



Published in final edited form as:

Math Comput Simul. 2009 November ; 80(3): 561–571. doi:10.1016/j.matcom.2009.09.002.

An examination of Bayesian statistical approaches to modeling change in cognitive decline in an Alzheimer's disease population

Al Bartolucci^{a,*}, Sejong Bae^b, Karan Singh^b, and H. Randall Griffith^c

^aDepartment of Biostatistics, Alzheimer's Disease Research Center, University of Alabama at Birmingham, USA

^bDepartment of Biostatistics, Institute for Aging and Alzheimer's Disease Research, University of North Texas Health Science Center, USA

^cDepartment of Neurology, Alzheimer's Disease Research Center, University of Alabama at Birmingham, Birmingham, Alabama, USA

Abstract

The mini mental state examination (MMSE) is a common tool for measuring cognitive decline in Alzheimer's Disease (AD) subjects. Subjects are usually observed for a specified period of time or until death to determine the trajectory of the decline which for the most part appears to be linear. However, it may be noted that the decline may not be modeled by a single linear model over a specified period of time. There may be a point called a change point where the rate or gradient of the decline may change depending on the length of time of observation. A Bayesian approach is used to model the trajectory and determine an appropriate posterior estimate of the change point as well as the predicted model of decline before and after the change point. Estimates of the appropriate parameters as well as their posterior credible regions or regions of interest are established. Coherent prior to posterior analysis using mainly non informative priors for the parameters of interest is provided. This approach is applied to an existing AD database.

Keywords

Alzheimer's Disease; Bayesian; Change Point; Mini Mental State; Trajectory

1. Introduction

Cognitive decline in Alzheimer's disease (AD) is often measured by the mini mental state examination (MMSE). It is generally known that over time the performance on this exam declines in AD subjects. The MMSE or Folstein test is a 30-point questionnaire test that is used to assess cognition. It is commonly used in neurological settings to screen for dementia. In an administration time of about 10 minutes it samples various functions including arithmetic, memory and orientation. It was introduced by Folstein *et al* [9] in 1975 and has been widely

*Corresponding Author: Al Bartolucci, 1665 University Blvd., Birmingham, Alabama 35294-0022. Albartol@uab.edu.

Al Bartolucci, Department of Biostatistics and Alzheimer's Disease Research Center, University of Alabama at Birmingham. Ryals Building, Suite 327, 1665 University Blvd., Birmingham, Alabama 35294-0022. Albartol@uab.edu

Sejong Bae, Department of Biostatistics and Institute for Aging and Alzheimer's Disease Research, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, Texas 76107, USA sbae@hsc.unt.edu

Karan P. Singh, Department of Biostatistics and Institute for Aging and Alzheimer's Disease Research, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, Texas 76107, USA ksingh@hsc.unt.edu

H. Randall Griffith, Department of Neurology and Alzheimer's Disease Research Center, University of Alabama at Birmingham. Sparks Center, Room 650 G, 1720 7th Ave. South, Birmingham, Alabama 35294-0017. rlgriffith@uabmc.edu

used with small modifications. A score over 27 (out of 30) is considered normal. Below this threshold, 20–26 indicates mild dementia; 10–19 moderate dementia, and below 10 severe dementia. There may be slightly different interpretations of what the threshold values represent in terms of severity of dementia. However, it can be said that the MMSE score declines as dementia progresses. The value of the MMSE is often corrected for degree of schooling and age in an analysis of covariance (Crum et al. [5]). Low to very low scores usually correlate well with the presence of dementia, although other mental disorders can also lead to abnormal findings on MMSE testing. The question often asked is “What is the rate of decline of this mental status indicator?” For a discussion of this issue please see Holmes et al. [15]. This rate of decline will, of course, depend on the severity of the disease, demographics, and genetic factors as noted in Holmes et al. [16], Perry et al. [24], and Dentino et al. [7] plus numerous other references in neurology. The research issue in our approach is to examine the rate of decline in a longitudinal cohort of AD subjects and determine through several statistical strategies at what time in the follow up period does this rate of decline change if at all. Alternatively, is there a time point at which the slope of the linear function changes or levels off which may be a real leveling off of cognitive disability or in part due to informative censoring? If so, then what is that time in the follow up period and what is the average MMSE at that time?

We work with a cohort of 271 AD subjects who have been followed for 4 years since their diagnosis of AD. In our data base we have subjects followed to about seven years. However, the sample size beyond 4 years is only 10 percent of the sample which is very limited for valid statistical inference due to the large amount of loss of data. Most subjects were mild to moderate AD with only 6 being severe. Since our analyses adjusted for initial MMSE we left the severe group in the cohort. Severe disease is the point at which memory difficulties continue to worsen, significant personality changes may emerge and affected individuals need extensive help with customary daily activities such as getting dressed. The primary response variable is the change in MMSE over time. We apply a mixed model analysis to determine the variables influencing the course of the MMSE. Once this is done a recursive partition analysis is applied to determine the cut points of these variables indicating a significant influence on the variation of the MMSE. We then model the MMSE over time using classical regression techniques to determine a possible point in time at which the rate of MMSE decline changes. To improve on the fixed effects analysis, our next approach is to incorporate a Monte Carlo Markov Chain (MCMC) Bayesian approach to determine the posterior densities of the regression parameters (slopes of these lines, change point and value of MMSE at the change point). Our approach is much like that of Carlin et. al [3]. This is a completely Bayesian parametric approach using weak priors sometimes referred to as “non informative” on the parameters of interest. Our method differs from previous authors such as Carter and Blight [4] who examine normally distributed sequential responses and use a Bayesian rule for detecting a change from a constant mean response. Here empirical data from an aggregate is used to derive prior information where our approach is primarily non informative. Their procedure is somewhat sequential in that they wish to determine a change point quickly for a medical application. Our approach is to analyze the data retrospectively to make inference about the change point. Lin [19] evaluates an approach for detecting covariate influence on certain change points which may impact on evaluation of reliability models. The process follows a Poisson distribution modeling decline of functionality in two stages. The Markov Chain Monte Carlo technique is used to simulate the posteriors resulting from gamma and multinomial priors. Elliott and Slope [8] use change modeling much like ours to examine data over time to determine the change in rate of vehicle crashes in Michigan before and after introducing the graduated driver’s license program. Their process follows a linear trend and they employ a prior structure much as we do and they utilize the MCMC approach for estimation. Multiple change points are possible in their application. Of course the classical approach to this endeavor of examining change points has also been pursued earlier. For example see Hinkley [13] for a two phase regression application. For our

MMSE application we demonstrate like Tombaugh [28] that the MMSE follows closely a linear decline over time. However, unlike Tombaugh[28], we do not assume a straight linear decline, but show a change point in which the decline may not be one linear trend, but is allowed to change at some point in the follow up. We do not discount that there could be more. However, for the short duration that this diagnostic group is observed, one can get a reasonable estimate of at least one.

