Published in final edited form as:
Stat Med. 2013 March 15; 32(6): 1038–1053. doi:10.1002/sim.5557.

Bivariate random change point models for longitudinal outcomes

Lili Yang[*,†] and **Sujuan Gao**
Department of Biostatistics, Indiana University School of Medicine, 410 W. 10th Street, Suite 3000, Indianapolis, IN 46202-3002, USA

Abstract

Epidemiologic and clinical studies routinely collect longitudinal measures of multiple outcomes, including biomarker measures, cognitive functions, and clinical symptoms. These longitudinal outcomes can be used to establish the temporal order of relevant biological processes and their association with the onset of clinical symptoms. Univariate change point models have been used to model various clinical endpoints, such as CD4 count in studying the progression of HIV infection and cognitive function in the elderly. We propose to use bivariate change point models for two longitudinal outcomes with a focus on the correlation between the two change points. We consider three types of change point models in the bivariate model setting: the broken-stick model, the Bacon–Watts model, and the smooth polynomial model. We adopt a Bayesian approach using a Markov chain Monte Carlo sampling method for parameter estimation and inference. We assess the proposed methods in simulation studies and demonstrate the methodology using data from a longitudinal study of dementia.

Keywords

random change point model; longitudinal bivariate outcomes; Bayesian method

1. Introduction

Longitudinal epidemiologic and clinical studies routinely collect repeated measures of multiple outcomes. For example, in longitudinal studies of dementia, cognitive function measures, activity of daily living measures, physical function measures such as height and weight, neurological measures, and psychosocial measures are collected repeatedly from participants over a relatively long follow-up period. In recent years, research on Alzheimer's disease (AD) has come to the consensus that both AD pathological processes and the clinical decline occur gradually, with dementia at the end stage of many years of accumulation of these pathological changes [1]. An additional feature of AD is that biological changes begin to develop decades before the presentation of earliest clinical symptoms. Longitudinal measures of biomarkers, cognitive functions, and clinical symptoms will enable researchers to establish the temporal order of relevant biological processes and their association with the onset of clinical symptoms.

Change point models are useful as an alternative to linear models to determine when changes have taken place in an event window. Change point models with one change point

[*]Correspondence to: Lili Yang, Department of Biostatistics, Indiana University School of Medicine, 410 W. 10th Street, Suite 3000, Indianapolis, IN 46202-3002, USA.
[†]yanglili@iupui.edu

and two linear phases are most commonly used because many biological mechanisms can be readily modeled. To account for individual variability, random change point models have been further formulated by including flexible subject-specific random effects to capture both population trends and individual-level variations. Univariate change point models have been used to model various clinical endpoints such as CD4 count in studying the progression of HIV infection and AIDS [2–4] and cognitive function in studying dementia in the elderly [5–8].

The simple change point model with an abrupt transition is referred to as the broken-stick model [3, 4, 7], which has the advantage of detecting a significant departure in direction and volatility from the immediate past. However, the broken-stick model is not always appropriate in practice because a sudden change in direction may not be realistic. The non-continuity at the change point of the broken-stick model may also cause numerical issues in parameter estimation. Two types of smooth change point models were proposed by van den Hout *et al*. [8]: the Bacon–Watts model [9] and a smooth polynomial model.

There have been a few studies on the joint modeling of bivariate random change point model for longitudinal outcomes. Hall *et al*. [10] simultaneously estimated two different change points of two longitudinal measures of cognitive function. Jacqmin-Gadda [6] constructed joint models between a random change point model for a longitudinal outcome and a lognormal model for time-to-event data. In this paper, we consider bivariate change point models for two longitudinal outcomes with a focus on the correlations between the two change points. Motivated by data from a longitudinal study of dementia, we develop joint models for bivariate longitudinal outcomes under the aforementioned modeling frameworks: the random broken-stick model, the random Bacon–Watts model, and the random smooth polynomial model. The proposed bivariate change point models take the correlation structure into account and provide a useful framework to assess the correlation between the two change points and their temporal order. The proposed methodology is applicable to other studies in which determining the order of biomarker changes is needed. We adopt a Bayesian estimation approach using Markov chain Monte Carlo (MCMC) for a computational and inferential framework for the bivariate random change point models. We assess the performance of the proposed method in simulation studies and demonstrate the methodology using data from a longitudinal study of dementia.

We organize the remainder of this paper as follows. Section 2 describes a longitudinal study of dementia as a motivating example. In Section 3, we present three bivariate random change point models, the Bayesian methodology for parameter estimation, and statistical inference. We carried out a series of simulation studies to compare the performances of the three joint models, and we present results in Section 4. In Section 5, we apply the proposed methods to the example data set. We conclude with a discussion in Section 6.

2. The Indianapolis–Ibadan Dementia Study

The Indianapolis–Ibadan Dementia Study (IIDS) is a longitudinal comparative epidemiology study designed to investigate risk factors associated with dementia and AD. The study enrolled and maintained two cohorts of elderly participants, one consisting of African Americans living in Indianapolis, Indiana, and the other consisting of Nigerians living in Ibadan, Nigeria. Details about the study have been published [11, 12]. The data used for the current paper come from the Indianapolis cohort. Briefly, 2212 African American adults aged 65 years and older living in Indianapolis were enrolled in the study in 1992. The study participants were followed for up to 17 years and underwent regularly scheduled cognitive assessments and clinical evaluations approximately every 2 or 3 years. In this ongoing study, there were seven evaluations by the end of 2009.

The cognitive function of study participants was measured by the Community Screening Interview for Dementia (CSID) at baseline and at years 3, 6, 9, 12, 15, and 17 with respect to the baseline. The CSID questionnaire [13] has been widely used as a screening tool for dementia. It evaluates multiple cognitive domains including language, attention, memory, orientation, praxis, comprehension, and motor response. For this analysis, we use a CSID score that incorporated all cognitive items from the screening exam, some of which had not been utilized previously [13]. The additional cognitive score items in the CSID include the East Boston story (immediate and delayed recall), three mental calculation items, the name of the state, the name of the president, and the name of the governor. In addition, unit weighting was used for object repetition, object recall, instruction command, and animal naming, with the exception that animal naming is capped at a maximum raw score of 23 (95th percentile). The CSID total score ranges from 0 to 80, with higher scores indicating better cognitive function.

Also, height and weight measures from all participants were collected at each evaluation starting from year 3. Because obesity is associated with increased risk for diabetes, hypertension, and cardiovascular diseases, conditions related to increased risk of dementia, it is therefore important to monitor weight change in this elderly cohort. It is widely known that subjects with dementia and cognitive impairment suffer weight loss, which can be attributed to the fact that these subjects often forget to eat. However, there are also reports that weight loss precedes dementia diagnosis [14]. In particular, in this cohort, we found that accelerated weight loss was associated with dementia or mild cognitive impairment as early as 6 years prior to clinical diagnosis, supporting the hypothesis that weight loss is an early marker for the manifestation of the dementia disorder, including the early stage of mild cognitive impairment [15]. It is important to examine the longitudinal trajectories of both cognitive function and weight measures to determine whether cognitive decline leads to weight loss or whether weight change proceeds cognitive impairment. Because both body weight and cognitive function are assumed to be stable over time and sudden changes may indicate underlying disease processes, we propose to use bivariate change point models to model cognitive trajectories and changes in body mass index (BMI) over time, with a particular focus on the correlation between the change points of the two trajectories. Here, BMI is defined as weight in kilograms divided by height in meters squared. We choose to use only two change points on the basis of the study design of the IIDS data. The IIDS followed normal subjects without dementia to dementia diagnosis, and no data were collected once a subject was diagnosed with dementia. Because both body weight and cognitive function in elderly subjects without dementia are assumed to be stable over time and sudden changes may indicate underlying dementia progression, we believe one change point for each longitudinal trajectory should capture the decline in the pre-dementia or earlier dementia stage. It is possible that there exists a second change point reflecting a rapid deterioration in both body weight and cognition just prior to death. However, because the IIDS did not conduct any follow-up evaluations in the subjects with dementia and our evaluation interval window of every 2 to 3 years may be too wide to capture the rapid changes in the second change points, we focused on models with only one change point.

