



Published in final edited form as:

*Psychol Aging*. 2013 June ; 28(2): 377–385. doi:10.1037/a0031000.

## Investigating Terminal Decline: Results From a UK Population-Based Study of Aging

**Graciela Muniz-Terrera,**

MRC Unit for Lifelong Health and Ageing, London, UK

**Ardo van den Hout,**

Department of Statistical Science, University College London

**Andrea M. Piccinin,**

Department of Psychology, University of Victoria

**Fiona E. Matthews,** and

MRC Biostatistics Unit, Cambridge, UK

**Scott M. Hofer**

Department of Psychology, University of Victoria

### Abstract

The terminal decline hypothesis states that in the proximity of death, an individual's decline in cognitive abilities accelerates. We aimed at estimating the onset of faster rate of decline in global cognition using Mini Mental State Examination (MMSE) scores from participants of the Cambridge City over 75 Cohort Study (CC75C), a U.K. population-based longitudinal study of aging where almost all participants have died. The random change point model fitted to MMSE scores structured as a function of distance to death allowed us to identify a potentially different onset of change in rate of decline before death for each individual in the sample. Differences in rate of change before and after the onset of change in rate of decline by sociodemographic variables were investigated. On average, the onset of a faster rate of change occurred about 7.7 years before death and varied across individuals. Our results show that most individuals experience a period of slight decline followed by a much sharper decline. Education, age at death, and cognitive impairment at study entry were identified as modifiers of rate of change before and after change in rate of decline. Gender differences were found in rate of decline in the final stages of life. Our study suggests that terminal decline is a heterogeneous process, with its onset varying between individuals.

### Keywords

random change point models; terminal decline; cognitive decline; MMSE; old age

---

The terminal decline hypothesis proposes an association between proximity to death and cognitive decline suggesting that “years to live” or proximity to death may be better indicators of cognitive performance than chronological age before death (Siegler, 1975). Two issues emerging as critical in the terminal decline literature are whether terminal decline is detectable and, when detectable, whether it is possible to identify its onset.

Although agreement exists among researchers that terminal decline is a within-person process and should be modeled as such, most previous attempts to identify its onset have assumed that all individuals experience a change in rate of decline at the same distance from death. Here, we test whether heterogeneity exists across individuals in the onset of terminal decline.

Several studies have reported results in support of the hypothesis (Batterham, Mackinnon, & Christensen, 2011; Dodge, Wang, Chang, & Ganguli, 2011; Gerstorf, Ram, Hoppmann, Willis, & Schaie, 2011; Laukka, MacDonald, & Bäckman, 2006, 2008; Muniz-Terrera, Matthews, Stephan, & Brayne, 2011; Wilson, Beck, Bienias, & Bennett, 2007; Wilson, Beckett, Bienias, Evans, & Bennett, 2003), although others reported mixed or no evidence of accelerating decline before death (Piccinin, Muniz, Matthews, & Johansson, 2011).

Differences in study design, sensitivity to change of the measures used to assess the cognitive domains of interest, and differences in the methodologies employed may explain these divergent reports. Study design features such as the number and separation of follow-up interviews, and length of follow-up are likely to impact reports. An insufficient number of data waves may hamper the examination of accelerating decline as longitudinal models considered for its examination should include terms to describe a nonlinear trajectory, and such models are only identifiable if data from at least three data waves are analyzed. In addition, if interviews are far apart, individuals may die between interviews without providing sufficient data to capture a change in cognitive performance or in rate of decline before death.

Different sensitivity to change of cognitive abilities prior to death may also impact reports (Bosworth & Siegler, 2002). Crystallized ability has been reported to be better maintained with aging than fluid abilities and decline in this ability more likely to be associated with impending death. Hence, investigations of change due to proximity to death are more likely to report changes where crystallized abilities are examined (White & Cunningham, 1988).

A recent publication by Gerstorf and colleagues (Gerstorf et al., 2011) also suggested the possibility of cohort differences explaining differences in terminal decline reports, although they reported few secular trends, and the literature about cohort differences in the terminal decline context is scarce.

Critical methodological discrepancies in reports have been extensively discussed in Piccinin et al. (2011). They include the choice of the time metric considered to structure change (Sliwinski, Hofer, Hall, Buschke, & Lipton, 2003); the separation of between- and within-person effects (Mehta & West, 2000); and the treatment of data from demented individuals and survivors.

As the terminal decline hypothesis implies a change in rate of decline, only models that assume a nonconstant rate of decline are adequate for its description. The two most commonly used models in the terminal decline literature are mixed effects models that include a squared time metric term (often referred as quadratic models), and mixed effects change point models (Smith, 1975). While quadratic models inform about rate of decline and its acceleration, they do not inform about the onset of such acceleration. Instead, change point models do. Most change point models considered in the aging literature describe change as consisting of two linear phases that intersect at a change point, with the change point either regarded as a fixed effect (i.e., assumed to be common to all individuals) or a random effect (change point varying across individuals). The application of change point models in cognitive aging increased substantially in recent years after a seminal publication by Hall, Lipton, Sliwinski, and Stewart (2000), who employed these models to estimate the onset of memory changes in preclinical dementia.

The onset of terminal decline has been identified in several publications where profile likelihood was used to estimate models where the change point was modeled as a fixed effect (MacDonald, Hultsch, & Dixon, 2011; Sliwinski et al., 2006; Thorvaldsson et al., 2008; Wilson et al., 2003). Estimates of the onset of faster decline varied from about 6 years before death (Sliwinski et al., 2006) to over 14 years before death (Thorvaldsson et al., 2008). Although these models advanced our understanding of cognitive changes in the final stages of life, the assumption of a common change point model in heterogeneous samples is questionable in the context of terminal decline (a process that is understood, per definition, to be as a within-person process) and deserves further investigation.