The computational resources at hand help us to eliminate the challenge of computing marginal posterior distributions of the change point and the model parameters. The method is conceptually straight forward as we utilize the random effects approach of Ashby et al [1]. Ashby et al have also listed a Web guide to the WinBugs computational approach to this challenge at stat.bus.utk.edu/techrpts/2003/2003-01.pdf. The BUGS language allows a concise expression of the model to denote stochastic (probabilistic) relationships and to denote deterministic(logical) relationships. The stochastic parameters, however specified, may be given proper but minimally informative prior distributions, while the logical expression for the variance in the model allows the standard deviation (of the random effects distribution) to be estimated. Fixed effect model approaches are also handled rather well with the software. As seen in the WinBugs manual by Spiegelhalter et. al [27] also at www.mrc.bsu.cam.ac.uk/bugs/winbugs/manual14.pdf, the WinBUGS software uses compound documents, which comprise various different types of information (formatted text, tables, formulae, plots, graphs, etc.) displayed in a single window and stored in a single file for application to the problem at hand. This manual describes the WinBugs software an interactive Windows version of the BUGS program for Bayesian analysis of complex statistical models using MCMC techniques. We have been careful to apply only our newly established versions of the models that WinBugs handles thus avoiding possibly spurious results with untested models as one is so warned. Some modifications have been made to accommodate the particular parameterization of the change point problem. The approach that is presented in this paper can be viewed as a combination of the models or mixed models over two separate time periods. We suggest a joint model in which, as in Carlin et al [3], the MMSE trajectories or regressions are described by means of early and late time periods of observation in the course of the study. The software and references can be made available by contacting the corresponding author.

2. Methods

Before getting into the Bayesian methodology of examining the MMSE trajectories and change in these trajectories, we provide some motivation for considering this issue.

A cohort of 271 AD patients were followed for at least 4 years. It comprised the diagnoses of 76% mild AD, 21% moderate and 03% severe. These were clinical diagnoses made in a consensus diagnostics conference according to the initial MMSE. The mean age was 73.31 years, sd= 8.81, range 50–94, median=75 years. Patients must have 1) a six month history of progressive short-term memory loss; 2) impairment in occupational or social functioning 3) impairment in two other higher cortical cognitive domains (e.g. language, executive function); 4) a Hachinski score (see McKhann et al.[21]) of less than four; 5) no evidence of a systemic illness or neurological disease that can cause dementia. Age could be a confounder with Time. We examined rate of decline of MMSE at two years for ages 50 to 74 and 75 to 96. The initial average MMSE for ages 50–74 was 22.10(95% confidence interval, 21.21, 22.99) and the initial average MMSE for ages ≥ 75 was 20.51 (95% confidence interval, 19.65, 21.37). The t-test of means yields $p=0.0503$. These values are close to statistical significance. One cannot assume normality of the MMSE score, Thus non parametrically via Wilcoxon's test, the $p=0.01$. However, an analysis of covariance comparing slopes of MMSE over time for the two age groups yielded, $p=0.1109$. We examine the role of age a little more closely in the following

paragraph. Also since most diagnoses were mild to moderate AD, the diagnosis did not impact on the analysis.

One of the purposes of the study was to determine the change in MMSE over time. Time was coded from 0, baseline to 4 years. In addition, over 50 demographic, clinical and psychiatric variables were also collected on these individuals. Among those variables were age, education in years of schooling coded 4 to 24, gender and race (white vs. non white). An initial analysis of the data using a mixed model approach with a spatial (referring to the time between visits at times 1 to 4) Gaussian covariance structure in the SAS 9.1.3 system yielded a strong association of change in MMSE with education ($p=0.0008$) with coefficient 0.278 with 95% confidence interval (0.117, 0.438) and of course, time ($p=0.0001$) with coefficient -1.522 with 95% confidence interval ($-1.866, -1.179$). This covariance structure accommodated the within subjects correlation over time and was the most appropriate structure for this model as the correlation decreased with increasing lag or space from time 0 as one would expect. That is to say, the correlation parameters of MMSE at time 0 with MMSE at times 1,2,3, and 4 were 0.75, 0.56, 0.55 and 0.66, respectively. All were decreasing as one can see with the exception of time 4, which may be due to the large variation from the small sample size at time 4. Time was strongly associated with the subject's age over the course of the study. Thus, although age was not included in this particular model, it could certainly be a possible confounder. We redid the analysis with age as a confounder and our results were such that education now had a p -value of 0.005 with coefficient 0.228 (0.068, 0.387). The parentheses, (), from here on out will indicate the 95% confidence interval unless noted otherwise. Time was still significant at $p=0.0001$ with coefficient -1.3639 ($-1.709, -1.019$). Thus education and time are both still significant. Their coefficients changed slightly with and without age in the model. However, although the coefficient on education changed by nearly 22%, from 0.278 to 0.228, the impact of education may have decreased slightly on the MMSE with age in the model, but not meaningfully as to change its interpretation or influence on the MMSE. For a discussion of age and time confounding effects in repeated measures analyses see Raudenbush, Johnson and Sampson [26]. Education has been repeatedly shown to have strong association with the MMSE as is seen recently in Raji et al [25]. From this analysis the investigators were interested in determining the partitions of the variables, education and time, which appear to give a division of the MMSE rated high or low. Using the recursive partitioning platform of the JMP version 8.0 software, the significant split in education was less than 9 years versus greater than or equal to 9 years of schooling. The former yielded a mean MMSE of 18.26 (16.74,19.77) and the latter yielded a mean MMSE of 21.57 (21.11,22.03). With respect to time, the significant splits were less than 2 years on study with mean MMSE= 22.07 (21.58,22.56) versus 2 years or more on study with mean MMSE=20.08 (19.31,20.85). This obviously was not a meaningful split in terms of differentiating dementia categories as it occurred in the education group of greater than 9 years and indicated mainly mild dementia on average as one would expect. In the less than 9 years education group the split of less than 2 years on study had an average MMSE of 19.45 (18.74,20.16) and greater than two years on study an average MMSE of 15.20 (11.68, 18.71) or both within the moderate or worse dementia group. All this time partitioning within the education groups just confirms that the longer one is observed there is a decline in the MMSE regardless of education status. As a matter of fact, if education were not in the model, the significant split for time is still at 2 years with mean MMSE =21.79 (21.54,22.03) for those on study less than two years and 20.06 (19.64,20.47) for those greater than two years. Note here again that the upper mean is in the mild dementia category and the lower mean is close to the threshold of moderate. A more analytic rationale for the partition of two years is that when regressing MMSE on time the slope of decline of MMSE prior to two years which was -1.24 ($-1.880, -0.599$) was statistically significant, $p=0.001$ and the slope of decline post 2 years which was -0.38 ($-2.08, 1.32$) was not significant, $p=0.629$.

