterior means for main changepoint and association parameters in Scenario 1 (disease cohort with shared random effect association structure). Panels A, B, and C represent the distributions of estimated posterior means for $\beta_{1\tau_0}$ (the mean changepoint parameter in Y_1), $\beta_{2\tau_0}$ (the mean changepoint parameter in Y_2), and ζ_3 (the association parameter between longitudinal and survival processes in h_3), respectively. In each panel, the boxes denote the posterior means across all simulated datasets based on each method, while the dashed line indicates the true value for the parameter. PW: piecewise model; BC: benticable model; RE: shared random effect structure; CV: current value structure.

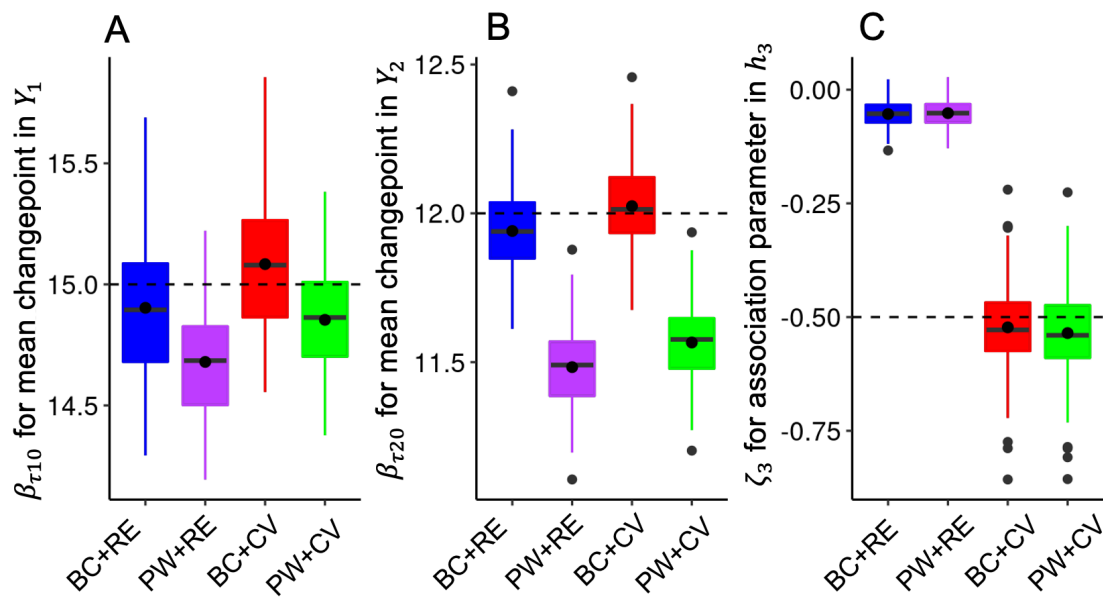


Figure 6. Distributions of posterior means for main changepoint and association parameters in Scenario 2 (disease cohort with current value association structure). Panels A, B, and C represent the distributions of estimated posterior means for $\beta_{1\tau_0}$ (the mean changepoint parameter in Y_1), $\beta_{2\tau_0}$ (the mean changepoint parameter in Y_2), and ζ_3 (the association parameter between longitudinal and survival processes in h_3), respectively. In each panel, the boxes denote the posterior means across all simulated datasets based on each method, while the dashed line indicates the true value for the parameter. PW: piecewise model; BC: bent-cable model; RE: shared random effect structure; CV: current value structure.