Out of the 2212 IIDS participants enrolled at baseline, 441 had at least five cognitive measurements, of which 238 also had at least five BMI measurements. For modeling purposes, we restrict the data to participants with at least five measurements for both of cognitive function and BMI ($N = 238$). Out of the 238 subjects with age ranges from 64.3 to 84.6 years at baseline, 190 (79.8%) subjects were female. The mean baseline age was 70.4 (SD=4.8) years, and the mean years of education was 10.8 (SD=2.6). The mean cognitive scores at baseline and visit 6 were 70.6 (SD=6.0) and 65.4 (SD=9.6), respectively. BMI measures were collected starting from visit 1, and the mean BMI at visits 1 and 6 were 29.1 (SD=5.1) and 26.7 (SD=5.3), respectively. The histogram plots of cognitive function and

BMI measures at baseline were also explored. Although CSID scores are slightly skewed toward lower scores, we assumed normal distributions for both CSID scores and BMI measures. We investigated the robustness of our proposed methods to non-normal distributions in simulation studies. Figure 1 shows the cognitive and BMI trajectories from five randomly selected IIDS participants. We can see from Figure 1 that, in general, both cognitive score and BMI decrease with age. We noted that these 238 participants used in our analysis are survivors with relatively long follow-up information, and they expected to be healthier than others in the cohort who did not provide five measurements. In Section 6, we provide further discussion on the impact of missing data due to death and its potential impact on our analysis results.

3. Statistical methods

In this section, we define notations and introduce three different random change point models for longitudinal outcomes. For each longitudinal outcome, we consider the random change point model with one change point that can be further extended to multiple change points. Let t_{ij} be the time of the j th longitudinal measurement for the i th subject, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m_i$; y_{1ij} and y_{2ij} are the bivariate longitudinal outcomes for the i th subject at time t_{ij}.

3.1. Broken-stick model

For the i th subject at time t_{ij},

$$y_{1ij}=\alpha_{1i}+\alpha_{2i}\left(t_{ij}-\alpha_{4i}\right)I_{(-\infty,\alpha_{4i})}\left(t_{ij}\right)+\alpha_{3i}\left(t_{ij}-\alpha_{4i}\right)I_{[\alpha_{4i},\infty)}\left(t_{ij}+\epsilon_{1ij}, \quad (1)\right.$$

$$y_{2ij}=\alpha_{5i}+\alpha_{6i}\left(t_{ij}-\alpha_{8i}\right)I_{(-\infty,\alpha_{8i})}\left(t_{ij}\right)+\alpha_{7i}\left(t_{ij}-\alpha_{8i}\right)I_{[\alpha_{8i},\infty)}\left(t_{ij}+\epsilon_{2ij}, \quad (2)\right.$$

where α_{4i} and α_{8i} denote the change points for y_{1ij} and y_{2ij}, respectively. α_{1i} and α_{5i} represent the intercepts in the two models and can be interpreted as the mean values of longitudinal outcomes at change points α_{4i} and α_{8i}, respectively. α_{2i} and α_{6i} denote the slopes before the change points, and α_{3i} and α_{7i} denote the slopes after the change points. ϵ_{1ij} and ϵ_{2ij} denote the residual errors of the longitudinal measurements, which are independently distributed as $\epsilon_{1ij}\tilde{}_{\text{iid}}N\left(0,\sigma_{e_1}^2\right)$ and $\epsilon_{2ij}\tilde{}_{\text{iid}}N\left(0,\sigma_{e_2}^2\right)$. $I_A(\cdot)$ is an indicator function with $I_A(x) = 1$ for $x \in A$ and $I_A(x) = 0$ for $x \notin A$.

In addition, we assume a multivariate distribution for the parameters in models (1) and (2),

$$\alpha_i=(\alpha_{1i},\alpha_{2i},\alpha_{3i},\alpha_{4i},\alpha_{5i},\alpha_{6i},\alpha_{7i},\alpha_{8i})^T\tilde{}\,\text{MVN}\left(\alpha,\Sigma_\alpha\right),$$

where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8)^T$ is an 8×1 vector with each entry representing the population mean, and Σ_α is the 8×8 variance–covariance matrix.

The broken-stick model can be implemented using a Bayesian framework and has simple parameter interpretation. However, it is not always appropriate because a sudden change in direction may not be a realistic assumption. Furthermore, the non-continuity at the change point can also cause numerical problems in parameter estimation using the frequentist method, such as the maximum likelihood method. Thus, there is a need to investigate other models not hampered by the disadvantages of the broken-stick model. Here, we use some IIDS data analysis results from Section 5 as an example to illustrate the three different models. In Figure 2, the black dots denote the cognitive function measures for a randomly

selected individual, and the black solid line illustrates the predicted broken-stick curve of this individual with a sudden transition that happened at age of 78.13 years.

3.2. Bacon–Watts model

An alternative to the broken-stick model is the Bacon–Watts model [9]. For the i th subject at time t_{ij},

$$y_{1ij}=\beta_{1i}+\beta_{2i}\left(t_{ij}-\beta_{4i}\right)+\beta_{3i}\left(t_{ij}-\beta_{4i}\right)\mathrm{trn}\left(\left(t_{ij}-\beta_{4i}\right)/\phi_1\right)+\epsilon_{1ij}, \quad (3)$$

$$y_{2ij}=\beta_{5i}+\beta_{6i}\left(t_{ij}-\beta_{8i}\right)+\beta_{7i}\left(t_{ij}-\beta_{8i}\right)\mathrm{trn}\left(\left(t_{ij}-\beta_{8i}\right)/\phi_2\right)+\epsilon_{2ij}, \quad (4)$$

where trn denotes the general transition function. Here, we choose to use the hyperbolic tangent function, tanh, a commonly used transition function; ϕ_1 and ϕ_2 are the transition parameters in the bivariate model and determine transition rates with larger values corresponding to slower transitions. In particular, if the transition parameter is close to zero, the Bacon–Watts model will work similarly to the broken-stick model. Parameters β_{1i} and β_{5i} denote the intercepts in each model, which have the same interpretation as in the random broken-stick models. Parameters β_{4i} and β_{8i} are change points in the bivariate model. However, the two slopes (β_{2i} and β_{3i}) in model (3) and the two slopes (β_{6i} and β_{7i}) in model (4) no longer have the same interpretation as in the random broken-stick model because of the formulation of the Bacon–Watts model. Again, we assume a multivariate normal distribution for all parameters in the bivariate model,

$$\beta_i=(\beta_{1i},\beta_{2i},\beta_{3i},\beta_{4i},\beta_{5i},\beta_{6i},\beta_{7i},\beta_{8i})^T \sim \mathrm{MVN}\left(\beta,\Sigma_\beta\right),$$

where

$$\beta=(\beta_1,\beta_2,\beta_3,\beta_4,\beta_5,\beta_6,\beta_7,\beta_8)^T$$

is the vector of means, and Σ_β is the 8 × 8 variance–covariance matrix corresponding to the parameter vector.

Compared with the broken-stick model, the Bacon–Watts model enjoys continuity over the entire parameter space. However, its applicability may be limited because its slope parameters are difficult to interpret with respect to practice. Continuing the previous example in Section 3.1, the black dash line in Figure 2 shows the predicted Bacon–Watts curve with a smooth transition at the age of 77.68 years and a transition parameter of 1.60 for the selected subject.

3.3. Smooth polynomial model

Another alternative to the broken-stick model is the smooth polynomial model in which the continuity in the regions around the change points is achieved by using a polynomial function [8]. The bivariate random smooth polynomial model for the i th subject at time t_{ij} is given by

$$y_{1ij}=\left(\eta_{1i}+\eta_{2i}t_{ij}\right)I_{(-\infty,\eta_{4i})}\left(t_{ij}\right)+g_1\left(t_{ij}|\eta_{1i},\eta_{2i},\eta_{3i},\varepsilon_1\right)I_{[\eta_{4i},\eta_{4i}+\varepsilon_1)}\left(t_{ij}\right)+\left(\lambda_{1i}+\eta_{3i}t_{ij}\right)I_{[\eta_{4i}+\varepsilon_1,\infty)}\left(t_{ij}\right)+\epsilon_{1ij} \quad (5)$$

and

$$y_{2ij}=\left(\eta_{5i}+\eta_{6i}t_{ij}\right)I_{(-\infty,\eta_{8i})}\left(t_{ij}\right)+g_2\left(t_{ij}|\eta_{5i},\eta_{6i},\eta_{7i},\varepsilon_2\right)I_{[\eta_{8i},\eta_{8i}+\varepsilon_2)}\left(t_{ij}\right)+\left(\lambda_{2i}+\eta_{7i}t_{ij}\right)I_{[\eta_{8i}+\varepsilon_2,\infty)}\left(t_{ij}\right)+\epsilon_{2ij} \quad (6)$$

where ε_1 and ε_2 denote the intervals around the change points that connect the two linear parts in each model and act as transition parameters as in the Bacon–Watts model but with a slightly different interpretation. As the transition parameter tends to zero, the interval around the change point tends to zero, and the smooth polynomial model becomes the broken-stick model. Note that the parameters in the smooth polynomial models have different interpretations from the previous two models. The change points in the smooth polynomial models are defined as η_{4i} + 1/2ε_1 and η_{8i} + 1/2ε_2, respectively. η_{1i} and η_{5i} are the mean values of longitudinal measurements at η_{4i} and η_{8i} for the i th subject, respectively. Parameters η_{2i} and η_{3i} specify the slopes for the two linear parts before and after the smooth interval, respectively, for y_{1ij}, and η_{6i} and η_{7i} are defined similarly for y_{2ij}.