Thus, what was noticeable when one plots the trajectories of MMSE from time 0 to time 2 the slope is steeper than the slope from time 2 until time 4 which indicates that at a certain point in time there is a slowing of the decline of the trajectory. The task now is to ascertain a more accurate assessment of the change point in time and the value of the MMSE at that change?

Suppose we divide our sample into two components. The first are all data at baseline and before time 2; call this time 0, and the second is all data from time 2 forward; call this time 1. The coding of time as 0 or 1 is inconsequential. They are merely place holders as in a repeated measures design. We are actually considering a repeated measures format and taking just two time points instead of 4 To demonstrate if there is a significant downward trend from time 0 to time 1 and that we actually have two different distributional formats of MMSE at times 0 and 1 we will examine a random effects growth curve model. This somewhat resembles the same situation as the mean shift model discussed by Hinckley and Schechtman [14]. However, in our case the shift is in the value of the slope of the curve from time 0 to time 1.

The model for this aspect of the problem is primarily a random effects linear growth curve much like the normal hierarchical model proposed by Gelfand and Smith [11]. Without loss of generality we will consider the quantiles of the MMSE at times 0 and time 1. The quantiles naturally range from 0.0% to 100% and include the inter quartile range of 25%, 50% and 75% with a total of eleven quantile values.

The model is:

$$\begin{aligned}
 Y_{ij} &\sim \text{Normal}(\alpha_i + \beta_i(x_j - \text{mean}[x]), \tau_c) \\
 \alpha_i &\sim \text{Normal}(\alpha_c, \tau_\alpha) \\
 \beta_i &\sim \text{Normal}(\beta_c, \tau_\beta)
 \end{aligned}
 \tag{1}$$

where Y_{ij} is the MMSE for the i observations ($i=1, \dots, 271$) at quantile j , $j=1, \dots, 11$; j is the j th quantile used as a default in the JMP 8.0 software. The value x is merely the placeholder time at 0 or 1 where 0 is less than 2 years and 1 is the index for greater than or equal to 2 years. The parameters, $\alpha_c, \tau_\alpha, \beta_c, \tau_\beta, \tau_c$, are given independent “noninformative” or weak priors by the software. For example, $\tau_c \sim \text{gamma}(0.001, 0.001)$. Also see equation (3). Note we have a growth curve for observations at quantile j , as well as an overall composite growth for all the observations across times 0 and 1 denoted by the subscript, c . The τ ‘s represent the variances for each of the distributions denoted above. This is a very simple model to demonstrate the behavior of the MMSE from the two time partitions.

One focuses on the values of the β ’s and naturally if their posterior means are negative then we have a downward trend of the MMSE across the two times and for each observation i , and if the posterior mean of the β_c is negative then the composite trend from times 0 to 1 is certainly decreasing. Although their posterior means are of interest, their posterior variances will be of interest as will be seen in the next “Results” section.

Now concerning an accurate estimate of the change point as well as the MMSE at the change point, our approach is much like that of Carlin et. al [3]. We assume a model with two straight lines that meet at a certain change point, x_k . Much in the same notation of Carlin et al. [3], we assume

$$\begin{aligned}
 Y_i &\sim \text{Normal}(\mu_i, \tau) \\
 \mu_i &= \alpha + \beta_{j_i}(x_i - x_k) \quad J_i=1 \text{ if } i \leq k \quad J_i=2 \text{ if } i > k,
 \end{aligned}
 \tag{2}$$

$i=0,1,2,3,4$ denoting our four time points of observing the MMSE, Y_i . We assume the MMSE follow a normal distribution in our population. A test of normality using the normal quantile plot and by the Shapiro-Wilks test confirmed this assumption. Following the usual regression notation, the μ_i is the mean response at time i .

Here $E(Y) = \alpha$ at the change point, with gradient or slope β_1 before, and gradient β_2 after the change point. We give independent "weak" priors to α , β_1 , β_2 and τ . A uniform prior is given to x_k . Here we use the actual data to calculate the change point and two separate curves, one prior to the change point and one after the change point. For a discussion of multiple change points and curves, please see Elliot and Slope [8]. The overall goal is to derive accurate posterior estimates of the change point as well as posterior estimates of the linear parameters. One could attempt a more complicated non linear mixture model for the curves. However, that was not necessary in this case as will be seen in the next section.

3. Results

Following equation (1) above, we wish to calculate the parameters for our growth curve model to examine the trend across the two major splits time 0 for observations less than two years and time 1 for observations greater than or equal to 2 years. Following the convention of our MCMC computations we define

σ as the inverse square root of τ . i.e.

$$\sigma = \sqrt{\tau}.$$

The weak priors are thus defined:

$$\begin{aligned} \tau_c &\sim \text{gamma}(0.001, 0.001) \\ \sigma &= 1/\sqrt{\tau_c} \\ \alpha_c &\sim \text{norm}(0.0, 1.0E-6) \\ \tau_\alpha &\sim \text{gamma}(0.001, 0.001) \\ \beta_c &\sim \text{norm}(0.0, 1.0E-6) \\ \tau_\beta &\sim \text{gamma}(0.001, 0.001), \end{aligned} \tag{3}$$

where norm and gamma are naturally the normal and gamma distributional assumptions. Also note the non informative nature of the priors due to the highly dispersed variances in each of the distributions. The application is not limited to the use of non informative hyperparametrization. If one so desires, a conjugate structure can be given to the prior parameters if one was so inclined by merely inserting more precise values for the mean and variances of the normal densities as well as more specific shape or scale hyperparameter values for the gamma components. An empirical Bayes approach for deriving coherent prior to posterior inferences can be used to determine specific values of the hyperparameters based on elicitation or other methods as seen in Birch and Bartolucci [2].

