Published in final edited form as:
Biometrics. 2010 March ; 66(1): 294–300. doi:10.1111/j.1541-0420.2009.01261.x.

Joint Modeling for Cognitive Trajectory and Risk of Dementia in the Presence of Death

Binbing Yu and
Laboratory of Epidemiology, Demography and Biometry, National Institute on Aging, Bethesda, Maryland 20892, U.S.A

Pulak Ghosh
Department of Biostatistics and Winship Cancer Institute, Emory University, Atlanta, GA, 30322, U.S.A

Binbing Yu: yubi@mail.nih.gov; Pulak Ghosh: pulakghosh@gmail.com

Summary

Dementia is characterized by accelerated cognitive decline before and after diagnosis as compared to normal aging. It has been known that cognitive impairment occurs long before the diagnosis of dementia. For individuals who develop dementia, it is important to determine the time when the rate of cognitive decline begins to accelerate and the subsequent gap time to dementia diagnosis. For normal aging individuals, it is also useful to understand the trajectory of cognitive function until their death. A Bayesian change-point model is proposed to fit the trajectory of cognitive function for individuals who develop dementia. In real life, people in older ages are subject to two competing risks, e.g, dementia and dementia-free death. Because the majority of people do not develop dementia, a mixture model is used for survival data with competing risks, which consists of dementia onset time after the change-point of cognitive function decline for demented individuals and death time for non-demented individuals. The cognitive trajectories and the survival process are modeled jointly and the parameters are estimated using the Markov chain Monte Carlo method. Using data from the Honolulu Asia Aging Study, we show the trajectories of cognitive function and the effect of education, apolipoprotein E 4 genotype and hypertension on cognitive decline and the risk of dementia.

Keywords

Change point; competing risks; dementia; Markov chain Monte Carlo

1. Introduction

Dementia is a progressive degenerative disease that generally presents with decline in cognitive function over a period of many years. In the preclinical phase, changes can be gradual and usually difficult to distinguish from the less marked decline associated with normal aging. As dementia progresses, cognitive impairments become more obvious and decline in function begins to accelerate. It is important to understand the shape of this decline and the time at which cognitive evolution of subjects who develop dementia becomes distinguishable from that of normal aged subjects (Sliwinski et al., 2006; Wilson et

7. Supplementary Materials
Web Appendices and Figures referenced in Sections 4 and 5 are available under the Paper Information link at the Biometrics website http://www.biometrics.tibs.org.\verb

al., 2003). As more effective treatments become available for dementia, it becomes increasingly important to develop and implement strategies to identify individuals with preclinical dementia at an earlier phase.

Change-point models have been used to describe the trajectory and trend of longitudinal measurements (Smith, 1975; Stephens, 1994). Hall et al. (2000) use a piecewise linear mixed model to compare the trajectories of cognitive functions for incident cases of dementia and for subjects free of dementia at their last follow-up. Later, Hall et al. (2001) use a Bayesian change-point model to describe the cognitive decline of demented subjects. As Jacqmin-Gadda et al. (2006) point out, because subjects who are not demented before the end of the follow-up are classified as non-demented, the analysis may be biased or lack statistical power due to misclassification. Subsequently, Jacqmin-Gadda et al. (2006) combine a piecewise polynomial mixed model with a random change-point for the evolution of the cognitive test and a log-normal survival model for the time from change-point to dementia onset. A Gauss-Hermite quadrature is used to approximate the likelihood function and the maximum likelihood estimates (MLEs) are obtained using the Maquardt optimization algorithm. The model is attractive, but the computation is difficult.

In addition, there are several complicate issues of modeling the trends of cognitive function in aging studies. First, because of the long history of disease progression, subjects without dementia at their last visits may be in the preclinical phase of dementia. Treating them as non-demented will bias the comparison between normal and pathological aging. Second, prevalent cases of dementia are excluded because their onset ages are not known, thus creating left truncation. Third, people may die from other causes without having dementia.

Jacqmin-Gadda et al. (2006) address the first two issues by modeling cognitive decline using a change-point model, but they did not consider the effect of competing risk. In their model, dementia onset time follows a proper survival function, which implies that all subjects would eventually develop dementia. Although dementia is a common disease for older people, most people do not develop dementia in their lifetime and are dementia free at death, producing a competing risk scenario. It is critically important to include the effect of competing risks because many elderly individuals will die while still in the preclinical phase of the development of dementia and cognitive function has been shown to decline significantly with proximity to death (Sliwinski et al., 2006; Wilson et al., 2003).

Here we provide a unified framework that accounts for all the issues together. We considered two competing risks, i.e., dementia and dementia-free death. The Bayesian change-point model for cognitive trajectories and the mixture survival model for dementia onset and death are estimated jointly. The proposed model extends the approach by Jacqmin-Gadda et al. (2006) in several ways: The joint model can handle left-truncated and interval-censored survival data and can easily incorporate multiple covariates. It takes the dementia-free death into account when estimating the time from acceleration of cognitive decline to dementia onset. The Markov chain Monte Carlo method is used for parameter estimation, which is simpler to implement than the direct maximization of likelihood function.