In this work, we examined whether the onset of terminal decline in global cognition varied across individuals in the sample. With this aim, using global cognitive scores from a subsample of deceased participants of the Cambridge City over 75 Cohort Study (CC75C) structured as a function of distance to death, we initially considered a change point model where the change point was modeled as a fixed effect, and consequently, estimated a common change point for all individuals in the sample. Second, we relaxed the common change point assumption and estimated a model where the change point was modeled as a random effects. This model estimated a potentially different change point for each individual in the sample. The fact that at the time of the analysis only very few study participants remained alive is a strong feature of our study.

## Method

### Participants

The Cambridge City over 75 Cohort Study (CC75C) is a prospective study of a representative population sample of Cambridge City residents aged 75 years and older in 1985. Its focus is prevalence and incidence of dementia, risk factors for cognitive decline and dementia, neuropsychology, and depression. All patients registered in six general practices and one third of a seventh practice, including patients living in residential and nursing homes, were invited to take part in the study by their doctors, and 95% accepted. After excluding participants from one practice ( $n = 445$ ) due to differential recruitment and their concurrent participation in an intervention study, all remaining participants ( $N = 2,166$ ) were screened by a trained interviewer who recorded patients' details, family contacts, health status, and use of health services. The screening interview was followed by a more detailed clinical interview of all individuals scoring 23 or less on the Mini Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975) and a third of those with scores of 24 and 25 points. After the baseline interview, up to six follow-up interviews were carried out to establish incidence of dementia.

### Measures

Cognitive status was assessed with the MMSE, a test originally designed as a screening tool for dementia, but widely used as an indicator for global cognition. The MMSE takes integer values in the 0–30 interval, with high values indicating good cognitive status. When a question was omitted, or not applicable due to sensory or physical impairment, the item was scored as zero for calculation of the final MMSE score. Sociodemographic information and information about physical ability was collected (see [www.cc75c.cam.ac.uk](http://www.cc75c.cam.ac.uk) for more details) at baseline and other survey interviews.

To identify the deceased subsample, information about participants' date of death was obtained from the National Health Service central register. On average, the interval between the last interview and death was 2.8 years ( $SD = 2.6$ , range = [0.7, 19.7]). As the time interval between dropout and death was large in some cases and our aim was to model

terminal decline, we selected the subsample of deceased CC75C participants ( $n = 1,896$ ) who were last seen less than 10 years before death. Excluded participants were only about 7% of individuals. The number of individuals who were in the selected subsample and who had missed from 1 to 6 interviews were 27, 92, 206, 318, 471, and 778, respectively. Distance to death from each interview was calculated as the difference between each person's date of death and the interview dates. On average, individuals were 81 years old at study entry and died at the age of 88 years old. Sixty-five per cent of individuals in the deceased sample were women and 37% per cent of individuals had a nonmanual profession (See Table 1 for descriptive characteristics of the sample).

Participants were asked if they could walk unaided around the block. Those who answered "no" were classified as physically impaired. According to this criterion, 43% of the individuals in the sample were physically impaired at study entry. Baseline cognitive scores were used to define a cognitive impairment indicator that took the value of 1 if the individual scores below 23 MMSE point at baseline and 0 otherwise. Fifteen percent of individuals in the sample analyzed were cognitively impaired at study entry.

## Procedure

We fitted two mixed models with a change point. We first modeled the change point as a fixed effect (assuming a common change point for all individuals [FCP]), then, we relaxed the common change point assumption and modeled the change points as random effects (RCP). Linear mixed effects models are usually formulated by an equation that describes within-individual change (level 1) and a set of equations that describe individual-level parameters as a function of both fixed effects and individual-level covariates (level 2). Following this convention, a mathematical formulation of the RCP model fitted in our analysis is:

$$\text{Level 1: } MMSE_{it} = (\alpha_{0i} + \alpha_{1i}T_{it})[1 - I(T_{it} - \gamma_i)] + (\alpha_{0i} + \alpha_{1i}\gamma_i + \alpha_{2i}(T_{it} - \gamma_i))I(T_{it} - \gamma_i) + \varepsilon_{it} \quad (1)$$

$$\begin{aligned} \text{Level 2: } \alpha_{0i} &= \alpha_{00} + v_{0i} \\ \alpha_{1i} &= \alpha_{10} + \beta_1^T Z_i + v_{1i} \\ \alpha_{2i} &= \alpha_{20} + \beta_2^T Z_i + v_{2i} \end{aligned} \quad (2)$$

where  $\gamma_i$  = individual  $i$ 's change point, and  $I(T_{it} - \gamma_i) = 1$  if  $T_{it} - \gamma_i \geq 0$  and 0 otherwise.  $T_{it}$  (for  $i = 1, \dots, N$  with  $N$  = number of individuals in the sample; and  $t = 1, \dots, n_j$  with  $n_j$  = number of times individual  $i$  was observed) is the time (in years) from each interview to individual  $i$ 's date of death. The mathematical formulation of the FCP can be obtained from the expression shown in equation (1) replacing  $\gamma_i$  by  $\gamma$ .

Distances from death to each interview were coded as negative values to maintain chronological ordering. For ease of interpretation, we set the intercept of the models at 2 years before death. Random effects  $\alpha_{1i}$ ,  $\alpha_{2i}$  represent individual  $i$ 's annual rate of change per year closer to death before and after the change point, respectively, and were modeled as deviations from  $\alpha_{10}$ ,  $\alpha_{20}$ , which represent the expected rates of change for individuals with the reference values in all covariates  $Z_i$ . Cognitive performance 2 years before death is calculated as  $\alpha_{0i} + (\alpha_{1i} - \alpha_{2i})\gamma_i$ . Given the model parameterization and time scale,  $\alpha_{0i}$  is not a parameter of direct relevance (it represents individual  $i$ 's performance in the absence of change in rate of decline) and so is not regressed on the baseline covariates.