For each individual growth curve based on the quantile values the β_i 's ranged from -0.724 to -2.15 demonstrating a downward trend in all observations from the earlier to the later time period which would be expected with decreasing MMSE over time. Table 1 shows the posterior mean values of the parameters from the composite growth curve model.

One can see from the table that the posterior mean slope value of -1.269 indicates the downward trend across the time periods. The 95% credible Bayesian region from -2.003 to -0.4953 indicates strong evidence of no positive trend. The constant term value of 20.02 is the intercept as one might expect this to be close to the value of the MMSE at the boundary between the two time periods. The posterior standard deviation, sigma or τ of Y_{ij} of equation (1), 0.6158 ,

is not very informative. However, if one examines the plot of this posterior density as seen in Figure 1, one notes a bimodal trend in this value.

This is explained as one examines two of the posterior density plots for the growth curve beta's at the 25th and 75th percentiles in Figure 2.

Note the spike early on and the smaller spike later. This is an indication that the MMSE values at the early time point are not as dispersed as those in the later time points. It will be seen later that this indicates a steeper consistent fall of MMSE values early and less of a slope and more dispersion in the curve at later time periods. Also note at the top of Figure 1 and Figure 2, the expression "...sample 20000". This refers to the number of iterations generated for a solution to the posterior parameter estimate. We discuss this further in the "Conclusion" section 4.

We now focus on equation (2) above to derive the change point value as well as the parameters for the possibly two different lines before and after the change point. Following the same convention as above for our MCMC computations we define

- σ as the inverse square root of τ . i.e.
- $\sigma = 1/\sqrt{\tau}$.

The non informative priors are thus defined:

$$\begin{aligned} \tau &\sim \text{gamma}(0.001, 0.001) \\ \sigma &= 1/\sqrt{\tau} \\ \alpha &\sim \text{norm}(0.0, 1.0E-6) \\ \beta_1 &\sim \text{norm}(0.0, 1.0E-6) \\ \beta_2 &\sim \text{normal}(0.0, 1.0E-6) \\ x_k &\sim \text{unif}(1.5, 3.0), \end{aligned} \quad (4)$$

where in this case, for the change point distribution, unif is the uniform distribution. We assume a liberal range of the uniform prior from 1.5 to 3.0. Also as above, we have the non informative hyperparameter assignments, but can easily change those values to any conjugate beliefs we may have without loss of precision. This issue as well as that of robustness of our procedure is addressed in the conclusion/discussion section.

In Table 2 we have the posterior mean values with appropriate statistics for our parameters:

Note that the solution for the change point yields a value of 2.402 years with 95% credibility region of interest to be from 1.56 years to 2.49 years or the region in which the change in slopes of the MMSE curve could occur. The posterior expected value of the MMSE at this divide is 20.05 with 95% credibility region to be from 18.94 to 21.23. Note also in Table 1 that the initial slope posterior estimate for β_1 is -0.7797 with a posterior 95% credible region not containing 0 indicating a significant decrease in MMSE early on.. However, the slope posterior estimate for β_2 is -0.0632 with a posterior 95% credible region containing 0 indicating perhaps little or no change in the MMSE after the change point divide of 2.402 years. Solving for each of these equations with the original data, the estimated predicted lines for these regions pre and post change point, respectively, as in a covariance model are:

$$\begin{aligned} \text{MMSE} &= 21.92 - 0.78 * (\text{Time} - 2.402) \\ \text{and} \\ \text{MMSE} &= 20.022 - 0.063 * (\text{Time} - 2.402). \end{aligned} \quad (5)$$

The lines are drawn in Figure 3.

We mentioned above that there may be a difference in the variance within the two samples pre and post change point. Note in Table 1 that the 95% posterior region for β_1 is more narrow compared to the region for β_2 . It is also obvious from the standard deviation of both that the first gradient is less dispersed than the second. This is also seen in Figure 4 of the posterior density for the two gradient parameters.

4. Conclusion

Thus given the problem at hand and the appropriate background data our attempt was to determine the time point at which the MMSE growth curve changes rate of decline. This was achieved in a reasonable manner given the computational resources at hand. The MCMC approach of ours is much like that of other authors such as Gelman et al [12]. This is a completely Bayesian parametric approach using non informative priors on the parameters of interest. However, the use of conjugate priors can also be accommodated by our method. The computational procedures help us to eliminate the challenge of computing marginal posterior distributions of the change point and the model parameters. The method utilizes the WinBugs software with modification for our problem. As is seen, we were able to derive all relevant posterior estimates with no problems of convergence. All of our procedures used no more than 20000 update iterations. This is due primarily to the well behaved functions and distributions used for our analyses in terms of computational convergence to an appropriate solution. Most underlying probability distributions with the exception of the uniform distribution were from the exponential family of distributions such as the normal and gamma which have a track record for computational friendliness. The issue of robustness is certainly a reasonable one to consider. Although not detailed here, we utilized initial values for all our procedures that appeared reasonable from past experience or from some initial pilot data. The use of initial values constitutes a chain in the software application. We used different chains for each of our procedures with reasonably alternate initial values to test the consistency of our results. Another way of assessing robustness is to try different hyperparameter values to determine if the form of the prior has undue influence on the overall result. The hyperparameters for our data were realistic. We thus want to be sure that if we vary these values, they are done so reasonably within the limits of the values one would expect for this type of population. The normal means could vary by 0.05 in either direction and the variances or τ could also change in either direction (plus or minus) by about 0.01 which may be reasonable for this data. Working within these ranges, which were appropriate, our results for the change point and other parameters were consistent. Thus in general our problem was fairly robust.

5. Discussion

Many clinical studies have evaluated the course of MMSE over time. These studies generally involve a collection of repeatedly measured data and observations on demographic, psychiatric and clinical data as well. Since the MMSE is usually influenced by some of these variables, repeated measurements of the MMSE are usually closely monitored with these variables. As mentioned earlier our data base involved all these data and we have considered the historical influence of education on the regression of the MMSE over time. Having done so we showed that the MMSE can pretty much stand on its own when considering the point in time in which the rate of decline may change. Thus the problem was focused on that goal.