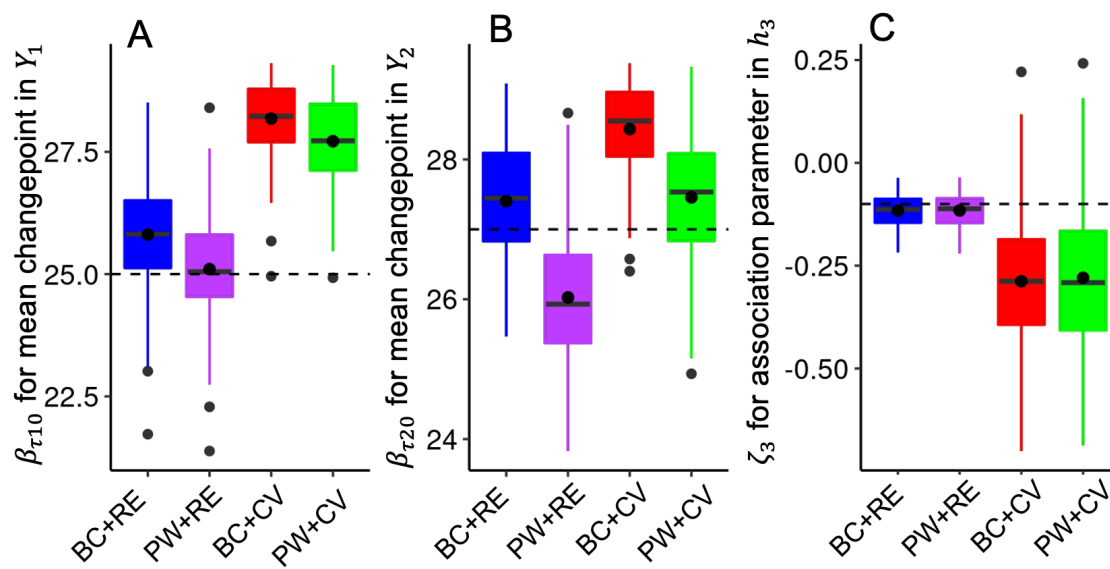


Figure 7. Distributions of posterior means for main changepoint and association parameters in Scenario 3 (community cohort with shared random effect association structure). Panels A, B, and C represent the distributions of estimated posterior means for $\beta_{1\tau_0}$ (the mean changepoint parameter in Y_1), $\beta_{2\tau_0}$ (the mean changepoint parameter in Y_2), and ζ_3 (the association parameter between longitudinal and survival processes in h_3), respectively. In each panel, the boxes denote the posterior means across all simulated datasets based on each method, while the dashed line indicates the true value for the parameter. PW: piecewise model; BC: bent-cable model; RE: shared random effect structure; CV: current value structure.

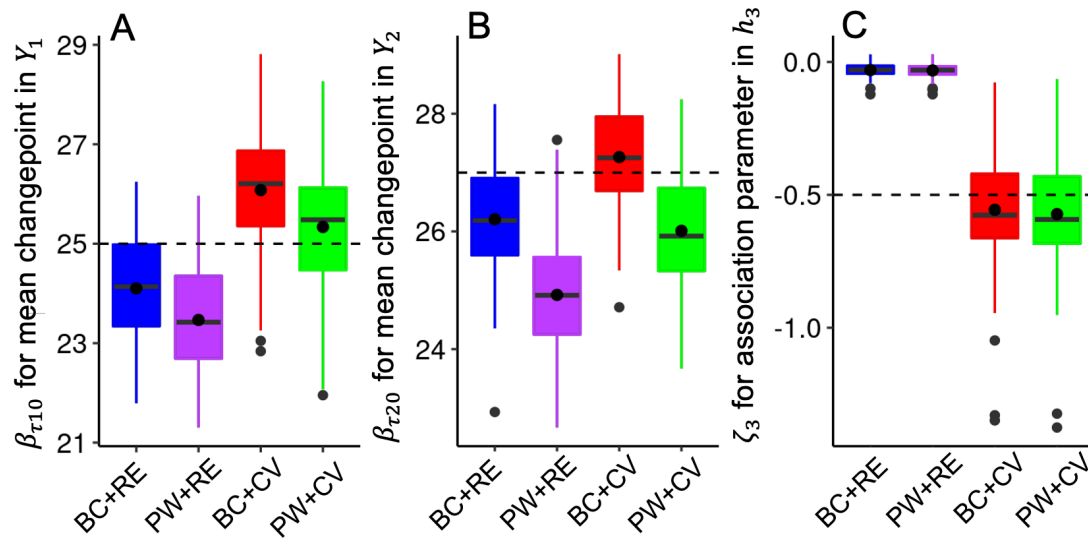


Figure 8. Distributions of posterior means for main changepoint and association parameters in Scenario 4 (community cohort with current value association structure). Panels A, B, and C represent the distributions of estimated posterior means for $\beta_{1\tau_0}$ (the mean changepoint parameter in Y_1), $\beta_{2\tau_0}$ (the mean changepoint parameter in Y_2), and ζ_3 (the association parameter between longitudinal and survival processes in h_3), respectively. In each panel, the boxes denote the posterior means across all simulated datasets based on each method, while the dashed line indicates the true value for the parameter. PW: piecewise model; BC: bent-cable model; RE: shared random effect structure; CV: current value structure.