Residuals  $\varepsilon_{it}$  were assumed to follow a normal distribution with mean 0 and variance  $\tau^2$ , that is,  $\varepsilon_{it} \sim N(0, \tau^2)$ , for all values of  $i$  and  $t$ . Level 2 residuals  $v_{0i}$ ,  $v_{1i}$ ,  $v_{2i}$  for  $i = 1, \dots, N$  were assumed to follow a normal distribution with mean 0 and covariance matrix  $\Omega$ .

Modifiers of rate of decline in our analysis were education, gender, age at death, and physical and cognitive impairment at study entry, as these have been identified as associated with cognitive decline (Piccinin et al., 2011). Education and age at death were considered continuous variables (centered at their mean value). Specifically, education was centered at age 14, the mean age at which individuals left full-time education, and age at death was centered at age 88, the average age of death of individuals in the sample. An indicator of physical impairment that took the value of 1 if the individual could not walk unaided around the block and 0 otherwise; a gender indicator that took the value of 1 for women and 0 for men and a cognitive impairment indicator that took the value of 1 if the baseline MMSE score was below 23 and 0 if the MMSE score was higher or equal to 23 MMSE points.

Models were fitted using a Bayesian framework (Gelman, Carlin, Stern, & Rubin, 2012), in which conclusions about a parameter of interest,  $\theta$ , are made in terms of probability statements conditional on the observed values of the data,  $y$ , written as  $p(\theta|y)$  and called the posterior distribution of parameter  $\theta$ . (This is in contrast to the traditional inferential framework, in which probabilities resulting from statistical tests define the likelihood of the data, given the parameters of the selected model. Often, researchers relying on the inferential framework interpret, however, the resulting probabilities of their analyses as the likelihood of the parameters in light of the data. Thus, although the inferential framework is applied, results are interpreted in a Bayesian framework). To make statements about the parameter of interest, the probability  $p(\theta|y)$  is factorized in terms of the likelihood function and a prior distribution that expresses the uncertainty about the parameter  $\theta$  before the data are considered. The posterior probability contains all the current information about the parameter  $\theta$ , and can be described using a range of numerical summaries that include, among others, the mean, standard deviation, and interquartile range. Because the posterior distribution is derived from a combination of the data and the prior distribution, it is standard practice to examine the sensitivity of results obtained with respect to the choice of priors for the key parameters of the model.

Model selection was based on the Deviance Information Criterion (DIC) (Spiegelhalter, Best, Carlin, & Van Der Linde, 2002), described in the Appendix.

## Results

### Sample Demographics

Initially, a comparison of characteristics of individuals who were excluded from the analysis due to having dropped out of the study more than 10 years before death was conducted. Excluded individuals died younger than individuals included in the analysis, were older at baseline and had lower cognitive function at study entry ( $t$  tests,  $p < .05$ ).

### Random Change Point Model

DIC values for the FCP (48151.9) and RCP (46684.1) models suggested that the RCP model supported the data better, when compared to the FCP model. Hence, we report results from the RCP model.

On average, individuals experienced a change in rate of decline at about 7.7 ( $SD = 0.21$ ) years before death. This value was calculated by subtracting 2 from the estimated mean change point ( $-5.7$ ), as time zero in our models corresponds to 2 years before death. The onset of change in rate of decline varied considerably between individuals as indicated by the variance in the distribution of the change points and its narrow credible interval (see bottom of Table 2), and is shown graphically in the bottom panel of Figure 1, where we plotted the distribution of the posterior mean estimates of the individuals change points.

MMSE performance 2 years before death was estimated as 23.9 ( $SD = 0.23$ ). Before the onset of change in rate of decline, annual rate of change was estimated at  $-0.13$  ( $SD = 0.02$ ) MMSE points per year closer to death for a not cognitively nor physically impaired man who died aged 88 years old and had left school aged 14 years. After the change point, the rate of decline was significantly faster, estimated at  $-0.61$  ( $SD = 0.05$ ) MMSE points per year closer to death. Between person variability in these rates of change were also much larger after the change point. Figure 2 shows the estimated mean trajectory and its 95% Credible Interval for an individual with reference values in all covariates. The correlation between mean rate of change before and after the change point was calculated as  $-0.09$ . This value suggests a negative but very small association between rate of decline before and after the onset of terminal decline.

Posterior mean estimates of each individual's slope before and after the change point were obtained and their difference calculated by subtracting the estimated slope after the change point of each individual from the corresponding estimated slope before the change point. The mean difference of the slopes was calculated at 0.76 ( $SD = 0.56$ ), and took values in the interval  $[-0.53, 2.86]$ . The distribution of the difference of the rate of decline before and after each individual's change point is shown in the top panel of Figure 1.

### Predictors of Change

Our results about the effect of potential modifiers of rate of decline before the onset of more rapid decline indicate that per extra year of older age at death, individuals declined at a slower rate than individuals who died younger (see Table 2). However, once they passed the onset of terminal decline, their rate of change was faster than the rate of change of individuals who died younger. More educated individuals declined more rapidly before the onset of terminal decline than less educated individuals, but after the onset of terminal decline, their rate of decline was slower than the rate of decline of less educated individuals.

Before the onset of terminal decline, no significant differences in rate of change were found between men and women, although women declined faster than men in the final stages of life. Estimates of the effect of physical impairment on rate of decline before and after the onset of faster rate of decline did not reach significance, although they suggest that individuals who were physically impaired at study entry declined more rapidly than those who were not impaired before and after the onset of faster rate of decline.

The direction of the effect of cognitive impairment at study entry on rate of decline before the onset of faster rate of change was not the one initially expected, as our results suggest that individuals who were labeled as impaired at study entry declined at a slower rate than those who were not impaired.

In an additional model, both slopes were adjusted for baseline age and the interaction of age and education. Baseline age was found to be significant on both slopes but had relatively small estimates ( $-0.030$  [ $SD = 0.009$ ] on the first slope and  $0.3$  [ $SD = 0.5$ ] on the second slope) indicating that older individuals at baseline decline slightly faster before the change point but decline at a slower rate after the change point. All other estimates remained relatively unchanged.