This model we employed with linear random effects describes the MMSE data very well. Our sensitivity analysis in terms of changing the initial chain and considering different prior structures appear to work well. Other authors such as Garre et al [10], who also took the WinBUGS approach to their analysis, were able to test their model consistency with another data set. We did not have access to another data set and thus we are open to applying our procedure to perhaps an enhanced data set as we move beyond and add to the 271 subjects that we now

have. Also only 3 percent of the sample was severely impaired which really did not impact on the results of declining cognition, mainly due to the fact that these individuals were already at floor on this measure at baseline.

Also we want to note that our posterior models are based on the assumption of normality and one can see from Figure 4 that it certainly looks so. We used the mixed model procedures regressing the MMSE on time adjusting for the covariates of age, sex and education and our residuals were constant and followed a normal distribution, $p=0.3352$, showing the adequacy of the modeling. We extended this assumption to the Bayesian modeling since one justification may be that the Bayesian Central Limit Theorem is a theoretical impetus for posteriors having a normal distribution even if the underlying data do not. Also see Kuskowski [18] in which MMSE scores were regressed on time of examination (measured at 6-month intervals) to estimate cognitive progression rates in individual patients. Thus this linear approach seemed to be a reasonable way to demonstrate our change point strategy to the reader. However, the original MMSE scores themselves at time=0 or visit 0 are not normally distributed. They are actually left skewed which could possibly impact on the slope estimate prior to the change point causing a slight reduction in the slope. Further we note that the MMSE values in this study may have approximated a normal distribution, but that will not necessarily be true in a healthy normal population and that the methods described here may not be applicable.

Also, a note of caution. The censoring for this type of sample is quite extensive over time. From years one to 4 the remaining sample was 61%, 37%, 20% and 10%, respectively. However, considering informative censoring at best and subjecting our data to the method of Dufouil et al. [6] who considered inverse probability weighting and imputation for missing MMSE values, our results were robust with the change point remaining close to the value of 2.4 as seen above. The data as collected was recorded by patient visit which we call time. Since time is inherently an interval scale variable, time scales are usually treated as continuous. This could potentially introduce a bias in both the point and interval estimates of the change point because of the sparsity of data between the times at which observations may actually be observed. Dufouil et al [6] refer to these as waves. Also, time between visits can be considered to be random and treated as such when the actual dates are used or gathered in the database. This certainly enhances the analysis. However, examining the trends noted in Dufouil et al [6], we are at least very close to the pattern of MMSE and monotonic loss that they note over time. Nevertheless, the censoring issue was worth investigating further. Like Dufouil [6], we had 4 waves and thus the probability of observing a particular subject in each wave is 0.25. One is then given a weight of 4 which is equivalent to imputing 3 missing values. Our software allows such weighting. Although the authors are not really comfortable with imputing data, as a check for consistency, the procedure was followed and the change point estimate with this procedure remained close to 2.4.

One could also argue, and rightfully so, that the composite measure of the MMSE and its trend may not represent the actual trends of all the sub domains of that instrument. Also in our database was coded the sub domains of orientation to place and time. When we re-examined our procedure as described above on these sub domains the trends were the same, but not so pronounced. The reason for this is that the sub domains do not have a range of 0 to 30 as the whole scale. We are confined to a very limited range, 0 to 5, in each of these cases which adds to the challenge of duplicating some results. Thus it is not surprising that such an instrument is certainly less sensitive to small changes in cognitive function. We also tried our procedure with the DRS or dementia rating scale (Hughes[17]) which positively correlated at 0.81 with the MMSE. The DRS had a range of 14 to 141 in our sample. We had a break or change also at about 2 years with that result. So our procedure appears to be consistent.

We note that a particular feature of our data application that we presented here is that over 93% of the subjects experienced a decrease in the MMSE whereas the MMSE trajectories remain stable for the rest of the patients. Again this could be due to the fact that some were at the floor measurement at baseline as mentioned above. We thus did not exclude any subjects not experiencing a decrease in the trajectory. Thus we may have been a little conservative, but such is unlikely and when we did exclude them, the results were consistent. Pauler and Finkelstein [22] took a slightly different approach and noted the presence of two subsets within their prostate cancer patient data—those who experienced a change point and those who did not, but they did not take this into account explicitly in their model. Instead their approach was to assume that reasonable estimates of the changepoint would only be reached if the data clearly indicate the existence of a changepoint. However, in the model that was proposed by Pauler and Laird [23] to identify subjects who switch regimes during a clinical trial, the presence of patient subsets was explicitly distinguished. Such was also the case in the latent class joint model that was proposed by Lin et al. [20] to describe heterogeneity in prostate-specific antigen trajectories in subpopulations of prostate cancer patients. Thus we see that the change point problem can be addressed in many different ways. We have provided one approach which we believe will be a starting point for future looks at the trajectories of other cognitive measures as well as other clinical variables of interest measured with time dependency in our AD dataset.

Acknowledgments

Part of this work was supported from research grants 5P50AG16582-10 and 1P03 AG022550-01 from the National Institute of Health (NIH/NIA).

References

1. Ashby, AJ.; Leon, RV.; Thyagarajan, J. Technical Report. Knoxville: Department of Statistics. University of Tennessee; 2003. Bayesian modeling of accelerated life tests with random effects.
2. Birch R, Bartolucci AA. Determination of the Hyperparameters of a Prior Probability Model in Survival Analysis. *Computer Programs in Biomedicine* 1983;17:89–94. [PubMed: 6362981]
3. Carlin BP, Gelfand AE, Smith AFM. Hierarchical Bayesian analysis of changepoint problems. *Applied Statistics* 1992;41(no 2):389–405.
4. Carter RL, Blight BJN. A Bayesian change- point problem with application to the prediction and detection of ovulation in women. *Biometrics* 1981;47(4):743–751. [PubMed: 7337793]
5. Crum RM, Anthony JC, Bassett SS, Folstein MF. Population based norms for the mini-mental state examination by age and educational level. *JAMA* 1993;269(21):2386–2391. [PubMed: 8479064]
6. Dufouil C, Brayne C, Clayton D. Analysis of longitudinal studies with death and drop out: a case study. *Statistics in Medicine* 2004;23(14):2215–2226. [PubMed: 15236426]
7. Dentino AN, Pieper CF, Rao MK, Currie MS, Harris T, Blazer DG, Cohen HJ. Association of interleukin 6 and other biologic variables with depression in older in older people living in the community. *J. Am. Geriatric. Soc* 1999;47:6–11.
8. Elliott MR, Slope JT. Use of a Bayesian changepoint model to estimate effects of a graduated driver's licensing program. *Journal of Data Science* 2003;1:43–63.
9. Folstein MF, Folstein SE, McHugh PR. Mini-Mental Status. A practical method for grading the cognitive status of patients for the clinician. *Journal of psychiatric research* 1975;12(3):189–198. [PubMed: 1202204]
10. Garre FG, Zwinderman AH, Geskus RB, Sijpkens YWJ. A joint latent class changepoint model to improve the prediction of time to graft failure. *J.R. Statist. Soc. A* 2008;171:299–308.
11. Gelfand AE, Smith AFM. Sampling based approaches to calculating marginal densities. *Journ. of Amer. Statist. Assoc* 1990;85:398–409.
12. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. Chapman and Hall/CRC. 2004
13. Hinkley DV. Inference about the intersection in two phase- regression. *Biometrika* 1969;56:495–504.