2. Statistical Models

Here, the primary time scale t is age. Let $Y_i(t)$ be the cognitive score of subject i at age t and let x_i be the corresponding covariate vector, which could be time-dependent, $i = 1, \ldots, N$.

2.1 Trajectory of cognitive function

Let $Y_{ij} = Y_i(t_{ij})$ be the cognitive score at at age t_{ij} for $j = 1,\ldots, n_i$ for the ith subject. The scores Y_{ij} can be described by a random change-point model, where

$$Y_{ij}=\mu_0+\sum_{k=1}^{K_1}\mu_{k,i}\times t_{ij}^k+\sum_{k=1}^{K_2}\mu_{K_1+k,i}\times\{(t_{ij}-\tau_i)^+\}^k+\varepsilon_{ij}, \quad (1)$$

where $a^+ = \max(0, a)$, τ_i is the change-point of cognitive function and K_1 and K_2 are the numbers of terms related to age and time from change-point. The independent random error $\varepsilon_{ij} \sim N(0, \sigma_\gamma^2)$. The coefficients $\mu_{k;i}$ may depend on covariates x_i and random effect b_{ki},

$$\mu_{k,i}=\alpha_k+\beta_k x_i+b_{ki},\ \ k=0,\ldots,K_1+K_2, \quad (2)$$

where the random effects $b_{ki} \sim N(0, \sigma_k^2)$. Equation (1) implies a continuous transition at change-point which agrees with the clinical belief of a progressive decline in the pre-diagnosis phase of dementia. Hall et al. (2001) used a quadratic model with $K_1 = K_2 = 2$ and found that the quadratic terms were not significant. While Jacqmin-Gadda et al. (2006) did not use the linear term $(t_{ij} - \tau_i)^+$, in part due computational difficulty.

2.2 Probability of developing dementia

Let D be the type of two competing risks, where $D = 1$ means dementia, $D = 2$ means dementia-free death. The probability of developing dementia is given by

$$\pi_i=P(D_i=1)=\frac{\exp(\delta x_i)}{1+\exp(\delta x_i)}. \quad (3)$$

2.3 Change point of cognitive decline

If a subject dies without dementia, then the change-point of cognitive decline does not occur and we set $\tau_i = \infty$, implying that the subject does not experience the acceleration of cognitive decline before death. If one develops dementia before his death, we assume that the change-point τ_i follows a normal distribution $N(\mu_\tau, \sigma_\tau^2)$ with the constraint that τ_i is between 0 and the maximum age in the data. Furthermore, μ_τ may also be related to covariates $\mu_\tau = \gamma_0 + \gamma_1 x$. Marginally the change-point τ_i takes a mixture distribution:

$$\tau_i\begin{cases}\sim N(\mu_\tau,\sigma_\tau^2) & \text{with probability } \pi_i,\\ =\infty & \text{with probability } 1-\pi_i.\end{cases}$$

2.4 Sub-survival function of competing risks

Conditional on event type D, the survival functions of event times are denoted by $S_d(t) = \Pr(T \geq t|D = d)$, $d = 1, 2$. Here T is the event time, which is the dementia onset age for $D = 1$ and the age of death for $D = 2$. For demented subjects ($D = 1$), we assume that dementia onset age follows a Weibull model with $S_1(t) = \exp(-\lambda_1 t^{r1})$, where λ_1 is related to both covariates and the change-point as $\lambda_1 = \exp(\eta_1 x + \zeta \log \tau)$. For subjects who die without

dementia, the death age also follows a Weibull model with $S_2(t) = \exp(-\lambda_2 t^{r2})$, where $\lambda_2 = \exp(\eta_2 x)$. Thus, the event time takes a mixture distribution with overall survival

$$S(t_i)=\pi_i S_1(t_i)+(1-\pi_i)S_2(t_i). \quad (4)$$

When dementia-free death is treated as censored and $\pi_i = 1$, it reduces to the dementia risk model by Jacqmin-Gadda et al. (2006).

Note that technically the covariates can be used in the components $\mu_{k;i}$, π_i, μ_τ, λ_1 and λ_2. In practice, the selection of a covariate into which component should be guided by scientific evidence and statistical criteria. If too many variables are incorporated or too many random effects are included, the computation could be slow and one may run into convergence problem. To ensure that the meaningful parameter estimates are not driven by the prior, one can pick priors with wide ranges and try different initial values to ensure the estimates are not driven by the prior. Another important issue is the interpretation of the parameter estimates corresponding to the covariates. Inclusion of a covariate in a regression model results in both the outcome and other covariates (X) being adjusted for that new covariate (Z). In a complex nonlinear model, the parameter for Z should be interpreted as the conditional effect of Z adjusting for covariates X.