In model (5), we derive λ_{1i} by assuming the equality of the two linear parts at change point η_{4i} + 1/2ε_1; eventually, we could represent it by a function of (η_{1i}, η_{2i}, η_{3i}, ε_1). We derive λ_{2i} in model (6) by following the same argument. Hence,

$$\lambda_{1i}=\eta_{1i}+\eta_{2i}\left(\eta_{4i}+1/2\varepsilon_1\right)-\eta_{3i}\left(\eta_{4i}+1/2\varepsilon_1\right),$$

$$\lambda_{2i}=\eta_{5i}+\eta_{6i}\left(\eta_{8i}+1/2\varepsilon_2\right)-\eta_{7i}\left(\eta_{8i}+1/2\varepsilon_2\right).$$

g_1 and g_2 are two pre-specified polynomial functions that connect the two linear parts in each model. As in van den Hout *et al.* [8], we achieve the smoothness of transition by imposing special constraints on g_1 so that the polynomial function will connect with the values of the linear function:

$$g_1\left(\eta_{4i}\right)=\eta_{1i}+\eta_{2i}\eta_{4i},\ g_1\left(\eta_{4i}+\varepsilon_1\right)=\eta_{1i}+\eta_{2i}\left(\eta_{4i}+\varepsilon_1\right),$$

$$\left(\frac{\partial}{t_{ij}}g_1\right)\left(\eta_{4i}\right)=\eta_{1i},\ \left(\frac{\partial}{t_{ij}}g_1\right)\left(\eta_{4i}+\varepsilon_1\right)=\eta_{2i}.$$

We define g_1 as a cubic polynomial, $g_1(x) = a_3x^3 + a_2x^2 + a_1x + a_0$, and solve the preceding linear system of four linear ordinal differential equations, where g_1 is a quadratic polynomial with the following coefficients (the coefficient of x^3 is zero):

$$a_2=\frac{\eta_{3i}-\eta_{2i}}{2\varepsilon_1},\ a_1=\eta_{2i}-\frac{\eta_{3i}-\eta_{2i}}{\varepsilon_1}\eta_{4i},\ a_0=\eta_{1i}+\frac{\eta_{3i}-\eta_{2i}}{2\varepsilon_1}\eta_{3i}^2.$$

We can specify the form of g_2 similarly as g_1.

We again assume a multivariate normal distribution for all parameters in models (5) and (6):

$$\eta_i=(\eta_{1i},\eta_{2i},\eta_{3i},\eta_{4i},\eta_{5i},\eta_{6i},\eta_{7i},\eta_{8i})^T\mathrm{MVN}\left(\eta,\Sigma_\eta\right),$$

where

$$\eta=(\eta_1, \eta_2, \eta_3, \eta_4, \eta_5, \eta_6, \eta_7, \eta_8)^T$$

represents the mean vector, and Σ_η is the 8 × 8 variance–covariance matrix corresponding to the parameter vector.

The smooth polynomial model not only maintains the advantages of the previous two models but also overcomes drawbacks of the previous two models. Thus, the smooth polynomial model is superior in interpretable parameters and continuity at the change point. Again, in Figure 2, assuming a fixed interval of 3 years around the change point, we illustrate the predicted smooth polynomial curve (black dot line) for the selected individual. It is observed that the smooth curve started at 80.51 years old and the change point was at 82.01 years old, calculated by adding half of the interval (1.5 years) to 80.51.

3.4. Estimation method

The maximum likelihood method is commonly used for parameter estimation in mixed-effects models. However, its use in models with multiple random effects can be challenging because of the need for multi-fold integrations. The Gaussian quadrature method, a numerical technique for approximating the multi-fold integration in mixed-effects models, can become computationally intractable when the number of random effects is large. In contrast, the Bayesian method using MCMC sampling avoids the direct multi-fold integration by taking repeated samplings from conditional posterior distribution for each parameter in the model, thus providing numerical solutions to a complex modeling situation.

Hall *et al.* [5, 10], Dominicus *et al.* [7], and van den Hout *et al.* [8] have considered the Bayesian method for parameter estimation from univariate random change point models. WinBUGS (MRC Biostatistics Unit, Cambridge, UK)[16] is a powerful and flexible statistical software for Bayesian inference using the Gibbs sampling technique. BRugs [17] is a package in **R** [18] that also uses the Gibbs sampling method for Bayesian inference. BRugs performs similarly as WinBUGS with an additional advantage of combining data manipulation with the Bayesian model's fitting process including model specification and the choice of priors. We chose to implement our methods using BRugs mostly because it can handle the simulations. For application to data analysis, we expect that both WinBUGS and BRugs will be adequate for implementing the bivariate change point model.

We base the quality of fit on two criteria, the deviance information criterion (DIC) [19] and the conditional predictive ordinate (CPO) [20]. We also monitor the trace plot of MCMC iterations for the purpose of convergence checking. The DIC has been widely used for Bayesian model comparison. Dominicus *et al.* [7] used DIC to compare models with different structures as well as models differing in prior distributions. The DIC consists of two parts: DIC= $\bar{D} + p_D$, where $\bar{D}$ is the posterior expectation of deviance, and p_D is the effective number of parameters measuring the complexity of model (defined as the posterior mean of the deviance minus the deviance of the posterior means). Similar to Akaike information criterion [21], a smaller DIC corresponds to a better fit. Another frequently used model-selection criteria in Bayesian inference is the CPO, a cross-validated predictive approach calculating the predictive distributions conditioned on the observed data by leaving out one observation each time. Chen *et al.* showed that there existed a Monte Carlo approximation of the CPO [22]. The models are compared using the log pseudo-marginal likelihood (LPML), which is defined as $\mathrm{LPML}=\sum_{i=1}^{n} \log\left(\widehat{\mathrm{CPO}}_i\right)$, where n is the total number of observations and $\widehat{\mathrm{CPO}}_i$ is the Monte Carlo approximation of CPO. Contrary to the DIC, the model with larger LPML indicates a better fit.

4. Simulation study

We used Monte Carlo (MC) simulations to assess the performance of the Bayesian approach for parameter estimation in the proposed bivariate random smooth polynomial models because the smooth polynomial model is more realistic in practice and more comprehensive than the other two models. We simulated data from a bivariate random smooth polynomial model using the estimated parameters from fitting this model to the real data (IIDS). Specifically, each simulated MC data set consists of bivariate longitudinal data from 238 subjects with seven non-missing bivariate repeated measurements per subject (equally spaced with 3 years between the two adjacent visits). We used baseline ages from IIDS subtracted by 65 years as ages at the first visit for each subject.

We present simulation results for two scenarios by varying the variances of change points and measurement errors:

$$\text{Scenario 1 (large variances)}: \sigma^2_{\eta_4}=64,\ \sigma^2_{\eta_8}=16,\ \sigma^2_{\epsilon_1}=20,\ \sigma^2_{\epsilon_2}=5;$$

$$\text{Scenario 2 (small variances)}: \sigma^2_{\eta_4}=16,\ \sigma^2_{\eta_8}=4,\ \sigma^2_{\epsilon_1}=5,\ \sigma^2_{\epsilon_2}=1.$$

We chose the other parameters to be close to the estimated parameters from IIDS data: η_1 = 70, $\sigma^2_{\eta_2}=15$, $\eta_2 = -0.2$, $\sigma^2_{\eta_3}=0.2$, $\eta_3 = -3$, $\sigma^2_{\eta_3}=2$, $\eta_4 = 15$, $\eta_5 = 28$, $\sigma^2_{\eta_5}=16$, $\eta_6 = 0.2$, $\sigma^2_{\eta_6}=0.2$, $\eta_7 = -0.4$, $\sigma^2_{\eta_7}=0.2$, $\eta_8 = 10$, $r_{\eta 2\eta 3} = 0.2$, $r_{\eta 6\eta 7} = -0.5$, $r_{\eta 4\eta 8} = 0.4$, $\varepsilon_1 = 3$, and ε_2 = 3. Here, $r_{\eta 2\eta 3}$ denoted the correlation between η_2 and η_3; we defined $r_{\eta 6\eta 7}$ and $r_{\eta 4\eta 8}$ similarly. Thus, in the 8 × 8 variance–covariance matrix Σ_η, only $\sigma_{\eta 2\eta 3}$, $\sigma_{\eta 6\eta 7}$, and $\sigma_{\eta 4\eta 8}$ were nonzero, and all the other of-diagonal elements were set to be zeros.