14. Hinkley DV, Schechtman E. Conditional bootstrap methods in the mean shift model. *Biometrika* 1987;74:85–93.
15. Holmes C, Ballard C, Lehmann D, Smith AD, Beaumont H, Day IN, Khan MN, Lovestone S, McCulley M, Morris CM, Munoz DG, Obrien K, Russ C, Ser TD, Warden D. Rate of progression of cognitive decline in Alzheimer's disease: effect of butyrylcholinesterase K gene variation. *J. Neurol. Neurosurg. Psychiatry* 2005;76:640–643. [PubMed: 15834019]
16. Holmes C, Cairns N, Lantos P, Mann A. The validity of current research diagnostic criteria for Alzheimer's disease vascular dementia and dementia with Lewy bodies. *Br. J. Psychiatry* 1999;174:45–50. [PubMed: 10211150]
17. Hughes CR, Berg L, Danziger WL, Cohen LA, Martin RL. A new clinical scale for the staging of dementia. *Brit. J. Psychiat* 1982;140:566–572.
18. Kuskowski MA, Mortimer JA, Morley GK, Malone SM, Okaya AJ. Rate of cognitive decline Alzheimer's disease is associated with EEG alpha power. *Biol. Psychiatry* 1993;33(8–9):650–652.
19. Lin J. A two stage failure model for Bayesian change point analysis. *IEEE Transactions on Reliability* 57(2):388–393.
20. Lin H, Turnbull BW, McCulloch CE, Slate EH. Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer. *Journ of. Amer. Statist. Assoc* 2002;97:53–65.
21. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan E. Clinical diagnosis for Alzheimer's disease; Report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services Task force on Alzheimer's disease. *Neurology* 1984;34:939–944. [PubMed: 6610841]
22. Pauler DK, Finkelstein DM. Predicting time to prostate cancer recurrence based on joint models for non-linear longitudinal biomarkers and event time outcomes. *Statist. Med* 2002;21:3897–3911.
23. Pauler DK, Laird NM. A mixture model for longitudinal data with application to assessment of noncompliance. *Biometrics* 2000;56:464–472. [PubMed: 10877305]
24. Perry E, McKeith I, Ballard C. The K variant of the butyrylcholinesterase enzyme is associated with significantly slower cognitive decline and predicts treatment response in dementia patients. *Neurology* 2003;10:1852–1853. [PubMed: 12796550]
25. Raji MA, Reyes-Ortiz CA, Kuo YF, Markides KS, Ottenbacher KJ. Depressive symptoms and cognitive change in older Mexican Americans. *J. Geriatr. Psychiatry Neurol* 2007;20(145):145–152. [PubMed: 17712097]
26. Raudenbusch SW, Johnson C, Simpson RJ. A multivariate, multilevel Rausch model for self reported criminal behavior. *Sociological methodology* 2003;33(1):169–211.
27. Spiegelhalter, D.; Thomas, A.; Best, N.; Lunn, D. WinBUGS Version 1.4 User Manual. Cambridge: 2003.
28. Tombaugh TN. Test-retest reliable coefficients and 5 year change scores for the MMSE and 3MS. *Archives of Clinical Neuropsychology* 2005;20(4):485–503. [PubMed: 15896562]

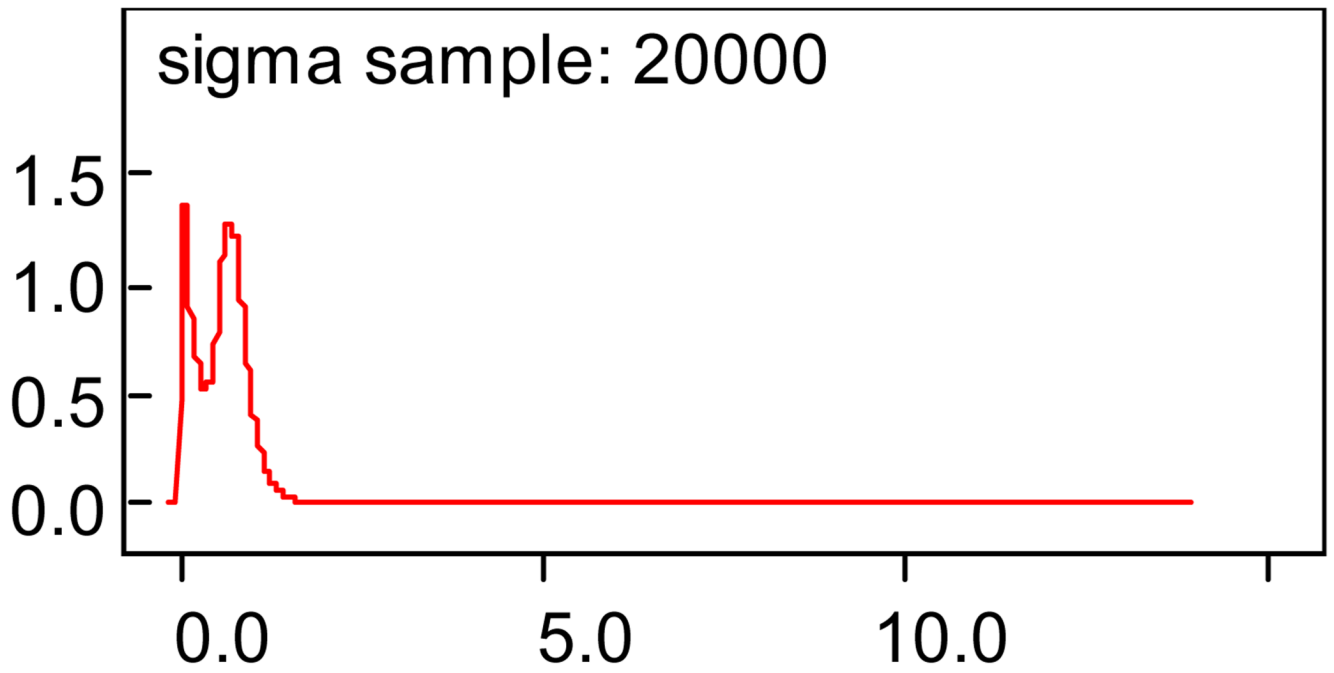


Figure 1.
Plot of posterior variance of the normal hierarchical model, equation (1).