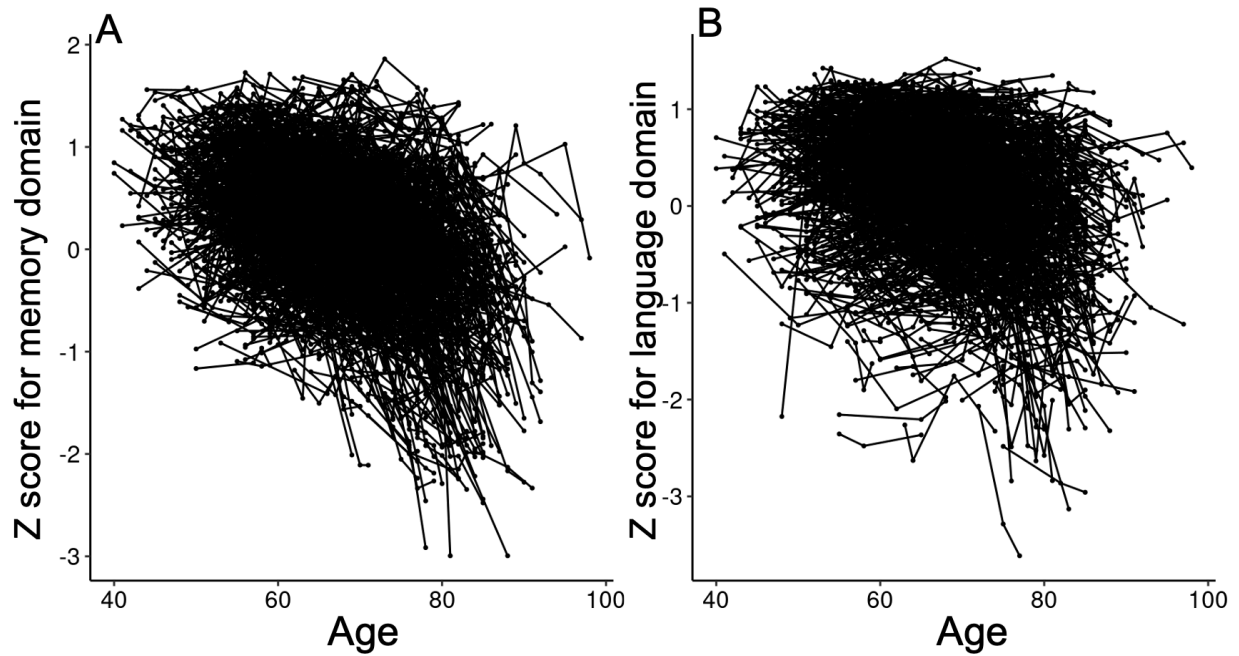


Figure 9. Z-score trajectories for the memory and language domains of neuropsychological tests. Panels A and B show individual z-score trajectories for the memory and language domains, respectively. In each panel, the x-axis represents age, and the y-axis represents z-scores. Each curve denotes a participant's trajectory.