3. Estimation Method

Because the prevalent cases of dementia are excluded (Joly et al., 1998), all subjects enter the study free of dementia and the event times are left truncated. Let a_i be the age of subject i at the study entry. Usually the death times t_i are observed exactly and the dementia onset times are interval censored at (t_i, r_i). For the subjects who are still alive and dementia-free at the end of followup t_i, the event time is right censored and the event type $D_i = 0$. Let $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_{K_1+K_2})^{\mathrm{T}}$, $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_{K_1+K_2})^{\mathrm{T}}$ and let $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_Y, \sigma_k, \gamma, \sigma_\tau, \delta, \eta_1, \eta_2, \zeta, r_1, r_2)$.

Let y_i be the vector of the n_i measurements for subject i and $f_Y(y_i|\tau_i)$ be the density of y_i given change-point τ_i. Let $f_\tau(\tau_i|D_i)$ be the density of τ_i given event status D_i and let $f_d(t)$ be the density function of event time for event type d, $d = 1, 2$. The likelihood contribution for a subject who develops dementia ($D_i = 1$) during (t_i, r_i) is

$$L_i(\theta;D_i=1)=\pi_i f_\gamma(y_i|\tau_i) f_\tau(\tau_i|D_i=1)\{S_1(t_i)-S_1(r_i)\}/S_1(a_i).$$

and the likelihood contribution for a subject dying at t_i without dementia ($D_i = 2$) is

$$L_i(\theta;D_i=2)=(1-\pi_i) f_\gamma(y_i|\tau_i=\infty) f_2(t_i)/S_2(a_i),$$

and the likelihood contribution for a subject who is alive and dementia free ($D_i = 0$) at t_i is

$$L_i(\theta;D_i=0)=\pi_i f_\gamma(y_i|\tau_i) f_\tau(\tau_i|D_i=1) S_1(t_i)/S_1(a_i)+(1-\pi_i) f_\gamma(y_i|\tau_i=\infty) S_2(t_i)/S_2(a_i).$$

The overall likelihood function is given by

$$L(\theta)=\prod_{i=1}^{n}\prod_{d=0}^{2}L_i(\theta;D_i=d)^{\mathbf{I}(D_i=d)}.$$

Jacqmin-Gadda et al. (2006) used the Maquardt optimization algorithm to find the MLEs. However, the computation is involved and only one covariate education is used.

Because the number of parameters is large and the joint likelihood is quite complicated, the Markov chain Monte Carlo (MCMC) method is used to obtain the parameter estimates (Chen et al., 2000). The MCMC method is implemented by the Bayesian inference package BUGS (Spiegelhalter et al., 1999) and the open source of BUGS, called OpenBugs (Thomas et al., 2006), is imbedded into the R package BRugs (R Development Core Team, 2007). Assuming the elements of $\boldsymbol{\theta}$ are independent of each other, the joint prior $p(\boldsymbol{\theta})$ can be specified as the product of each individual prior. This is a common approach of obtaining "noninformative" prior for joint parameters, which is often justified on the grounds that "ignorance" is consistent with "independence" (Carlin and Louis, 2000, p36). In order that the parameter estimation is driven by the observed data, we assign weakly informative priors. In particular, we specify multivariate (univariate) normal priors for the location parameters $\boldsymbol{\alpha},\boldsymbol{\beta}$, γ, η_1, η_2, δ, ζ, uniform priors for the standard deviations σ_Y, σ_τ, σ_k, k = 0, …, $K_1 + K_2$ (Gelman et al., 2003, p593), and exponential priors for the shape parameters r_1, r_2. The hyperparameters are assumed to be known.

In order to see how stable the final estimates are, multiple independent runs should be carried out. The convergence of the MCMC samples of the parameters $\boldsymbol{\theta}$ after excluding the initial burn-in samples can be diagnosed using several criteria. The common methods of assessing convergence are proposed by Geweke (1992), Heidelberger (1983) and Gelman and Rubin (1992). The Gelman and Rubin method calculates the 'potential scale reduction factor' for each parameter in $\boldsymbol{\theta}$, together with upper and lower confidence limits. Approximate convergence is achieved when the upper limits are close to 1.

4. Application

The Honolulu Heart Program (HHP) is a prospective study of heart disease and stroke involving a cohort of Japanese-American men born between 1900 and 1919 in Hawaii. Clinical and demographic information were collected during three examinations in 1965, 1968–1970, and 1971–1974 (examinations 1–3). As an extension of the HHP, the Honolulu Asia Aging Study (HAAS) was started in 1991 to study dementia prevalence, incidence and risk factors. At the initial examination (examination 4) of the HAAS, 3,734 members of the HHP cohort participated and they were at least 71 years old. Three subsequent follow-up examinations were conducted in 1994, 1997 and 2000 (examinations 5–7). The death times of the participants were recorded accurately until June 20, 2002.