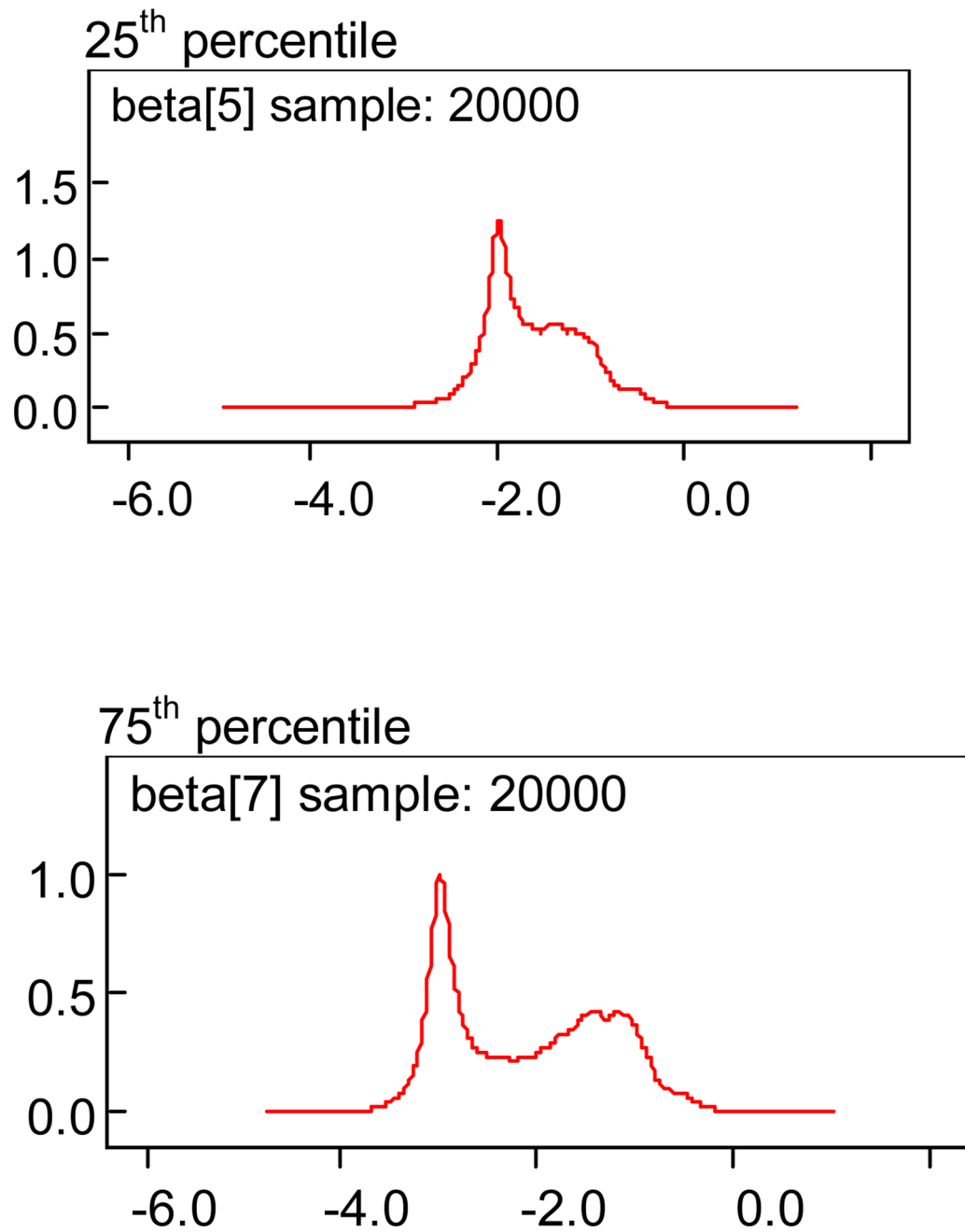


Figure 2. Plot of the posterior variances of the beta estimates at the 25th and 75th percentiles.