4.1. Estimation using bivariate random smooth polynomial models

In the Bayesian model fitting of the bivariate random smooth polynomial model, we chose prior distributions of parameters for scenario 1 as the following:

$$\eta_1 \tilde{} \mathcal{N}(65,\ 0.01),\ \eta_5 \tilde{} \mathcal{N}(25,\ 0.01),$$

$$\sigma^2_{\eta_1} \tilde{} \text{invGamma}(0.001,\ 0.001),$$

$$\begin{pmatrix} \eta_2 \\ \eta_3 \end{pmatrix} \tilde{} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}\right),$$

$$\Sigma_{\eta_2\eta_3} \tilde{} \text{invWishart}\left(\begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix},\ 2\right),$$

$$\begin{pmatrix} \eta_4 \\ \eta_8 \end{pmatrix} \tilde{} N\left(\begin{pmatrix} 15 \\ 10 \end{pmatrix}, \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}\right),$$

$$\Sigma_{\eta_4\eta_8} \tilde{} \text{invWishart}\left(\begin{pmatrix} 100 & 0 \\ 0 & 100 \end{pmatrix}, 2\right),$$

$(\eta_6, \eta_7)^T$ and $\Sigma_{\eta 6 \eta 7}$ have the same prior distribution as $(\eta_2, \eta_3)^T$ and $\Sigma_{\eta 2 \eta 3}$, respectively; $\sigma^2_{\eta_5}$, $\sigma^2_{\epsilon_1}$, and $\sigma^2_{\epsilon_2}$ also have the same priors as $\sigma^2_{\eta_1}$. In Bayesian analysis, in particular, conjugate prior is a natural and popular choice because of its flexibility and mathematical convenience. We chose invGamma(α, β) as it is commonly used as the conjugate prior to the variance of univariate normal distribution, where α is the shape parameter and β is the scale parameter. On the other hand, invWishart(Σ, k) was a conjugate prior to the variance–covariance matrix of a multivariate normal distribution, where Σ is a positive definite inverse scale matrix and the positive integer k denotes the degree of freedom. We chose priors for scenario 2 the same as in scenario 1 except the variance–covariance of the two change points:

$$\Sigma_{n_4 n_8} \tilde{} \text{invWishart}\left(\begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}, \quad 2\right).$$

We first treated the two transition parameters ε_1 and ε_2 as fixed (equal to the true values) in the model fitting for the two scenarios. For each scenario, we generated and fitted 500 MC samples by the bivariate random smooth polynomial model. For each MC sample, we considered 20, 000 additional iterations following 2000 burn-in iterations. We present the simulation results in Table I. For each scenario, we reported mean, mean squared error, mean standard error, empirical standard error, and coverage probabilities of 95% posterior intervals. The simulation results showed that the Bayesian method generally performed well for fitting the bivariate smooth random polynomial model: estimated parameters had low bias, and coverage probability rates of 95% posterior credible intervals were around the nominal level. It is also observed that the variances of change points and variance of measurement errors influence model fitting. Specifically, smaller variances of change points or variance of measurement errors led to parameter estimates with smaller bias, as well as smaller mean square errors (MSEs). We also conducted a simulation study treating the two transition parameters as unknown parameters and setting uniform prior distributions for them. We found few differences in parameter estimations between the two situations.

4.2. Estimation using broken-stick and Bacon–Watts models

We have been focusing on investigating the performance of the bivariate random smooth polynomial model via simulation studies. However, the random smooth polynomial model is much more complex in model structure than the other two models and consequently more computationally expensive in practice; thus, there is a need to study the performance of the other two simplified bivariate models under the assumption that the true model is the bivariate random smooth polynomial model.

We chose prior distributions for the bivariate random broken-stick model and the bivariate random Bacon–Watts model similarly as that in the bivariate random smooth polynomial

model. We treated the two transition parameters in the bivariate random Bacon–Watts model as unknown parameters with uniform prior distributions,

$$\phi_1 \sim \text{Unif}(0.1,\ 5),$$

$$\phi_2 \sim \text{Unif}(0.1,\ 5).$$

Table II summarizes the simulation results of the three different bivariate models for the two scenarios. As most model parameters were not directly comparable because of different model parameterizations, we compared only the following parameters among the three bivariate models: change points, variances of change points, and correlations between change points. Under the assumption that the true model is a bivariate random smooth polynomial model, simulation results confirmed that the bivariate random smooth polynomial model had the best performance among the three modeling frameworks with smaller bias, smaller MSEs, and better posterior interval coverage. In contrast, the bivariate random broken-stick model and the bivariate random Bacon–Watts model showed larger bias, larger MSEs, and worse posterior interval coverage in parameter estimations than those obtained under the bivariate random smooth polynomial model.

The bivariate random broken-stick model and the bivariate random Bacon–Watts model had similar parameter estimation results. When the variances of random change points were larger, the change points were underestimated by approximately 2 years; estimates of variances of change points and correlation between change points also deviated from the true values. However, when the variances of measurement error and variances of change points were small, the bivariate random broken-stick model and the bivariate random Bacon–Watts model had much improved performance, nearly as well as the bivariate random smooth polynomial model.

4.3. Sensitivity analysis

Estimation of change point of random change point model is usually sensitive to the distributional assumption of the data. To study the sensitivity and robustness of the proposed methods, we replaced the normal distribution in generating the simulated data for random effects and error density by lognormal distributions. Again, for each scenario, we generated 500 Monte Carlo samples, each with 238 subjects and seven non-missing bivariate repeated measurements per subject, from the bivariate random smooth polynomial model with correlated slopes in each univariate model. We selected the true parameters to be the same as those in the previous simulation study (Section 4.1) and then transformed them to the mean and standard deviations of the lognormal distribution to ensure that the generated data have the similar range as when using normal distributions. We then fitted the generated MC samples by the three bivariate change point models assuming normal distributions of the random effect and error terms. We chose the prior distributions for parameters for each model in similar fashion as in Section 4.1. For each MC sample, we considered 20, 000 additional iterations following 2000 burn-in iterations. We present simulation results, including the estimates of change points, variances of change points, and the correlation estimates between change points in Table III. The results show that for both of the small and large scenarios the bivariate random smooth polynomial model has the best performance in smaller MSE and better 95% PI coverage. We also observe that under the assumption of the lognormal distribution and the smooth polynomial model, the bivariate random broken-stick model and the bivariate Bacon–Watts model are sensitive in estimating change points. In addition, the variance of change points and measurement errors have some impact on the

model-fitting results. Specifically, the change point estimates from the small variance scenario are less sensitive than those from the large variance scenario.

5. Application to the Indianapolis–Ibadan Dementia Study data

In this section, we fitted the three proposed models to the IIDS data using the previously described Bayesian method. Specifically, age was centered at 65 years. For each bivariate model, we consider two different models that differ in the variance–covariance structure. Of particular interest is the relationship between the change points of cognitive and BMI measurements, and we assume that cognitive function and BMI are correlated only through their change points.

For bivariate random broken-stick models, we denote by BS_1 the model with the following specified variance–covariance structure:

$$\Sigma_\alpha = \begin{pmatrix} \sigma^2_{\alpha 1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_{\alpha 2} & \sigma_{\alpha 2\alpha 3} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\alpha 2\alpha 3} & \sigma^2_{\alpha 3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_{\alpha 4} & 0 & 0 & 0 & \sigma_{\alpha 4\alpha 8} \\ 0 & 0 & 0 & 0 & \sigma^2_{\alpha 5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma^2_{\alpha 6} & \sigma_{\alpha 6\alpha 7} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{\alpha 6\alpha 7} & \sigma^2_{\alpha 7} & 0 \\ 0 & 0 & 0 & \sigma_{\alpha 4\alpha 8} & 0 & 0 & 0 & \sigma^2_{\alpha 8} \end{pmatrix}$$

. This model allows correlations between the slopes before and after the change points. We further define model BS_2 by setting $\sigma_{\alpha_2\alpha_3} = 0$ and $\sigma_{\alpha_6\alpha_7} = 0$, without correlations between the two slopes. We obtained Bayesian estimation for the preceding two models by imposing prior distributions similar to those in the simulation study. We found that the choice of non-informative prior distributions had little influence on the marginal posterior distributions. Model BS_1 (DIC = 15, 810, LPML = −6, 752) is superior to BS_2 (DIC = 15, 820, LPML = −6, 766) in terms of smaller DIC, larger LPML, and better convergence on the basis of the history trace plots of model parameters. Table IV shows the summary statistics of the posterior distributions of model parameters from model BS_1. The mean (95% posterior interval) age of cognitive function change point is 22.7 (19.5, 25.7) years; the mean slope before change point is −0.1 (−0.2, −0.1) points/year, and the mean slope after change point is −1.7 (−2.6, −1.0) points/year. Cognitive function decreases steadily before the change point and plummets after the change point. On the other hand, for BMI the mean age of change point is 12.5 (9.0, 16.8) years; the mean slope before the change point is 0.2 (0.03, 0.4) points/year, and the mean slope after the change point is −0.4 (−0.6, −0.3) points/year. BMI steadily increases before its change point and decreases after. The posterior means of correlation between the two slopes for cognitive function and BMI are both positive but with wide 95% posterior intervals. The posterior mean of the correlation between the two change points is 0.5 (0.1, 0.8), suggesting that the change in cognitive function is positively correlated with the change in BMI. Furthermore, the estimated change point of BMI is found to be on average 10 years ahead of the estimated change point of cognitive function. The posterior means of the variances of the change points of cognitive function and BMI are 52.3 (30.5, 82.8) and 22.5 (9.0, 43.6) respectively.