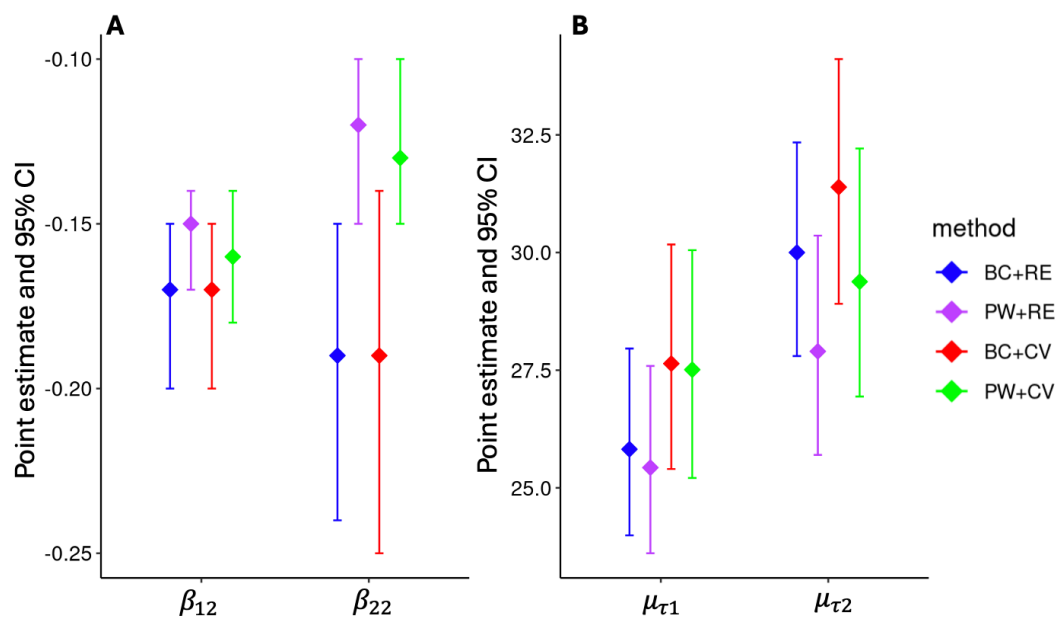


Figure 10. Comparison of estimates for slope differences after the changepoint and mean changepoint location. In Panel A, the bars denote the posterior means and its 95% credible intervals for slope differences after the changepoint based on each method. β_{12} and β_{22} in the x-axis denote slope differences after the changepoint in memory and language cognitive domains, respectively. In Panel B, the bars denote the posterior means and its 95% credible intervals for estimated mean changepoints based on each method. $\mu_{\tau 1}$ and $\mu_{\tau 2}$ in the x-axis denote estimated mean changepoints in memory and language cognitive domains, respectively. PW: piecewise model; BC: bent-cable model; RE: shared random effect structure; CV: current value structure. CI: credible interval.

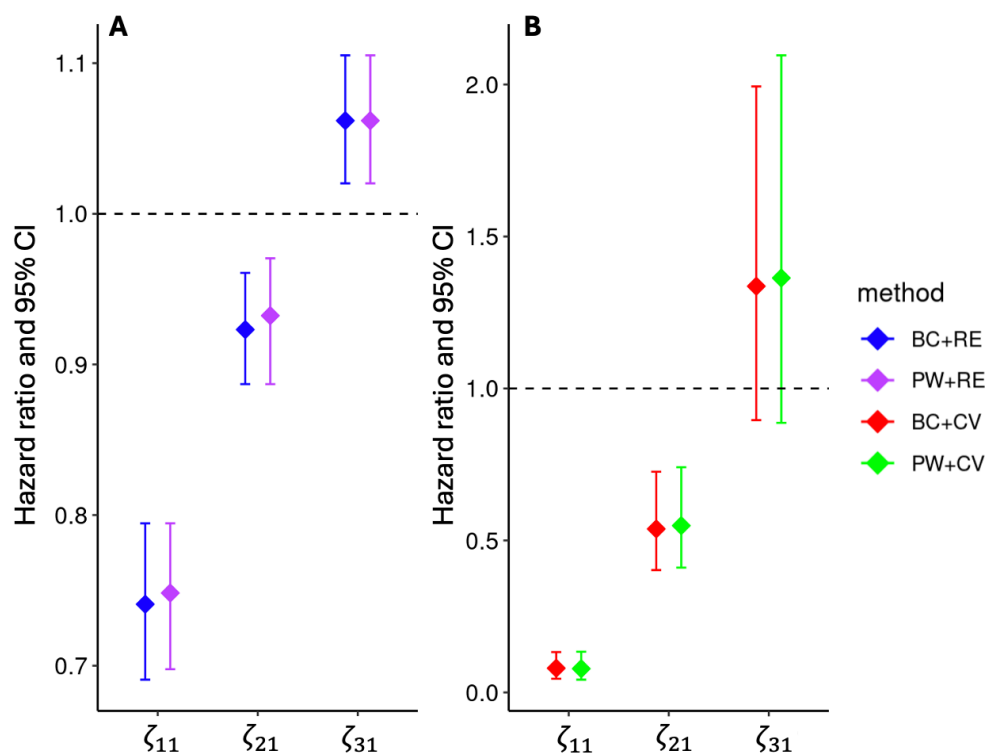


Figure 11. Comparison of estimates for association parameters linking the longitudinal and survival processes. In Panel A, the bars denote the posterior means and its 95% credible intervals for the association parameters between the changepoint in the longitudinal process and the survival process based on each method. ζ_{11} , ζ_{11} , and ζ_{31} in the x-axis denote the association parameters in h_1 , h_2 , and h_3 , respectively. In Panel B, the bars denote the posterior means and its 95% credible intervals for the association parameters between the true unobserved value of the longitudinal cognitive measures and the survival process based on each method. ζ_{11} , ζ_{11} , and ζ_{31} in the x-axis denote the association parameters in h_1 , h_2 , and h_3 , respectively. PW: piecewise model; BC: bent-cable model; RE: shared random effect structure; CV: current value structure. CI: credible interval.