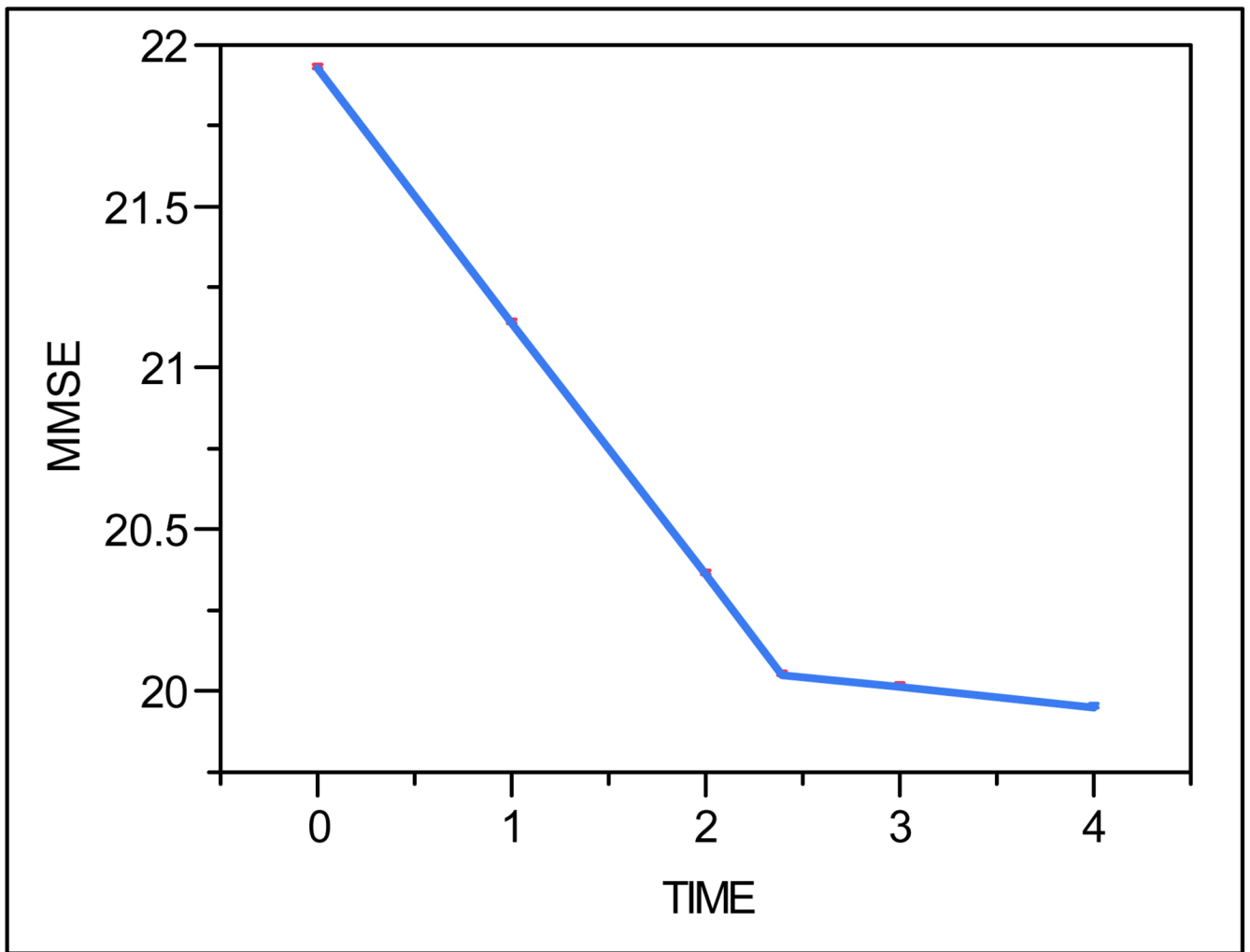
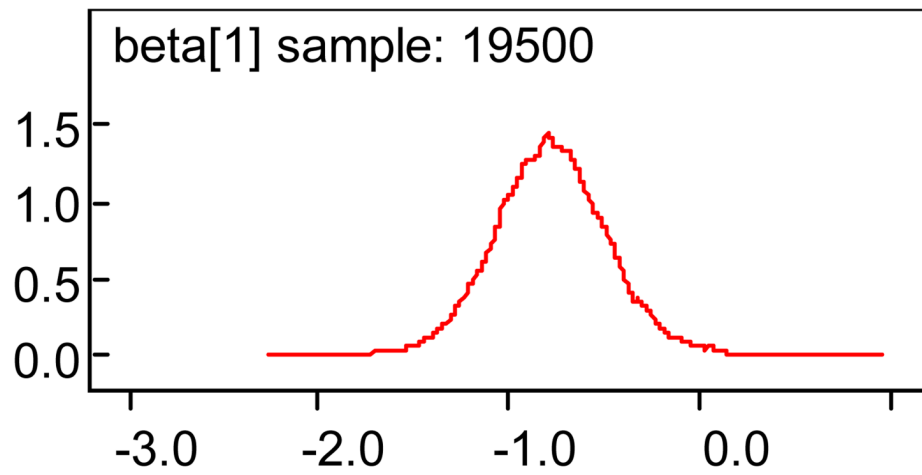


Figure 3.
MMSE by Time. Change point =2.4 years.



Posterior Density of β_1

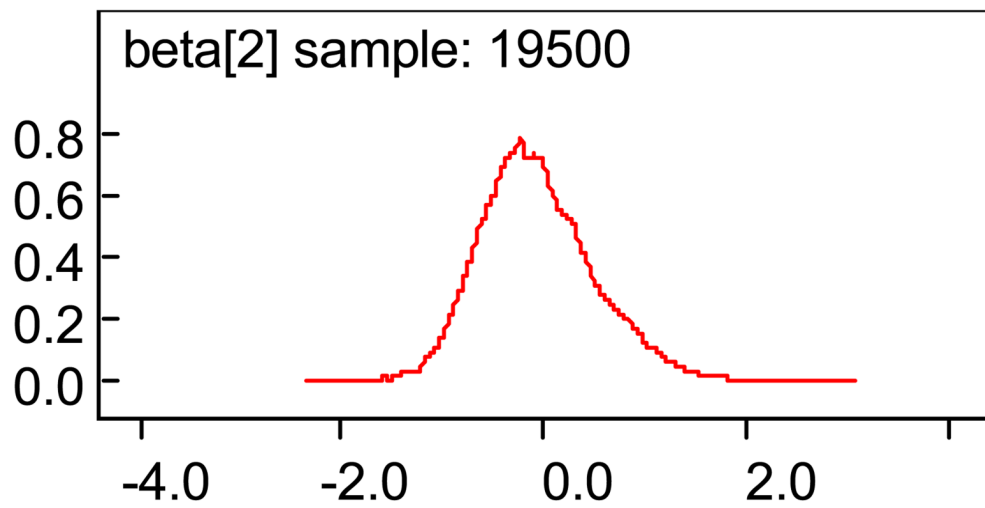


Figure 4.
Posterior Densities

Table 1

Parameter estimate values for equation (1), the normal hierarchical model.

| Parameter | mean | stand dev | 95% Credible Region | |
|------------|--------|-----------|---------------------|---------|
| | | | 2.5% | 97.5% |
| α_c | 20.02 | 0.2952 | 14.16 | 25.85 |
| β_c | -1.269 | 0.4883 | -2.003 | -0.4953 |
| σ | 0.6158 | 0.7370 | 0.0373 | 1.3100 |

Table 2

Parameter estimate values for equation (2), the changepoint model.

| Parameter | mean | stand dev | 95% Credible Region | |
|-----------|---------|-----------|---------------------|---------|
| | | | 2.5% | 97.5% |
| α | 20.05 | 0.5835 | 18.94 | 21.23 |
| β_1 | -0.7797 | 0.3099 | -1.384 | -0.7829 |
| β_2 | -0.0632 | 0.5781 | -1.060 | 1.2010 |
| σ | 1.3040 | 0.1755 | 1.0150 | 1.6990 |
| x_k | 2.4020 | 0.4303 | 1.5580 | 2.4860 |