For the bivariate random Bacon–Watts model, we denote by BW_1 the model with the same variance–covariance structure as for BS_1, and by BW_2 the model with the same variance–covariance structure as for BS_2 earlier. We choose prior distributions to be similar to the settings for the bivariate random Bacon–Watts model in the simulation study. Different prior

distributions led to similar posterior distributions of model parameters, DIC and LPML. The two slopes in each model are correlated according to the model parameterization of the random Bacon–Watts model. Therefore, BW_1 as a more faithful model should be a better model than BW_2. This is evident from the history trace plots of model parameters, the DIC, and LPML. Specifically, the BW_1 had a smaller DIC 15, 790 versus 15, 870 from model BW_2 and a greater LPML −6, 738 versus −6, 806 from model BW_2. We summarize the posterior distributions of model parameters for model BW_1 in Table III. The mean change points for cognitive function and the BMI are 22.1 (18.8, 25.5) and 11.3 (7.9, 15.3) years, respectively. The posterior mean of transition parameters ϕ_1 and ϕ_2 are 1.7 (0.2, 4.4) and 3.3 (0.4, 4.9), respectively. The posterior mean of correlation between change points is 0.5 (0.1, 0.8), confirming that the change in cognitive function is positively correlated with the change in BMI. Similar to the previous model BS_1, the estimated change point of BMI is around 11 years ahead of the estimated change point of cognitive function.

Again, for the bivariate random smooth polynomial model, model SP_1 enjoys the same variance–covariance structure as for BS_1, and model SP_2 has the same variance–covariance structure as for BS_2. We specified similarly the prior distributions for parameters in the preceding two models as in the simulation study. Different from the bivariate random Bacon–Watts model, the transition parameters in the smooth polynomial model are held constant at 3 years, according to the roughly 3-year intervals between two visits. We assume that the change of cognitive function and BMI occurred within 3 years (we also implemented model fitting with the two transition parameters equal to 1 or 6 years, but parameter estimates are quite similar). Because of the limited number of repeated measurements per subject, we choose not to include those parameters in the model. This also resulted in a convergence problem in the Bayesian computation. Model SP_1 is superior with a smaller DIC of 15, 540 (compared with model SP_2's 15, 720) and a larger LPML of −6, 659 (compared with model SP_2's −6, 742) as well as better convergence profiles, as is evident from the history trace plot. Table IV shows a summary of the posterior distributions of model parameters for model SP_1. The mean age of cognitive function change point is 27.5 (= 26.0 + 1.5) (22.3, 30.8) years with variance 60.8 (33.5, 105.6); the mean slope before smooth interval is −0.2 (−0.2, −0.1) points/year, and the mean slope after smooth interval is −3.0 (−4.8, −1.67) points/year. Cognitive function steadily decreases before the smooth interval and sharply declines after. For BMI measurement, the mean age of BMI change point is 11.0 (= 9.5 + 1.5) (7.3, 12.4) years with variance 14.5 (8.0, 23.9); the mean slope before the smooth interval is 0.2 (0.1, 0.4) points/year, and the mean slope after the smooth interval is −0.4 (−0.5, −0.3) points/year. The BMI has a similar trend—gradually increasing before smooth interval and decreasing after. The posterior mean of correlation between change points is 0.6 (0.2, 0.8), which again implies that change in cognitive function is positively correlated with the change in BMI. Compared with the change point of cognitive function, the change point of BMI appears to be 16 years ahead on average.

The model fitting of the three bivariate models uses the same 10,000 burn-in and 40,000 additional iterations. It takes about 20, 27, and 55 min for the bivariate random broken-stick, the bivariate random Bacon–Watts, and the bivariate random smooth polynomial models, respectively, in a PC with Pentium(R) 4 CPU 2.00 GB of RAM.

In the previous paragraphs, we have presented model-fitting results of the three different bivariate random change point models for cognitive function and BMI. We show the fitted trajectories of nine random selected individuals from IIDS for models BS1, BW1, and SP1 in Figure 3. We compared the three models on the basis of both DIC and LPML; model SP_1 appears to be the best model under consideration of the smallest DIC (15, 540) and the largest LPML (−6, 659). We observed differences in parameter estimation among the three models (Table I). The estimated change points of cognitive function measurements are 22.7,

22.1, and 27.5 years for models BS_1, BW_1, and SP_1, respectively. The estimated change points of models BS_1 and BW_1 are very close, whereas the estimated change point of cognitive function of SP_1 is around 5 years later than those from the other two models. Various reasons may cause such a big change point estimate for cognitive function in the random smooth polynomial model. First, as we have observed in scenario 1 of the simulation study, if the true model is the bivariate random smooth polynomial model with larger variances, the estimated change points tend to be years later than those of the other two models. Second, the special model structure of the smooth polynomial model (with an additional smooth interval between two linear trends) allows the seeking of change points in a later time window. This is observed in the individual trajectory plot as well. On the other hand, estimated change points for BMI in three models are comparable, which are all around 12 years (12.5, 11.3, and 11.0 years for models BS_1, BW, and SP_1, respectively). This may be because the BMI increases firstly and decreases later, making it easy to detect a change point for change point models. Another possibility is that the variance of measurement errors and the variance of change point are both much smaller compared with the cognitive function, so the estimation of change point of BMI becomes more stable for different models. The estimated correlation between the change points of cognitive function and BMI in the three joint models are similar ($r_{\alpha 4 \alpha 8} = 0.5$, $r_{\beta 4 \beta 8} = 0.5$, and $r_{\eta 4 \eta 8} = 0.6$), and all of them have good 95% posterior interval coverage.

6. Conclusion

We have developed joint modeling frameworks of bivariate longitudinal outcomes under three different bivariate random change point models: the bivariate random broken-stick model, the bivariate random Bacon–Watts model, and the bivariate random smooth polynomial model. We applied the proposed methodology to the IIDS data. We used the Bayesian method for model fitting using the BRUGS package in **R**. We assessed the goodness of model fitting using DIC and LPML.

The Bayesian method has been a useful tool for parameter estimation of mixed-effects models with several advantages compared with traditional frequentist methods. First, the highly complex model structure can be still easily handled in WINBUGS and BRUGS. Second, the Bayesian method can deal with the mixed-effect model with multiple random effects. The maximum likelihood method using Gaussian quadrature is commonly used for parameter estimation in non-linear mixed-effect models, but computation of multi-fold integrations can become intractable with large number of random effects. Third, the Bayesian method is also advantageous in its interpretability and ability to deal with missing data [8]. Finally, from a practical implementation perspective, the Bayesian method using an MCMC sampling method is conveniently available in various popular statistical software, such as WINBUGS and BRUGS in **R**. The often cited disadvantage of heavy computation overhead of Bayesian MCMC has become less an issue with the rapid advances of modern computing technology. Bayesian methods require one to specify the prior distributions, which sometimes may be challenging. In many cases, knowledge of priors of parameters is either unknown or even non-existent, which makes it very difficult to specify a unique prior distribution. Careful sensitivity analysis is needed to assess the influence of different priors on the posterior estimates. In our analysis, it appears that the model-fitting results are not sensitive to the choices of priors. On the other hand, the Bayesian method is a useful technique that can incorporate available prior knowledge of the model parameters into the prior distributions.

The restriction of subjects with at least five measurements to be included in our analysis makes the models conditional on the subjects having to survive to a relatively long period of time during follow-up and inevitably limits the modeling to a subset of healthier individuals

than the rest of the cohort. It is known that missing data, especially under the non-ignorable missing data mechanism, could significantly impact model results. In the IIDS data and in most longitudinal studies involving elderly subjects, subject dropouts due to death account for the majority of missing data. Our current proposed method is limited in its capability to deal with non-ignorable missing data and also in the ability to detect potential change point, followed by rapid death. Ghosh *et al*. recently studied the effect of informative dropouts in longitudinal outcomes with multiple change points [23]. It will be an important future research topic to study bivariate change point models that incorporate informative dropouts so that inference on the entire longitudinal cohort can be made. Another interesting and important future research is to take censored change points into account in the bivariate change point model.