Assessment of cognitive function and dementia was assessed during examinations 4–7 by means of a multistep procedure. The 100-point Cognitive Abilities Screening Instrument (CASI) was used to screen the entire sample. Dementia cases diagnosed in examination 4 were called prevalent cases and the cases diagnosed in examinations 5–7 were called incident cases. There were 226 prevalent cases and 135, 112 and 52 incident cases diagnosed at examinations 4–7, respectively. For details of the dementia diagnosis procedure, see White et al. (1996). Because of the gap time between two examinations, the dementia onset times were interval censored. The prevalent cases were excluded from the analysis because their onset ages were left censored and they did not contribute to the risk set of developing dementia (Joly et al., 1998). The 858 subjects who did not participate in

any follow-up examination were also excluded because they did not carry any information about the change in cognitive function. As a result, there were 2650 people with at least one follow-up exam. The follow-up status and dementia diagnosis are shown in Table 1. The 82 subjects with non-monotone missing visits were excluded because their dementia diagnosis times could not be accurately ascertained. The final data included 2568 subjects with only monotone missing visits.

Three binary covariates, i.e., edu, hyp and apoe are used, where edu is the indiator of having education more than 10 years, hyp is the self reported history of high blood pressure at examination 1 and apoe is the indicator of being homozygous or heterozygous ApoE $\varepsilon 4$ allele carrier. In the cohort, 59% have more than 10 years of education, 17% self-report a history of hypertension at exam 1 and 18% are ApoE $\varepsilon 4$ carriers. The reference group are the subjects with lower education, no hypertension and without AopE allele.

First, we specify the longitudinal model for the trajectory of cognitive function. Because the CASI scores are left skewed, we use the transformed score $Y = \log(101 - \text{CASI})$ as the response (Yip et al., 2002). So a higher transformed score means a lower cognitive function. In Equation (2), only edu is used because education is the most important factor related to cognitive function among the three. The mean change-point $\mu_\tau = \gamma_0 + \gamma_1 \text{edu}$. Initially we let $K_1 = 2$ and $K_2 = 2$, which assumes the cognitive decline for healthy aging and acceleration of cognitive decline are both quadratic. We find that the quadratic term and random slope before the change-point are close to 0 and its 95% confidence interval also covers 0, hence these two terms are dropped. In the final model for cognitive function, $K_1 = 1$, $K_2 = 2$, and

$$Y_{ij}=(\alpha_0+\beta_0\text{edu}+b_{0i})+(\alpha_1+\beta_1\text{edu})t_{ij}+(\alpha_2+\beta_2\text{edu}+b_{2i})(t_{ij}-\tau_i)^+ +(\alpha_3+\beta_3\text{edu}+b_{3i})\{(t_{ij}-\tau_i)^+\}^2+\varepsilon_{ij}. \quad (5)$$

This model is similar to the one by Jacqmin-Gadda et al. (2006), but here equation (5) contains a term $(t_{ij} - \tau_i)^+$ after the change-point. All three covariates are used for the probability of developing dementia. Both the variables edu and hyp are used in the scale parameters λ_1 and λ_2 for the conditional sub-survival functions. In summary, $\mu_k = \alpha_k + \beta_k \text{edu} + b_k$ in Equation (2), $\pi = [1 + \exp\{-(\delta_1 + \delta_2 \text{apoe} + \delta_3 \text{edu} + \delta_4 \text{hyp})\}]^{-1}$ in Equation (3), $\lambda_1 = \exp(\eta_{11} + \eta_{12}\text{hyp} + \eta_{13}\text{edu} + \zeta \log \tau)$, and $\lambda_2 = \exp(\eta_{21} + \eta_{22}\text{hyp} + \eta_{23}\text{edu})$.

Weakly informative priors are used for parameters. In particular, the prior for each parameter in $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \gamma, \delta, \eta_1, \eta_2, \zeta)$ is normal with mean 0 and variance 0.001. Each of the standard deviation parameters is given a uniform prior U(0, 10). All these priors are highly dispersed compared to their parameter ranges. Two MCMC chains with widely dispersed initial values are simulated. The initial values for the fixed parameters are selected by starting with the prior mean and covering ±2 standard deviations. Because of the large sample size and slow convergence of the change-point model, 30, 000 samples are excluded and the next 30, 000 samples with thinning factor 5 are used to calculate the posteriors. The convergence of the MCMC simulations is reached based on the criteria described in Section 3.