### Discussion

In this study, we examined terminal decline in global cognition measured by MMSE scores in a population-based longitudinal study of older persons where the vast majority of the individuals in the sample died over the course of the study. The RCP model was found to fit the data better than the FCP model. This extends previous reports by demonstrating that the



onset of terminal decline varies across individuals. For a physically and cognitively able man who left school aged 14 years and died aged 88 years, the onset of faster rate of decline was estimated at about 7.5 years before death. The difference in rate of decline before and after the change point suggests that most individuals experience a period of relatively slow decline followed by a period of sharper decline in the years prior to death.

Going beyond our previous examination of the terminal decline hypothesis in the CC75C study (Muniz-Terrera, Matthews, Denning, Huppert, & Brayne, 2009), the random change point model fitted here allowed us to identify the onset of change in rate of decline for each person in the subsample examined. In addition, to minimize the effect of between-person differences in estimates of key model parameters such as change points, and to maximize the model's ability to capture the within-person nature of the terminal decline process, we have now considered the subsample of individuals who were last seen within 10 years before death. Models converged more quickly when individuals who contributed observations far from death were removed from the sample.

We adjusted our model for possibly relevant risk factors for terminal decline. Education emerged as a modifier of rate of decline before and after the onset of more rapid decline, however, against our initial expectations and compensation or cognitive reserve theories (Park & Reuter-Lorenz, 2009; Stern, 2002) as more educated individuals experienced a more rapid decline before, and slower decline after the change point. Batterham and colleagues (Batterham et al., 2011), who examined the effect of education in a range of cognitive measures, also reported that the only hint of a protective effect of education against decline was found in their global measure. Others have reported that more educated individuals performed better before death in a range of cognitive tasks (Laukka et al., 2006; Piccinin et al., 2011). Wilson et al. (2007) reported no association between terminal decline and education in an investigation of change in global cognitive function. Differences in reports may stem from several sources, including, among others, variations in study design, how education was measured, and features of the samples, as discussed extensively by Anstey and Christensen (2000) and Batterham et al. (2011).

In our study, we derived a variable to account for differences in physical ability by asking individuals if they could walk unaided around the block. Although our results suggest faster rate of decline before and after the onset of the terminal phase for physically impaired individuals, estimates of the effect of physical impairment on the corresponding parameters did not reach statistical significance. Wilson and colleagues (2007) considered a disability measure on the Katz scale and similarly reported no significant effect on terminal decline. Although the longitudinal association between physical ability and cognitive change in the proximity of death is not well studied, evidence exists for their cross sectional association (Gillum & Obisesan, 2010) and for the effect of physical function on age-related cognitive decline (Nilsson et al., 2007; Shipley, Der, Taylor, & Deary, 2007).

Our results about the effect of cognitive impairment on rate of change before and after the change point went against expectations, as cognitively impaired individuals appeared to have a slower rate of decline before the change point than those categorized as nonimpaired. These results could be explained by reasons that include the method used to code item nonresponse due to physical or sensory impairment, as when individuals could not answer a question due to impairment, questions were scored as zero. This practice may have led to a misclassification of some impairment cases, as we used a cutpoint of 23 in the MMSE score to define dementia cases because of the lack of formal diagnosis. It is also possible that this was influenced by severely impaired individuals who were nearing the MMSE scoring floor and had already experienced a change point prior to the start of the study.

The treatment of dementia cases in examination of terminal decline is important, as cognitive changes may be driven by dementia (Laukka et al., 2008), although evidence of terminal decline has also been reported in analyses where dementia was explicitly accounted for (Piccinin et al., 2011). Dementia cases have been treated in different ways in the literature, with some researchers excluding individuals demented at baseline from the analysis (Sliwinski et al., 2003; Wilson et al., 2007), while others also excluded incident cases (Sliwinski et al., 2006; Thorvaldsson et al., 2008) and some either included demented and incident cases without adjustment, or accounted for them explicitly in the models (Ghisletta, McArdle, & Lindenberger, 2006; Piccinin et al., 2011).

We adjusted our models for differences by age at death to separate between-person effects of being old at death from within-person effects of approaching death. This is common practice in studies of aging that structure time as time in study (usually these models include terms to account for differences in initial age), but not so common in terminal decline studies. Piccinin and colleagues (Piccinin et al., 2011) accounted for distance to death at study entry and age at baseline with a similar aim.

Older age at death was associated with slower rate of change before the onset of terminal decline and with faster rate of change after it. However, once the terminal phase started, individuals who died at an older age declined faster than individuals who died younger. These results are similar to the survivor effect reported in the context of studies of aging and has been reported previously in studies of terminal decline (Piccinin et al., 2011). The impact of older age at death on terminal decline of perceptual speed has been associated with poorer performance but not rate of change or acceleration (Thorvaldsson et al., 2008).

An excellent overview of the terminal decline literature can be found in Bäckman and MacDonald (2006). Laukka et al. (2006) examined MMSE scores using a linear random effects model where analyses were conducted for the total sample and separately for individuals younger than 81 years old and those older than 81 years old. Cognitive performance 3 years before death for individuals who remained free of dementia and survived to the end of the study was estimated at 26.56, with a loss of almost half a point for both those in the preclinical dementia phase and those in the impending death group. Rate of decline was estimated as  $-0.18$ ,  $-0.02$  and  $-0.15$  points per year closer to event for the different groups. Although this analysis represents an interesting cross group comparison, it provides limited information about terminal decline, as the model fitted, which assumes constant rate of change, does not allow for a change in rate of decline, an essential feature of the terminal decline hypothesis.

In reports where change point models were fitted to results of cognitive tests other than MMSE, estimates varied widely, from 2.8 to 15 years before death, depending on the cognitive domain (Thorvaldsson et al., 2008; Wilson et al., 2003). These varied results, and our own, suggest that terminal decline does not occur uniformly across domains and individuals. The disparity of results may reflect the differential impact of a range of possible influences on individuals' biological systems before death.