The proposed bivariate random change point models not only estimate the change points of bivariate longitudinal outcomes but also investigate the correlation between the change points. This methodology is useful for disease prognosis using biomarkers in medical science, and it also possesses flexibility in model fitting. Although in this paper we have focused on investigating the correlation between the change points, one can readily extend to more complex models by specifying and estimating other correlation parameters such as correlation between the two slopes before the change point as well as the two slopes after the change point in the bivariate model. The extension to multivariate change point models for multiple longitudinal outcomes is also applicable.

Acknowledgments

The research is supported by National Institutes of Health Grants R01 AG019181, R01 AG09956, and P30 AG10133.

References

1. Jack CR, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, Petersen RC, Trojanowski JQ. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. The Lancet Neurology. 2010; 9:119–128.
2. Lange N, Carlin BP, Gelfan AE. Hierarchical Bayes models for the progression of HIV infection using longitudinal cd4 t-cell numbers. Journal of the American Statistical Association. 1992; 87:615–626.
3. Kiuchi AS, Hartigan JA, Holford TR, Rubinstein P, Stevens CE. Change points in the series of t4 counts prior to aids. Biometrics. 1995; 51:236–248. [PubMed: 7766779]
4. Ghosh P, Vaida F. Random change point modelling of HIV immunologic responses. Statistics in Medicine. 2007; 26:2074–2087. [PubMed: 16969894]
5. Hall CB, Ying J, Kuo L, Lipton RB. Bayesian and profile likelihood change point methods for modeling cognitive function over time. Computational Statistics and Data Analysis. 2003; 42:91–109.
6. Jacqmin-Gadda H, Commenges D, Dartigues JF. Random change point model for joint modeling of cognitive decline and dementiae. Biometrics. 2006; 62:254–260. [PubMed: 16542253]
7. Dominicus A, Ripatti S, Pedersen NL, Palmgren J. A random change point model for assessing variability in repeated measures of cognitive function. Statistics in Medicine. 2008; 27:5786–5798. [PubMed: 18680123]
8. van den Hout A, Muniz-Terrera G, Matthews FE. Smooth random change point models. Statistics in Medicine. 2010; 30:599–610. [PubMed: 21337356]
9. Bacon DW, Watts DG. Estimating the transition between two intersecting straight lines. Biometrika. 1971; 58:525–534.
10. 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]

11. Hendrie HC, Osuntokun BO, Hall KS, Ogunniyi AO, Hui SL, Unverzagt FW, Gureje O, Rodenberg CA, Baiyewu O, Musick BS, Adeyinka A, Farlow MR, Oluwole O, Class CA, Komolafe O, Brashear A, Burdine V. Prevalence of Alzheimer's disease and dementia in two communities: Nigerian Africans and African Americans. The American Journal of Psychiatry. 1995; 152:1485–1492. [PubMed: 7573588]
12. Hendrie HC, Ogunniyi A, Hall KS, Baiyewu O, Unverzagt FW, Gureje O, Gao S, Evans RM, Ogunseyinde AO, Adeyinka AO, Musick BS, Hui SL. Incidence of dementia and Alzheimer disease in 2 communities: Yoruba residing in Ibadan, Nigeria, and African Americans residing in Indianapolis, Indiana. The Journal of the American Medical Association. 2001; 285:739–747.
13. Hall KS, Ogunniyi AO, Hendrie HC, Osuntokun BO, Hui SL, Musick BS, Rodenberg CA, Unverzagt FW, Gureje O, Baiyewu O. A cross-cultural community based study of dementias: methods and performance of the survey instrument: Indianapolis, U.S.A. and Ibadan, Nigeria. International Journal of Methods in Psychiatric Research. 1996; 6:129–142.
14. Buchman AS, Wilson RS, Bienias JL, Shah RC, Evans DA, Bennett DA. Change in body mass index and risk of incident Alzheimer disease. Neurology. 2005; 65:892–897. [PubMed: 16186530]
15. Gao S, Nguyen JT, Hendrie HC, Unverzagt FW, Hake A, Smith-Gamble V, Hall K. Accelerated weight loss and incident dementia in an elderly African-American cohort. The American Geriatrics Society. 2011; 59:18–25.
16. Lunn DJ, Thomas A, Best N, Spiegelhalter D. Winbugs—a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing. 2000; 10:325–337.
17. Thomas A, O'Hara B, Ligges U, Sturtz S. Making BUGS Open. R News. 2006; 6:12–17. http://cran.rproject.org/doc/Rnews/.
18. Team RDC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2007.
19. Spiegelhalter DJ, Best NG, Carlin BP, Linde AVD. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society, Series B. 2002; 64:583–639.
20. Gelfand AE, Dey DK, Chang H. Model determination using predictive distributions with implementation via sampling-based methods (with discussion). Bayesian Statistics. 1992; 4:147–167.
21. Akaike H. Factor analysis and AIC. Psychometrika. 1987; 52:317–332.
22. Chen, MH.; Shao, QM.; Ibrahim, JG. Monte Carlo Methods in Bayesian Computation. Springer-Verlag; NewYork: 2000.
23. Ghosh P, Ghosh K, Tiwari RC. Joint modeling of longitudinal data and informative dropout time in the presence of multiple changepoints. Statistic in Medicine. 2010; 30:611–626.

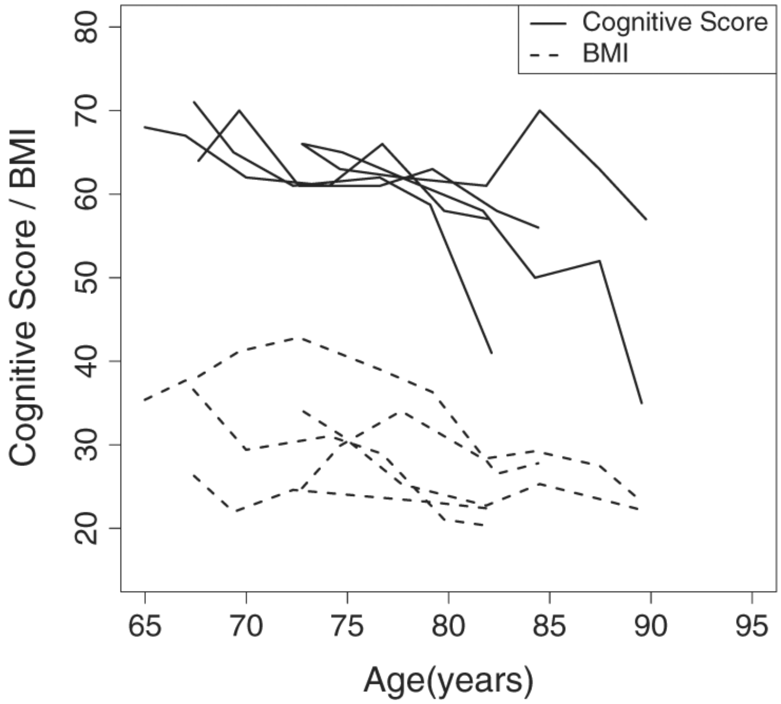

Figure 1.
Cognitive scores and body mass index measures for five randomly selected participants from the Indianapolis–Ibadan Dementia Study.

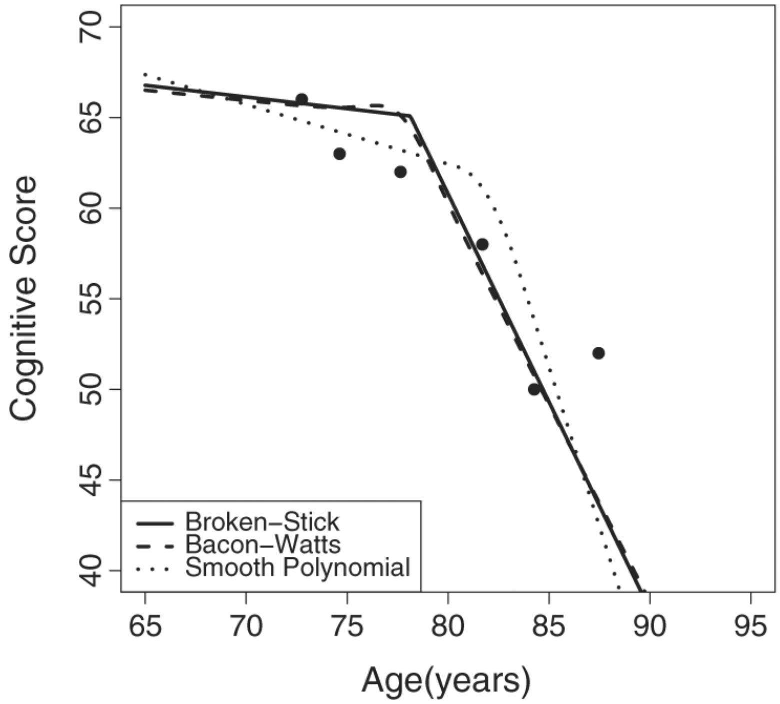

Figure 2.
Predicted curves of the three types of change point model for the cognitive scores of an individual from the Indianapolis–Ibadan Dementia Study.