Table 2 shows the parameter estimates and their standard deviations (SD) and 95% confidence intervals (CI) for the longitudinal model for cognitive function and for the mixture survival model for competing risks. The trajectories of the CASI score by dementia status and education level are plotted in Figure 1. Jacqmin-Gadda et al. (2006) only show the cognitive trajectories for subjects who developed (or will develop) dementia, here we are able to calculate the trajectories for both demented and non-demented. For non-demented, the scores are 90 and 93 at age 70 for individuals with lower and higher education and the

trajectories are parallel by education level. For demented, the change-points are at age 73 and 85 for lower and higher education levels, respectively. The change-point for subjects with higher education occurs later but the cognitive score declines with a faster rate compared to those with lower education. At age 95, all demented subjects have similar cognitive function regardless of their education level. The change-points are slightly different from those found by Jacqmin-Gadda et al. (2006). These can be due to several reasons: The study populations are different; the test instruments for cognitive function are different; the education cutoffs are different and the diagnostic criteria might be applied more or less aggressively. But both studies yield similar trajectories of cognitive function by education level. The late change-point for subjects with higher education suggests the effect of cognitive reserve.

For the probability of developing dementia, the subjects with higher education have a significantly lower risk of developing dementia. Although the AopE 4 genotype and hypertension increases the risk of dementia, their effects are not significant. The parameters for sub-survival functions are also presented in Table 2. To evaluate the fit of the Weibull models, we compare the marginal sub-incidence functions for dementia onset and dementia-free death from the Weibull model and the method by Fine and Gray (1999). The cumulative incidence curves from the two methods are almost identical for dementia onset and are reasonably close for death (see supplementary documents for details), which indicates the departure from the model assumption is not great. For demented, if two subjects have the same change-point, the one with higher education has higher risk of dementia onset. For non-demented, having hypertension significantly increases the risk of death, where education is not significant.

Table 3 shows the estimates and the standard errors (SEs) of median years from the change-point to dementia onset by education. When the change-point is at age 75, the medians are 8.3 and 4.1 years for the subjects with lower and higher education. For the subjects with lower education, the medians from change-point to dementia remain similar for different change-points. For the subjects with higher education, the medians are smaller for late change-point, which implies that the cognitive function detioriates even faster if the changes occurs later. The SEs for the subjects with lower education are wider. The trends of the SEs by change point and education are consistent with the trends by Jacqmin-Gadda et al. (2006).

5. Simulation

Conceptually, the difference of the proposed method and the method by Jacqmin-Gadda et al. (2006) is whether death is treated as censored or as a competing risk. A simulation study was carried out to examine the effect of modeling dementia-free death. For details of simulation setting and parameter specifications, see supplementary documents.

A single binary covariate x is considered and the cognitive function is given by

$$Y(t)=\alpha_0+\beta_0 x+(\alpha_1+\beta_1 x)t+(\alpha_2+\beta_2 x)(t-\tau)^+ +\varepsilon_\gamma,\ \ \varepsilon_\gamma \sim N(0,\sigma_\gamma^2),$$

where the change-point τ takes a mixture distribution with dementia probability $P(D = 1) = 1/[1 + \exp\{-(\delta_0 + \delta_1 x)\}]$ and the mean of τ is $\mu_\tau = \gamma_0 + \gamma_1 x$. The survival function for dementia onset is $S_1(t) = \exp(-\lambda_1 t^{r1})$ with $\lambda_1 = \exp(\eta_{11} + \eta_{12} x + \zeta \log \tau)$ and the survival function for dementia-free death is $S_2(t) = \exp(-\lambda_2 t^{r2})$ with $\lambda_2 = \exp(\eta_{21} + \eta_{22} x)$. The truth parameter values are specified based on the application. We set $(\eta_{21}, \eta_{22}) = (-9.2, -1)$ or $(-10.2, 0)$ to examine the effect of competing risk. The standard deviation of the cognitive

function $\sigma_Y = 0.1$ or 0.05. We assume the maximum follow-up time is 10 years and the cognitive functions are observed every 2 years until the event, i.e., dementia onset or death. Each dataset consists of 200 observations and x=0 or 1 for half of the data. Both the proposed method (PM) and the Jacqmin-Gadda et al's method (JGM) are applied to each dataset. The simulation shows the parameters for cognitive trajectories are unbiased for both methods. Of primary interest are the parameters for the risk of dementia onset (η_{11}, η_{12}, ζ, r_1) and for the change-point (γ_0, γ_1). The final estimates are the averages based on 100 replicates and are summarized in Table 4.