Our modeling approach has several methodological advantages. To begin, we maximized the use of all available data, as we were able to use data from all individuals, even of individuals who contributed with only one measurement. Estimates of population and individual change points reflect a combination of information contributed mainly by individuals with at least four observations, although individuals who were seen in fewer occasions still contribute to the estimation. Change points are estimated for individuals whose change points are censored. For these individuals, estimated change points are a



combination of the information provided by the remaining sample and the prior distribution considered.

Furthermore, Bayesian estimation has extended computational facilities compared to an inferential approach, and this framework allowed us to use prior beliefs about change point location. Win-BUGS (Lunn, Thomas, Best, & Spiegelhalter, 2000), a user-friendly package, offers extensive flexibility to fit mixed effects models in terms of the choice of prior distributions. Further, its use of Monte Carlo Markov Chain techniques facilitates estimation of random effects, and predicted values for individuals can be easily obtained, even in the context of more complex random effect structures. In a marginal likelihood framework, instead, random effects are integrated out and extra steps are required to obtain their estimates.

Our study also has some limitations. First, the model still requires normality of the errors, an assumption that MMSE scores may not fulfill as they have a skewed distribution. Data transformations could aid in fulfilling model distributional assumptions; however, such transformations complicate interpretation of model parameters and do not preserve distances, potentially biasing estimates of the change points. Second, we investigated change in a crude measure of global cognition, total MMSE scores. Domain-specific partial scores were not available for analysis, limiting our ability to examine whether these results were consistent across domains. Initially we would not expect consistency across domains, as evidence suggests the onset of change in rate of decline varies widely across abilities (Thorvaldsson et al., 2008). A third possible limitation arises from the fact that individual questions were scored zero when the individual could not perform a task because of physical or sensory impairment. Although this (standard) practice is likely to bias results, it was beyond our ability to examine these potential biases as scores had already been calculated at the time of analysis. Fourth, although we would have favored an analysis where formerly diagnosed dementia cases, incident dementia, and disease severity were accounted for, due to the lack of formal dementia diagnosis, we only used a crude indicator of cognitive impairment at study entry. Finally, the lack of information about cause of death did not allow us to differentiate patterns of decline of individuals who died from accidental causes.

In this investigation, we used data from a population-based study where the vast majority of participants have died over the course of the study. This feature allowed us to maximize the use of data while fitting a model with an appropriate time metric for the description of the process under examination. The identification of heterogeneity in individual's onset of faster rate of decline before death suggests that loss of abilities preceding death is not a homogeneous process across individuals. Extensions of our analysis to investigate the effect of covariates on the location of the change point, which will permit study of the impact of modifiable risk factors on the onset of terminal decline, are in progress.

## Acknowledgments

This research was supported by an MRC grant UC\_US\_A030\_0031 awarded to Fiona E. Matthews and a grant from the National Institute on Aging, National Institutes of Health Grant AG026453 awarded to Scott M. Hofer and Andrea M. Piccinin.

## References

- Anstey K, Christensen H. Education, activity, health, blood pressure and apolipoprotein E as predictors of cognitive change in old age: A review. *Gerontology*. 2000; 46:163–177.10.1159/000022153 [PubMed: 10754375]
- Bäckman L, MacDonald SW. Death and cognition: Synthesis and outlook. *European Psychologist*. 2006; 11:224–235.

- Batterham PJ, Mackinnon AJ, Christensen H. The effect of education on the onset and rate of terminal decline. *Psychology and Aging*. 2011; 26:339–350.10.1037/a0021845 [PubMed: 21480720]
- Bosworth HB, Siegler IC. Terminal change in cognitive function: An updated review of longitudinal studies. *Experimental Aging Research*. 2002; 28:299–315.10.1080/03610730290080344 [PubMed: 12079580]
- Dodge HH, Wang CN, Chang CC, Ganguli M. Terminal decline and practice effects in older adults without dementia: The MoVIES project 1. *Neurology*. 2011; 77:722–730.10.1212/WNL.0b013e31822b0068 [PubMed: 21832224]
- Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189–198.10.1016/0022-3956(75)90026-6 [PubMed: 1202204]
- Gelman, A.; Carlin, J.; Stern, H.; Rubin, D. Bayesian data analysis. 2. Boca Raton: Chapman & Hall/CRC; 2012. Posterior simulation; p. 290-318.
- Gelman A, Rubin D. Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*. 1992; 7:457–472.10.1214/ss/1177011136
- Gerstorff D, Ram N, Hoppmann C, Willis SL, Schaie KW. Cohort differences in cognitive aging and terminal decline in the Seattle Longitudinal Study. *Developmental Psychology*. 2011; 47:1026–1041.10.1037/a0023426 [PubMed: 21517155]
- Ghisletta P, McArdle J, Lindenberger U. Longitudinal cognition–survival relations in old and very old age: 13-year data from the Berlin Aging Study. *European Psychologist*. 2006; 11:204–223.10.1027/1016-9040.11.3.204
- Gillum RF, Obisesan TO. Physical activity, cognitive function, and mortality in a US national cohort. *Annals of Epidemiology*. 2010; 20:251–257.10.1016/j.annepidem.2010.01.005 [PubMed: 20171901]
- Hall CB, Lipton RB, Sliwinski M, Stewart WF. A change point model for estimating the onset of cognitive decline in preclinical Alzheimer’s disease. *Statistics in Medicine*. 2000; 19:1555–1566.10.1002/(SICI)1097-0258(20000615/30)19:11/12<1555::AID-SIM445>3.0.CO;2-3 [PubMed: 10844718]
- Laukka EJ, MacDonald SW, Bäckman L. Contrasting cognitive trajectories of impending death and preclinical dementia in the very old. *Neurology*. 2006; 66:833–838.10.1212/01.wnl.0000203112.12554.f4 [PubMed: 16567699]
- Laukka EJ, MacDonald SW, Bäckman L. Terminal-decline effects for select cognitive tasks after controlling for preclinical dementia. *The American Journal of Geriatric Psychiatry*. 2008; 16:355–365. [PubMed: 18448848]
- Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS-A Bayesian modelling framework: Concepts, structure and extensibility. *Statistics and Computing*. 2000; 10:325–337.10.1023/A:1008929526011
- MacDonald SW, Hulstsch DF, Dixon RA. Aging and the shape of cognitive change before death: Terminal decline or terminal drop? *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*. 2011; 66B:292–301.10.1093/geronb/gbr001
- Mehta PD, West SG. Putting the individual back into individual growth curves. *Psychological Methods*. 2000; 5:23–43.10.1037/1082-989X.5.1.23 [PubMed: 10937321]
- Muniz-Terrera G, Matthews F, Denning T, Huppert FA, Brayne C. Education and trajectories of cognitive decline over 9 years in very old people: Methods and risk analysis. *Age and Ageing*. 2009; 38:277–282.10.1093/ageing/afp004 [PubMed: 19252209]
- Muniz-Terrera G, Matthews FE, Stephan B, Brayne C. Are terminal decline and its potential indicators detectable in population studies of the oldest old? *International Journal of Geriatric Psychiatry*. 2011; 26:584–592.10.1002/gps.2566 [PubMed: 21480375]
- Nilsson SE, Read S, Berg S, Johansson B, Melander A, Lindblad U. Low systolic blood pressure is associated with impaired cognitive function in the oldest old: Longitudinal observations in a population-based sample 80 years and older. *Aging Clinical and Experimental Research*. 2007; 19:41–47. [PubMed: 17332720]
- Park DC, Reuter-Lorenz P. The adaptive brain: Aging and neurocognitive scaffolding. *Annual Review of Psychology*. 2009; 60:173–196.10.1146/annurev.psych.59.103006.093656