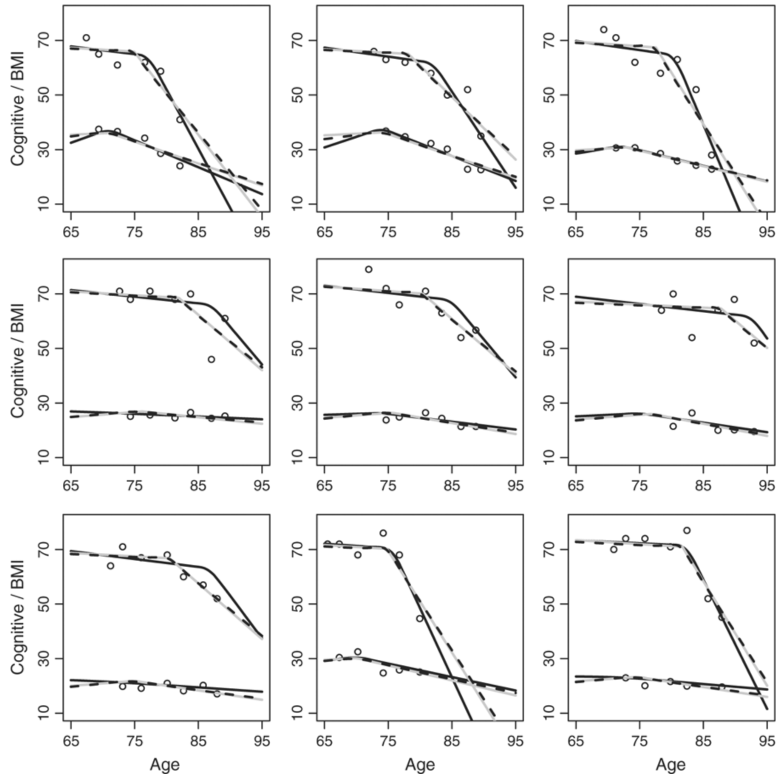

Figure 3.

Plots of nine randomly selected participants from the Indianapolis–Ibadan Dementia Study (black circle), fit for bivariate random broken-stick model BS_1 (solid gray line), bivariate random Bacon–Watts model BW_1 (dashed black line), and bivariate random smooth polynomial model SP_1 (solid black line).

Table I

Simulation results of bivariate random smooth polynomial model.

	Scenario 1 (large variances)						Scenario 2 (small variances)					
	True	Mean	MSE	Mean SE	Empirical SE	95% PI coverage	True	Mean	MSE	Mean SE	Empirical SE	95% PI coverage
η_1	70	70.0	0.21	0.46	0.46	95	70	70.0	0.11	0.32	0.34	95
$\sigma^2_{\eta_1}$	15	15.3	8.55	2.88	2.91	96	15	15.1	3.79	1.92	1.95	94
η_2	−0.2	−0.2	0.002	0.05	0.05	96	−0.2	−0.2	0.001	0.03	0.03	96
$\sigma^2_{\eta_2}$	0.1	0.1	0.001	0.03	0.03	95	0.1	0.1	0.0002	0.01	0.02	95
η_3	−3	−3.0	0.04	0.20	0.20	95	−3	−3.0	0.02	0.13	0.14	95
$\sigma^2_{\eta_3}$	2	1.9	0.15	0.38	0.38	95	2	2.0	0.07	0.27	0.26	95
$r_{\eta_2 \eta_3}$	0.2	0.2	0.07	0.24	0.26	93	0.2	0.2	0.02	0.11	0.12	92
η_4	15	15.1	1.04	0.98	1.02	94	15	15.0	0.13	0.35	0.36	95
$\sigma^2_{\eta_4}$	64	66.1	121.53	11.16	10.83	95	16	16.2	5.47	2.28	2.33	94
$\sigma^2_{\epsilon_1}$	20	20.1	0.83	0.88	0.91	95	4	4.0	0.04	0.19	0.19	96
η_5	28	28.0	0.14	0.38	0.37	96	28	28.0	0.09	0.31	0.31	95
$\sigma^2_{\eta_5}$	16	16.4	5.81	2.31	2.37	93	16	16.2	3.66	1.88	1.90	94
η_6	0.2	0.2	0.003	0.05	0.06	93	0.2	0.2	0.001	0.04	0.04	93
$\sigma^2_{\eta_6}$	0.2	0.2	0.001	0.03	0.03	95	0.2	0.2	0.001	0.02	0.02	94
η_7	−0.4	−0.4	0.002	0.04	0.04	95	−0.4	−0.4	0.001	0.03	0.03	95
$\sigma^2_{\eta_7}$	0.2	0.2	0.001	0.03	0.03	94	0.2	0.2	0.0004	0.02	0.02	97
$r_{\eta_6 \eta_7}$	−0.5	−0.6	0.02	0.12	0.12	90	−0.5	−0.5	0.01	0.07	0.07	94
η_8	10	9.9	0.65	0.83	0.79	95	10	10.0	0.07	0.27	0.26	94
$\sigma^2_{\eta_8}$	16	21.0	47.54	5.26	4.72	87	4	4.2	0.93	0.97	0.94	96

NIH-PA Author Manuscript

NIH-PA Author Manuscript

	Scenario 1 (large variances)						Scenario 2 (small variances)					
	True	**Mean**	**MSE**	**Mean SE**	**Empirical SE**	**95% PI coverage**	**True**	**Mean**	**MSE**	**Mean SE**	**Empirical SE**	**95% PI coverage**
$\sigma^2_{\epsilon_2}$	5	5.0	0.05	0.23	0.23	95	1	1.0	0.002	0.05	0.05	95
$r_{\eta_4 \eta_8}$	0.4	0.3	0.02	0.14	0.13	95	0.4	0.4	0.02	0.13	0.12	96

MSE, mean square error; SE, standard error; 95% PI, 95% posterior interval.

Table II

Simulation results for comparing three bivariate models with data generated from a bivariate random smooth polynomial model.

Broken-stick						Bacon–Watts						Smooth polynomial					
True	**Mean**	**MSE**	**Mean SE**	**Empirical SE**	**95% PI coverage**	**True**	**Mean**	**MSE**	**Mean SE**	**Empirical SE**	**95% PI coverage**	**True**	**Mean**	**MSE**	**Mean SE**	**Empirical SE**	**95% PI coverage**
Scenario 1 (large variances)																	
$\alpha_4(16.5)$	14.3	5.33	0.77	0.78	20	$\beta_4(16.5)$	14.3	5.45	0.77	0.79	20	$\eta_4(15)$	15.1	1.04	0.98	1.02	94
$\sigma^2_{\alpha_4}$ (64)	48.9	282	7.43	7.28	52	$\sigma^2_{\beta_4}$ (64)	48.6	288	7.34	7.19	50	$\sigma^2_{\eta_4}$ (64)	66.1	121	11.16	10.83	95
$\alpha_8(11.5)$	9.5	5.14	0.88	1.05	36	$\beta_8(11.5)$	9.6	4.48	0.86	1.00	41	η_8	9.9	0.65	0.83	0.79	95
$\sigma^2_{\alpha_8}$ (16)	20.5	73.22	6.75	7.31	88	$\sigma^2_{\beta_8}$ (16)	19.7	65.05	6.50	7.15	89	$\sigma^2_{\eta_8}$ (16)	21.0	47.55	5.26	4.72	87
$r_{\alpha_4\alpha_8}(0.4)$	0.3	0.04	0.17	0.17	90	$r_{\beta_4\beta_8}(0.4)$	0.3	0.04	0.17	0.17	92	$r_{\eta_4\eta_8}(0.4)$	0.3	0.02	0.14	0.13	95
Scenario 2 (small variances)																	
$\alpha_4(16.5)$	15.9	0.52	0.34	0.36	56	$\beta_4(16.5)$	15.9	0.50	0.34	0.36	58	$\eta_4(15)$	15.0	0.13	0.35	0.36	95
$\sigma^2_{\alpha_4}$ (16)	15.3	5.08	2.10	2.13	91	$\sigma^2_{\beta_4}$ (16)	15.2	5.09	2.09	2.12	92	$\sigma^2_{\eta_4}$ (16)	16.2	5.47	2.28	2.33	94
$\alpha_8(11.5)$	11.5	0.08	0.26	0.28	94	$\beta_8(11.5)$	11.4	0.08	0.27	0.28	93	$\eta_8(10)$	10.0	0.07	0.27	0.26	94
$\sigma^2_{\alpha_8}$ (4)	3.8	1.12	0.94	1.04	90	$\sigma^2_{\beta_8}$ (4)	3.7	1.22	0.98	1.06	91	$\sigma^2_{\eta_8}$ (4)	4.2	0.93	0.97	0.94	96
$r_{\alpha_4\alpha_8}(0.4)$	0.4	0.02	0.14	0.14	94	$r_{\beta_4\beta_8}(0.4)$	0.4	0.02	0.14	0.15	94	$r_{\eta_4\eta_8}(0.4)$	0.4	0.02	0.13	0.12	96

MSE, mean square error; SE, standard error; 95% PI, 95% posterior interval.