We see that the results are similar for $\sigma_Y = 0.1$ or 0.05. The estimates from the proposed method for all four scenarios are nearly unbiased, which is reasonable because they are based on the true model. Without considering the competing risk, the estimates from the JGM are biased, reflecting the impact of model misspecification. For example, in the true model $\eta_{12} = 1.5$ indicating a faster decline for $x = 1$. Because x has simultaneous effect on dementia-free death, it could lead a contrary effect of x on dementia if the competing risk is not taken into account. The median dementia onset years from the beginning of the study is $(\log(2)/\exp(\eta_{11}))^{\frac{1}{r_1}}$ for the reference group. By treating dementia-free death as censored, the JGM tends to over estimate the median onset age. But the magnitude of the over-estimation depends on the effect of x on dementia-free death. For the location of the change-points, the JGM tends to over estimate the change-point for $x = 0$. The underlying model assumes a mixture distribution for the change-point. The JGM assumes that every subject develops dementia eventually and the trajectory, if not censored, consists of a change-point, so the estimate from the JGM is in fact an average of true change-point and maximum age. The simulation shows that the estimates by treating death as censored might be biased if the underlying model is a mixture model. The amount of bias depends on the simultaneous effect of covariates on the dementia-free death.

6. Discussion

In the spirit of Hall et al. (2001), we propose a Bayesian model with greater complexity but simpler implementation than that of Jacqmin-Gadda et al. (2006). The alleviates the practical difficulties in fitting cognitive trajectories and incidence model jointly through maximum likelihood. This model can be used to model the trends of biomarkers in the presence of competing risks. By modeling the longitudinal trajectories of cognitive function, one can calculate the posterior probability of being healthy or demented (Skates et al., 2001), which can be used for effective screening of dementia. Here we assume a truncated normal distribution for the change-point. Because the probability of cognitive decline acceleration should increase with age, an asymmetric model with higher probability for older ages might be more appropriate. A possible extension is to use a Dirichlet process prior to estimate the nonparametric distribution of a change-point.

However, there are several limitations about the current study. First, the cognitive examinations in the HAAS had intervals of approximately three years. Many longitudinal studies have observations at more frequent intervals, which would seem to offer improved estimation of change-points. Statistically, it is useful to determine the ideal gap time between two consequent examinations to improve the diagnosis of dementia as well as the estimation of change-point. Second, the dementia cases are identified using a multi-step procedure and there exists verification bias. The misclassification may cause bias to the results. A subject who misses the previous diagnosis may either die before the next examination or would be diagnosed in later examinations. Using the mixture model, the subject who dies without diagnosis could be classified as preclinical phase of dementia. The subject who is diagnosed in later examinations then has an 1–2 year delay in dementia

diagnosis. Third, although the parametric Weibull models provides a good fit to the mortality and dementia incidence rates, for general modeling of age-specific disease rate, the parametric model may not be ideal and the penalized likelihood approach (Joly et al., 1998) is a useful alternative.

The current approach provides a different perspective to disease history in the presence of competing risks. The approach by Jacqmin-Gadda et al. (2006) considers disease progression by treating deaths as censored and it can help understand the neurological progression and biological etiology of the disease, while the current approach is appropriate to describe the actual disease history because death is the primary competing risk of dementia for the old. If the primary end point is cognitive function and dementia onset time, one does not need to model the age of dementia-free death. This amounts to fitting a cure model for dementia onset. The application of the cure model to dementia onset with longitudinal biomarkers is an extension of the current model.

Here we focus on cognitive function before the clinical diagnosis of dementia, and the cognitive test scores after the diagnosis are excluded. It is also of interest to know the trajectory of cognitive function after dementia diagnosis in order to help understand the progression of the disease. One common complication of the cognitive measure is the floor and ceiling effects. This makes the assessment of cognitive change difficult because there is not enough variation in the outcomes across individuals. Harvey et al. (2003) suggests transformation as a partial remedy and we use the transformation log(101−CASI). One option is to model the observed test score as a truncated sample of the latent cognitive score. However, this will involve more complicated models for the outcome distributions. In the analysis of competing risk data, the censoring or the competing risk might be correlated, it is also useful to examine the effect of informative censoring on the parameter estimates.

In summary, the findings in this paper indicate the effect of cognitive reserve, i.e., the acceleration of cognitive decline for the subjects with higher education occurs later but with a faster rate compared to the subjects with lower education. This offers an independent reinforcement of the results first reported empirically in Hall et al. (2007), which confirms an important theoretical prediction by Stern et al (1999).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported in part by the Intramural Research Program of the National Institute on Aging. We thank two anonymous reviewers and the associate editor for their helpful suggestions.

References

Carlin, BP.; Louis, TA. Bayes and Empirical Bayes Methods for Data Analysis. 1. Chapman and Hall; London: 1996.

Chen, M-H.; Shao, Q-M.; Ibrahim, JG. Monte Carlo Methods in Bayesian Computation. New York: Springer; 2000.

Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. Journal of the American Statistical Association. 1999; 94:496–509.

Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Statistical Science. 1992; 7:457–511.

Gelman, A.; Carlin, JB.; Stern, HS.; Rubin, DB. Bayesian Data Analysis. 2. New York: Chapman and Hall; 2003.