- Piccinin AM, Muniz G, Matthews FE, Johansson B. Terminal decline from within- and between-person perspectives, accounting for incident dementia. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*. 2011; 66B:391–401.10.1093/geronb/gbr010
- Riegel KF, Riegel RM. Development, drop, and death. *Developmental Psychology*. 1972; 6:306–319.10.1037/h0032104
- Shiple BA, Der G, Taylor MD, Deary IJ. Association between mortality and cognitive change over 7 years in a large representative sample of UK residents. *Psychosomatic Medicine*. 2007; 69:640–650.10.1097/PSY.0b013e31814c3e7c [PubMed: 17846257]
- Siegler IC. The terminal decline hypothesis: Fact or artifact? *Experimental Aging Research*. 1975; 1:169–185.10.1080/03610737508257957 [PubMed: 802856]
- Sliwinski MJ, Hofer SM, Hall C, Buschke H, Lipton RB. Modeling memory decline in older adults: The importance of preclinical dementia, attrition, and chronological age. *Psychology and Aging*. 2003; 18:658–671.10.1037/0882-7974.18.4.658 [PubMed: 14692855]
- Sliwinski MJ, Stawski RS, Hall CB, Katz M, Verghese J, Lipton R. Distinguishing preterminal and terminal cognitive decline. *European Psychologist*. 2006; 11:172–181.10.1027/1016-9040.11.3.172
- Smith A. A Bayesian approach to inference about a change-point in a sequence of random variables. *Biometrika*. 1975; 62:407–416.10.1093/biomet/62.2.407
- Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*. 2002; 64:583–639.10.1111/1467-9868.00353
- Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*. 2002; 8:448–460.10.1017/S1355617702813248 [PubMed: 11939702]
- Thorvaldsson V, Hofer SM, Berg S, Skoog I, Sacuiu S, Johansson B. Onset of terminal decline in cognitive abilities in individuals without dementia. *Neurology*. 2008; 71:882–887.10.1212/01.wnl.0000312379.02302.ba [PubMed: 18753475]
- White N, Cunningham WR. Is terminal drop pervasive or specific? *Journal of Gerontology*. 1988; 43:P141–P144. [PubMed: 3183309]
- Wilson RS, Beck TL, Bienias JL, Bennett DA. Terminal cognitive decline: Accelerated loss of cognition in the last years of life. *Psychosomatic Medicine*. 2007; 69:131–137.10.1097/PSY.0b013e31803130ae [PubMed: 17327212]
- Wilson RS, Beckett LA, Bienias JL, Evans DA, Bennett DA. Terminal decline in cognitive function. *Neurology*. 2003; 60:1782–1787.10.1212/01.WNL.0000068019.60901.C1 [PubMed: 12796531]

## Appendix

### Bayesian Implementation of Random Change Point Model

Bayesian inference was applied to estimate all models. MCMC methods to construct Markov chains that have the posterior distribution as its stationary distribution were used as in WinBUGS (Lunn et al., 2000).

We can express the model fitted as:

$$MMSE_{it} = f(T_{it}, \gamma_i, \alpha_i)$$

$$\alpha_{ki} = g_i(\alpha_{kk}, Z_i, \beta)$$

where  $f(T_{it}, \gamma_i, \alpha_i)$  is as in equation (1) in the text and  $g_i(\alpha_{kk}, Z_i, \beta)$  as in equation (2) with  $\alpha_{kk}$ ,  $\beta$ , and  $Z$  are also as before for  $k = 0, 1, 2$  and  $i = 1, \dots, N$ .

For sake of brevity, we only describe the distributional assumptions for the RCP:

$$\begin{aligned}
 MMSE_{it}|T_{it}, \gamma_i, \alpha_i, \tau &\sim N(f_i(T_{it}, \gamma_i, \alpha_i), \tau^2) \\
 \alpha_i|\alpha_{kk}, Z_i, \beta, \Omega &\sim MVN(g_i(\alpha_{kk}, Z_i, \beta), \Omega) \\
 \gamma_i &\sim N(\mu_\gamma, \zeta^2)I(-20, 2)
 \end{aligned}$$

with  $f_i$  and  $g_i$  as before.