Table III

Simulation results for comparing three bivariate models with data generated from a bivariate random smooth polynomial model using lognormal distribution for all random effects and errors.

Broken-stick						Bacon–Watts						Smooth polynomial					
True	Mean	MSE	Mean SE	Empirical SE	95% PI Coverage	True	Mean	MSE	Mean SE	Empirical SE	95% PI coverage	True	Mean	MSE	Mean SE	Empirical SE	95% PI coverage
Scenario 1 (large variances)																	
$\alpha_4(16.5)$	13.0	13.12	0.51	0.82	12	$\beta_4(16.5)$	13.0	13.23	0.51	0.82	11	$\eta_4(15)$	12.9	6.23	0.66	1.30	94
$\sigma^2_{\alpha_4}$ (64)	20.7	1890	3.37	4.27	35	$\sigma^2_{\beta_4}$ (64)	20.6	1899	3.33	4.19	35	$\sigma^2_{\eta_4}$ (64)	27.0	1443	4.96	8.53	95
$\alpha_8(11.5)$	9.9	6.96	0.75	2.10	0.2	$\beta_8(11.5)$	9.9	6.78	0.75	2.08	0.4	$\eta_8(10)$	7.6	14.07	0.84	2.91	95
$\sigma^2_{\alpha_8}$ (16)	19.2	524	5.56	22.70	67	$\sigma^2_{\beta_8}$ (16)	18.3	429	5.36	20.62	67	$\sigma^2_{\eta_8}$ (16)	16.8	102	4.00	10.11	84
$r_{\alpha_4\alpha_8}(0.4)$	0.3	0.05	0.17	0.16	75	$r_{\beta_4\beta_8}(0.4)$	0.3	0.05	0.17	0.16	75	$r_{\eta_4\eta_8}(0.4)$	0.3	0.04	0.15	0.15	93
Scenario 2 (small variances)																	
$\alpha_4(16.5)$	15.5	1.08	0.29	0.35	2	$\beta_4(16.5)$	15.5	1.08	0.29	0.35	3	$\eta_4(15)$	14.5	0.42	0.30	0.39	94
$\sigma^2_{\alpha_4}$ (16)	11.4	24.36	1.56	1.79	41	$\sigma^2_{\beta_4}$ (16)	11.4	24.65	1.56	1.79	42	$\sigma^2_{\eta_4}$ (16)	12.3	18.16	1.75	2.15	93
$\alpha_8(11.5)$	11.1	2.12	0.32	1.40	7	$\beta_8(11.5)$	11.0	2.20	0.33	1.40	9	$\eta_8(10)$	9.4	0.60	0.31	0.44	95
$\sigma^2_{\alpha_8}$ (4)	5.2	80.51	1.29	8.90	47	$\sigma^2_{\beta_8}$ (4)	4.8	40.72	1.20	6.34	52	$\sigma^2_{\eta_8}$ (4)	4.6	2.00	1.04	1.30	90
$r_{\alpha_4\alpha_8}(0.4)$	0.4	0.03	0.15	0.16	82	$r_{\beta_4\beta_8}(0.4)$	0.4	0.03	0.15	0.17	81	$r_{\eta_4\eta_8}(0.4)$	0.4	0.02	0.14	0.14	96

MSE, mean square error; SE, standard error; 95% PI, 95% posterior interval.

Table IV

Bayesian estimates of population parameters and 95% PI for bivariate random BS_1, bivariate random BW_1, and bivariate random SP_1 from IIDS data.

Parameter	BS_1 Estimate	95% PI	Parameter	BW_1 Estimate	95% PI	Parameter	SP_1 Estimate	95% PI
α_1	68.6	(67.4, 69.7)	β_1	68.5	(67.4, 69.7)	η_1	71.4	(70.6, 72.1)
$\sigma^2_{\alpha_1}$	14.8	(10.8, 19.4)	$\sigma^2_{\beta_1}$	14.7	(10.7, 19.3)	$\sigma^2_{\eta_1}$	15.1	(11.3, 19.5)
α_2	−0.1	(−0.2, −0.1)	β_2	−0.9	(−1.3, −0.5)	η_2	−0.2	(−0.2, −0.1)
$\sigma^2_{\alpha_2}$	0.02	(0.01, 0.03)	$\sigma^2_{\beta_2}$	0.4	(0.1, 1.0)	$\sigma^2_{\eta_2}$	0.04	(0.02, 0.1)
α_3	−1.7	(−2.6, −1.0)	β_3	−0.7	(−1.2, −0.4)	η_3	−3.0	(−4.8, −1.7)
$\sigma^2_{\alpha_3}$	1.8	(0.6, 3.9)	$\sigma^2_{\beta_3}$	0.4	(0.12, 1.0)	$\sigma^2_{\eta_3}$	3.5	(0.9, 8.1)
α_4	22.7	(19.5, 25.7)	β_4	22.1	(18.8, 25.5)	η_4	26.0	(22.3, 30.8)
						$\eta_4 + 1/2\varepsilon_1$	27.5	
$\sigma^2_{\alpha_4}$	52.3	(30.5, 82.8)	$\sigma^2_{\beta_4}$	51.4	(27.8, 83.0)	$\sigma^2_{\eta_4}$	60.8	(33.5, 105.6)
α_5	30.8	(29.9, 31.7)	α_5	30.7	(29.8, 31.6)	η_5	28.6	(27.7, 29.5)
$\sigma^2_{\alpha_5}$	25.9	(21.1, 31.7)	$\sigma^2_{\beta_5}$	26.1	(21.2, 32.0)	$\sigma^2_{\eta_5}$	15.9	(11.9, 20.6)
α_6	0.2	(0.03, 0.4)	β_6	−0.9	(−0.2, 0.1)	η_6	0.2	(0.1, 0.4)
$\sigma^2_{\alpha_6}$	0.03	(0.01, 0.1)	$\sigma^2_{\beta_6}$	0.03	(0.01, 0.05)	$\sigma^2_{\eta_6}$	0.2	(0.1, 0.3)
α_7	−0.4	(−0.6, −0.3)	α_7	−0.3	(−0.4, −0.2)	η_7	−0.4	(−0.5, −0.3)
$\sigma^2_{\alpha_7}$	0.1	(0.04, 0.1)	$\sigma^2_{\beta_7}$	0.04	(0.01, 0.1)	$\sigma^2_{\eta_7}$	0.1	(0.1, 0.2)
α_8	12.5	(9.0, 16.8)	β_8	11.3	(7.9, 15.3)	η_8	9.5	(7.3, 12.4)
						$\eta_8 + 1/2\varepsilon_2$	11.0	
$\sigma^2_{\alpha_8}$	22.5	(9.0, 43.6)	$\sigma^2_{\beta_8}$	18.4	(8.5, 36.1)	$\sigma^2_{\eta_8}$	14.5	(8.0, 23.9)
$r_{\alpha_4\alpha_8}$	0.5	(0.1, 0.8)	$r_{\beta_4\beta_8}$	0.5	(0.1, 0.8)	$r_{\eta_4\eta_8}$	0.6	(0.2, 0.8)
$\sigma^2_{\epsilon_1}$	20.1	(18.4, 21.9)	$\sigma^2_{\epsilon_1}$	20.0	(18.3, 21.9)	$\sigma^2_{\epsilon_1}$	20.0	(18.3, 21.7)
$\sigma^2_{\epsilon_2}$	5.8	(5.2, 6.4)	$\sigma^2_{\epsilon_2}$	5.7	(5.0, 6.4)	$\sigma^2_{\epsilon_2}$	5.0	(4.5, 5.6)
$r_{\alpha_2\alpha_3}$	0.04	(−0.6, 0.6)	$r_{\beta_2\beta_3}$	1.0	(0.9, 1.0)	$r_{\eta_2\eta_3}$	0.1	(−0.7, 0.8)
$r_{\alpha_6\alpha_7}$	0.1	(−0.65, 0.7)	$r_{\beta_6\beta_6}$	0.2	(−0.5, 0.7)	$r_{\eta_6\eta_7}$	−0.9	(−1.0, −0.8)
			Φ_1	1.7	(0.2, 4.4)			
			Φ_2	3.3	(0.4, 4.9)			

95% PI, 95% posterior interval; BS_1, broken-stick model; BW_1, Bacon–Watts model; SP_1, smooth polynomial model; IIDS, Indianapolis–Ibadan Dementia Study.