Geweke, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: Bernado, JM.; Berger, JO.; Dawid, AP.; Smith, AFM., editors. Bayesian Statistics. Vol. 4. Clarendon Press; Oxford, UK: 1992.

Hall CB, Lipton RB, Sliwinski M, Stewart WF. A change-point model for estimating the onset of cognitive in preclinical Alzheimer's disease. Statistics in Medicine. 2000; 19:1555–1566. [PubMed: 10844718]

Hall CB, Ying J, Kuo L, Sliwinski M, Buschke H, Katz M, Lipton RB. Estimation of bivariate measurements having different change-points, with application to cognitive ageing. Statistics in Medicine. 2001; 20:3695–3714. [PubMed: 11782027]

Hall CB, Derby C, LeValley A, Katz MJ, Verghese J, Lipton RB. Education delays accelerated decline on a memory test in persons who develop dementia. Neurology. 2007; 69:1657–1664. [PubMed: 17954781]

Harvey PD, Green MF, McGurk SR, Meltzer HY. Changes in cognitive functioning with risperidone and olanzapine treatment: a large-scale, double-blind, randomized study. Psychopharmacology. 2003; 169:404–411. [PubMed: 12590356]

Heidelberger P, Welch PD. Simulation run length control in the presence of an initial transient. Operations Research. 1983; 31:1109–1144.

Jacqmin-Gadda H, Commenges D, Dartigues JF. Random changepoint model for joint modeling of cognitive decline and dementia. Biometrics. 2006; 62:254–260. [PubMed: 16542253]

Joly P, Commenges D, Letenneur L. A prenalized likelihood approach for aritrarily censored and truncated data: application to age-specific incidence of dementia. Biometrics. 1998; 54:185–194. [PubMed: 9574965]

Larson MG, Dinse GE. A mixture model for the regression analysis of competing risks data. Applied Statistics. 1985; 34:201–211.

Skates SJ, Pauler DK, Jacobs IJ. Screening based on the risk of cancer calculation from Bayesian hierarchical changepoint and mixture models of longitudinal markers. Journal of the American Statistical Association. 2001; 96:429–439.

Smith AFM. A Bayesian approach to inference about a changepoint in a sequence of random variables. Biometrika. 1975; 62:407–416.

Spiegelhalter DJ, Best NG, Carlin BP, Linde AVD. Bayeisan measures of model complexity and fit. Journal of the Royal Statistical Society, Seris B. 2002; 64:583–639.

Stephens DA. Bayesian retrospective multiple-changepoint identification. Applied Statistics. 1994; 43:159–178.

Stern Y, Albert S, Tang MX, Tsai WY. Rate of memory decline in AD is related to education and occupation: cognitive reserve? Neurology. 1999; 53:1942–1957. [PubMed: 10599762]

R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: 2007. http://www.R-project.org

Thomas A, O'Hara B, Ligges U, Sturtz S. Making BUGS Open. R News. 2006; 6:12–17.

Sliwinski MJ, Stawski RS, Katz M, Verghese J, Lipton R. On the importance of distinguishing pre-terminal and terminal cognitive decline. European Psychologist. 2006; 11:172–181.

Spiegelhalter, DJ.; Thomas, A.; Best, NG. WinBUGS Version 1.2 User Manual. MRC Biostatistics Unit; 1999.

White L, Petrovitch H, Ross GW, Masaki KH, Abbott RD, Teng EL, Rodriguez BL, Blanchette PL, Havlik RJ, Wergowske G, Chiu D, Foley DJ, Murdaugh C, Curb JD. Prevalence of dementia in older Japanese-American men in Hawaii: The Honolulu-Asia Aging Study. Journal of the American Medical Association. 1996; 276:955–960. [PubMed: 8805729]

Wilson RS, Beckett LA, Bienias JL, Evans DA, Bennett DA. Terminal decline in cognitive function. Neurology. 60:1782–1787. [PubMed: 12796531]

Yip AG, Brayne C, Easton D, Rubinsztein DD. An investigation of ACE as a risk factor for dementia and cognitive decline in the general population. Journal of Medical Genetics. 2002; 39:403–406. [PubMed: 12070247]

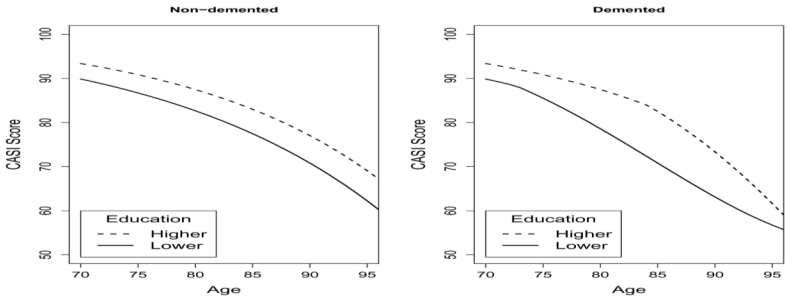

Figure 1.
Trajectories of cognitive test CASI scores for demented and non-demented subjects by education level *
*The demented subjects include those who already developed dementia and who will develop dementia according to their cognitive history