## Specification of Prior Distributions

Fixed effects  $\alpha_{kk}$  ( $k = 1, 2$ ) were modeled using vague normal prior distributions. The standard deviation of the level 1 residuals ( $\tau$ ) was modeled using a uniform distribution prior distribution, that is,  $\tau \sim 0 \text{ Uni}(0, 10)$  and  $\Omega^{-1}$ . The inverse of the variance covariance matrix of random effects residuals was modeled using a Wishart (R,3) distribution where R is a diagonal matrix with entries  $R(1,1) = R(3,3) = 4$  and  $R(2,2) = 2$ . In all models, normal distributions with large variances were considered as prior distributions of all regression coefficients.

Although all efforts were made to collect information on covariates, 14 individuals were missing data on disability or age at death. To avoid listwise deletion, we specified a model for the distribution of these variables, from which the MCMC in WinBUGS simulates values for the missing observations. A Bernoulli distribution with parameters that were given uniform priors was used to model disability status. Age at death was modeled using a normal distribution.

Under an exchangeability assumption, we assumed random change points  $\gamma_i$ ,  $i = 1, \dots, N$  to follow a truncated normal distribution, that is,  $\gamma_i \sim N(\mu_\gamma, \zeta^2)I(-20, 2)$ . The truncation was used to avoid the estimation of the change points after the death of the individuals and the truncation interval chosen in consideration of the fact that the time metric takes negative values and the intercept of the models was set to be 2 years prior to death.

We imposed a hierarchical structure with hyper-parameters  $\mu_\gamma$  and  $\zeta^2$  that were modeled using vague priors. More specifically,  $\mu_\gamma \sim \text{Uni}(-20, 2)$  and  $\zeta \sim \text{Uni}(0, 10)$  respectively. These parameters and distributions were chosen to take into account previous reports about the onset of terminal decline (Riegel & Riegel, 1972). Distributional assumptions of the FCP model fitted are similar to the ones considered for the RCP model.

## Posterior Distributions

The posterior distribution often does not have a closed form, requiring numerical methods to draw values of the parameter of interest,  $\theta$ . One technique to draw these values is to construct a Markov chain that converges to the posterior distribution  $p(\theta|y)$ . Intuitively, the idea behind Markov chain simulation is “to sample iteratively in such a way that at each step of the process we expect to draw from a distribution that is closer and closer to the posterior distribution  $p(\theta|y)$ ” (Gelman et al., 2012). This is the technique used in WinBUGS (Lunn et al., 2000), the software package used to estimate the models in this paper.

## Convergence Assessment

Two tools commonly used to assess convergence of the constructed Markov chain to the posterior distribution are traceplots and the Gelman-Rubin diagnostic (Gelman & Rubin, 1992). Traceplots plot the parameter value at each time against the iteration number. In the presence of nonconvergence, traceplots show some trending in the sample space. The

Gelman-Rubin diagnostic is such that convergence is shown when, after starting multiple chains from different starting points, these chains display a similar behavior.

We discarded the initial 50,000 burn-in samples and based our inferences on the subsequent 50,000 samples. Convergence was assessed visually using the Gelman-Rubin statistic that requires inspecting samples and trace plots from runs based on three chains initiated with different initial values.

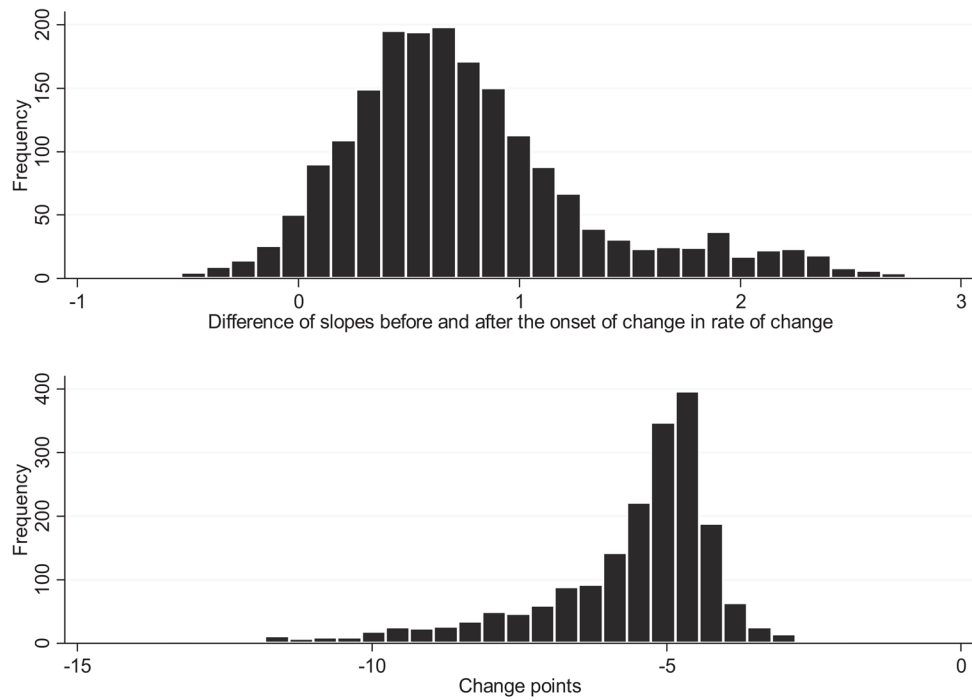
## Deviance Information Criterion

To compare the fit of the four models considered, we used the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002). The DIC is a generalization of Akaike's Information Criterion that combines information about the fit and complexity of a model. It is calculated as  $DIC = \overline{D(\theta)} + p_D$ , where  $\overline{D(\theta)} = E_{\theta|y}[D(\theta)]$  is the posterior mean deviance and  $p_D$  is the effective number of parameters.  $D(\theta)$ , the deviance of a model measures its fit. It is calculated as  $-2^* \loglikelihood$  of the model, standardized by a term that is a function of the data alone. Models with smaller DIC are considered to be better supported by the data.