Table 1

Follow-up status and dementia diagnosis for the 2650 subjects with at least one follow-up examination

				Dementia diagnosis			
EX5	**EX6**	**EX7**	**All**	**Not diagnosed**	**EX5**	**EX6**	**EX7**
No	No	Yes	22	21	.	.	1
No	Yes	No	16	12	.	4	.
No	Yes	Yes	20	16	.	3	1
Yes	No	Yes	24	22	.	.	2
Yes	No	No	648	579	69	.	.
Yes	Yes	No	477	406	34	37	.
Yes	Yes	Yes	1443	1295	32	68	48
	Overall		2650	2351	135	112	52

Table 2

Parameter estimates for cognitive trajectories and the mixture model for competing risks

Parameter	Mean	Median	SD	95% CI
Longitudinal model for cognitive function				
α_0	2.410	2.406	0.033	(2.359, 2.505)
α_1	0.050	0.050	0.003	(0.039, 0.055)
α_2	0.038	0.037	0.012	(0.017, 0.065)
α_3	−0.143	−0.137	0.104	(−0.378, 0.071)
β_0	−0.379	−0.375	0.032	(−0.448, −0.320)
β_1	0.007	0.008	0.003	(0.000, 0.013)
β_2	−0.008	−0.010	0.024	(−0.049, 0.041)
β_3	0.042	0.054	0.075	(−0.143, 0.167)
Parameters for the change-point				
γ_1	1.682	1.591	0.405	(1.030, 2.580)
γ_2	11.410	11.647	0.493	(10.043, 11.940)
Logistic model for the probability of developing dementia				
δ_1	−1.368	−1.370	0.122	(−1.605, −1.128)
δ_2 (apoe)	0.122	0.124	0.171	(−0.212, 0.453)
δ_3 (edu)	−0.454	−0.456	0.166	(−0.771, −0.116)
δ_4 (hyp)	0.103	0.104	0.198	(−0.291, 0.488)
Sub-survival function for dementia after change-point				
η_{11}	−4.648	−4.622	1.991	(−9.450, −0.129)
η_{12} (hyp)	−0.105	−0.096	0.233	(−0.599, 0.321)
η_{13} (edu)	0.992	1.150	1.025	(−1.445, 2.974)
$\zeta(\tau)$	−1.210	−1.386	1.303	(−4.284, 1.775)
r_1	2.373	2.360	0.235	(1.957, 2.895)
Sub-survival function for dementia-free death				
η_{21}	−10.086	−10.077	0.314	(−10.73, −9.475)
η_{22} (hyp)	0.274	0.278	0.092	(0.092, 0.444)
η_{23} (edu)	0.016	0.016	0.077	(−0.136, 0.161)
r_2	3.442	3.439	0.103	(3.242, 3.654)

Table 3

Estimate and the standard error (SE) of median years from change-point of cognitive decline to dementia

	Median years from change-point to dementia			
	Lower Education		Higher Education	
Change point τ	Estimate	SE	Estimate	SE
75	8.3	1.72	4.1	2.34
80	9.6	3.07	3.3	0.62
85	9.3	4.98	1.4	0.42

Table 4

Estimates of the parameters of dementia onset and change points from the two methods

	σ_Y = 0.05				σ_Y = 0.10			
(η_{21}, η_{22})	**(−9.2, −1)**		**(−10.2, 0)**		**(−9.2, −1)**		**(−10.2, 0)**	
Parameter value	**PM**	**JGM**	**PM**	**JGM**	**PM**	**JGM**	**PM**	**JGM**
Parameters of the incidence of dementia onset								
η_{11} = −4.6	−4.51	−2.30	−4.46	−3.30	−4.99	−1.90	−4.55	−2.82
η_{12} = 1.5	1.66	−0.49	1.67	−0.58	1.55	−0.18	1.40	−0.26
ζ = −1.5	−1.74	−0.76	−1.58	−0.22	−1.52	−1.10	−1.56	−0.62
r_1 = 2.4	2.88	0.85	2.37	0.73	2.60	0.86	2.50	0.75
Median onset time for the reference group (year)								
Median (5.8)	4.2	9.7	5.6	55.6	5.9	5.9	5.3	25.7
Parameters of change-point								
γ_1 = 5	5.04	6.64	4.76	10.54	4.89	5.96	4.46	8.09
γ_2 = 8	8.01	7.13	7.52	1.50	7.90	6.56	8.36	3.64

Biometrics. Author manuscript; available in PMC 2011 February 10.