## Sensitivity of Results to Prior Specification

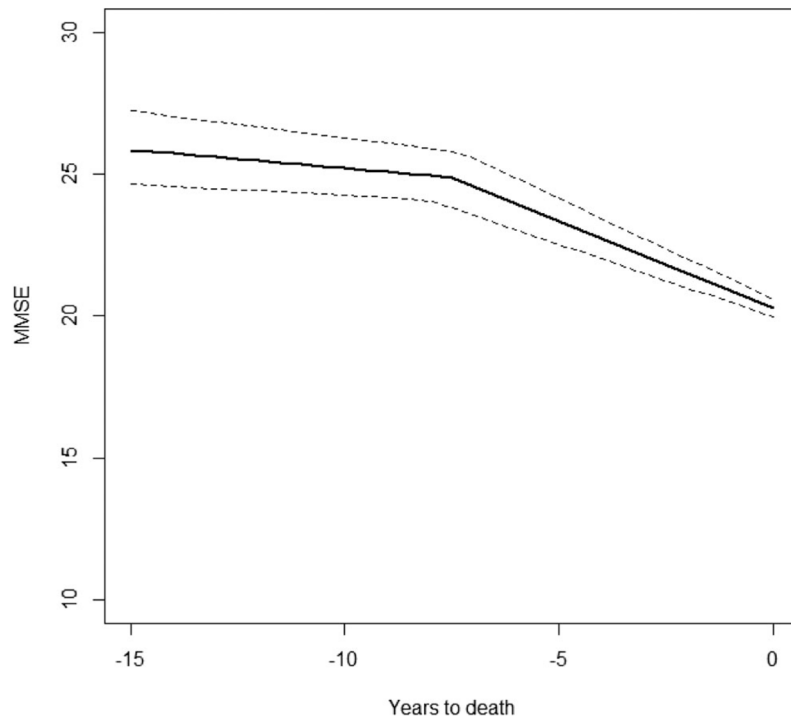
Robustness of results to the specification of the prior distributions was investigated and confirmed after a series of sensitivity analyses with different prior distributions. As the random change points are key parameters in the model, we considered a series of alternative distributions for the distribution of the hyper-parameter  $\mu_\gamma$ .

We considered  $\mu_\gamma \sim N(-7, 100)I(-20, 2)$  and  $\mu_\gamma \sim N(-3, 100)I(-20, 2)$ . In addition, we considered a uniform distribution for the hyper-parameter  $\mu_\gamma$ ; specifically, we modeled it as  $\mu_\gamma \sim Unif(-10, 2)$ . Model results were robust to the specification of other prior distributions.



**Figure 1.** Histogram of the difference of posterior means of individual slopes before ( $\alpha_{1i}$ ) and after the change point  $\alpha_{2i}$  and of the posterior mean estimates of individuals' change points  $\gamma_i$  (as estimated by the model).





**Figure 2.** Graphical representation of the estimated mean trajectory and its 95% Credible Interval for a physically and cognitively able man who left school aged 14 yrs and died aged 88 yrs.

**Table 1**

Mean and Standard Deviation of Cognitive Scores and Distance to Death of CC75C Study Participants and of the Subsample of Deceased Individuals Last Seen Less Than 10 Years Before Death

(Piccinin et al., 2011)	Whole sample			Last seen less than 10 years before death		
	N	MMSE	Mean (SD) Years to death	N	MMSE	Mean (SD) Years to death
Baseline	2,039	25.7 (4.0)	7.3 (5.1)	1,896	25.6 (4.1)	6.8 (4.9)
Two	1,106	24.5 (4.0)	6.7 (4.5)	1,088	24.5 (4.0)	6.6 (4.5)
Seven	630	22.9 (5.6)	5.3 (3.7)	628	22.9 (5.6)	5.3 (3.7)
Nine	354	22.9 (5.8)	4.7 (3.0)	353	22.9 (5.8)	4.7 (3.0)
Eleven	137	22.6 (5.4)	3.6 (2.4)	137	22.6 (5.4)	3.6 (2.4)
Thirteen	51	21.0 (6.8)	2.5 (1.6)	51	21.0 (6.9)	2.5 (1.6)
Seventeen	9	17.8 (6.4)	0.7 (0.4)	9	17.7 (6.4)	0.7 (0.4)

**Table 2**

Random Change Point Model: Posterior Mean, Standard Deviation and 95% Credible Interval of Fixed Effects and Standard Deviation of Residuals

	Estimate (SD)	95% Credible interval
Fixed effects		
Cognitive performance 2 years before death	23.9 (0.23)	[23.45, 24.37]
Rate of change before change point ( $\alpha_{22}$ )	-0.13 (0.02)	[-0.19, -0.08]
Age at death	0.01 (0.002)	[0.009, 0.01]
Physical impairment	-0.03 (0.02)	[-0.07, 0.007]
Gender	0.03 (0.02)	[-0.007, 0.07]
Education	-0.02 (0.003)	[-0.03, -0.01]
Cognitive impairment	0.73 (0.04)	[0.64, 0.82]
Rate of change after change point ( $\alpha_{22}$ )	-0.61 (0.05)	[-0.71, -0.51]
Age at death	-0.04 (0.004)	[-0.05, -0.03]
Physical impairment	-0.05 (0.05)	[-0.13, 0.04]
Gender	-0.17 (0.05)	[-0.26, -0.08]
Education	0.003 (0.009)	[0.01, 0.05]
Cognitive impairment	-0.48 (0.08)	[-0.66, -0.31]
Change point	-5.66 (0.21)	[-6.08, -5.24]
Random effects		
Intercept	1.03 (0.46)	[0.03, 0.95]
Slope before change point	0.03 (0.005)	[0.001, 0.03]
Slope after change point	0.36 (0.03)	[0.001, 0.44]
Change point	2.9 (0.1)	[2.6, 3.1]