Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
0(0) 1–13

DOI: 10.1177/0962280216645632
smm.sagepub.com
SAGE

Sigmoidal mixed models for longitudinal data

Ana W Capuano,[1] Robert S Wilson,[1] Sue E Leurgans,[1] Jeffrey D Dawson,[2] David A Bennett[1] and Donald Hedeker[3]

Abstract

Linear mixed models are widely used to analyze longitudinal cognitive data. Often, however, the trajectory of cognitive function is nonlinear. For example, some participants may experience cognitive decline that accelerates as death approaches. Polynomial regression and piecewise linear models are common approaches used to characterize nonlinear trajectories, although both have assumptions that may not correspond with the actual trajectories. An alternative is to use a flexible sigmoidal mixed model based on the logistic family of curves. We describe a general class of such a model, which has up to five parameters, representing (1) final level, (2) rate of decline, (3) midpoint of decline, (4) initial level before decline, and (5) asymmetry. Focusing on a four-parameter symmetric sub-class of the model, with random effects on two of the parameters, we demonstrate that a likelihood approach to fitting this model produces accurate estimates of mean levels across time, even in the case of model misspecification. We also illustrate the method on deceased participants who had completed at least 5 years of annual cognitive testing and annual assessment of body mass. We show that departures from a stable body can modify the trajectory curves and anticipate cognitive decline.

Keywords

Nonlinear models, longitudinal data, mixed models, cognitive decline, Alzheimer's disease, terminal decline

1 Introduction

With the aging of the population in developed countries, there is great interest in understanding how complex human behaviors change in old age. Longitudinal studies with repeated behavioral assessment over time have documented age-related changes in diverse domains, including cognitive function,[1] motor function,[2] sensory function,[3] and wellbeing.[4] However, trajectories of change are not always linear.[5] In particular, common neurodegenerative conditions[5] and impending death[4] are associated with acceleration in rates of behavioral change.

Researchers have tried different approaches to model nonlinear trajectories of behavioral change. Piecewise models[4] or mixed-effects change point models[6–9] are informative strategies and may provide a good model fit in some cases, but they are unlikely to capture the true trajectory. Mixed-effects change point models are often fitted in a Bayesian framework using Markov chain Monte Carlo methods that can be slow to converge, especially with several risk factors and multiple change points.[10,11] Smoothing splines are another approach, but they yield many parameters that may be difficult to interpret.[12] Polynomials can approximate nonlinear models but, with several parameters, providing a biologically based interpretation can be challenging.[13] Nonlinear mixed effect models are not typically used to analyze late-life changes in behavior with rare exceptions.[14] Without multiple covariates[15] or random effects, and with more control over the experimental time, these models are more commonly applied to studies in areas such as microbiology, pharmacokinetics,[16] and agriculture.[13] Here, we demonstrate a useful

[1]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA
[2]Department of Biostatistics, University of Iowa, Iowa City, IA, USA
[3]Department of Public Health Sciences, Biostatistics, The University of Chicago, Chicago, IL, USA

Corresponding author:
Ana W Capuano, Rush Alzheimer's Disease Center, Rush University Medical Center, 600 South Paulina, Suite 1023B, Chicago, IL 60612, USA.
Email: ana_capuano@rush.edu

nonlinear mixed model, the sigmoidal model. This model allows the inclusion of covariates that are related to parametric quantities such as the rate of decline and half decline (the decline midpoint), as well as early and final levels. In contrast to piecewise linear models, the model can potentially provide better year by year fit due to its smoothness and flexibility.

In this paper, we study late-life change in cognitive function using sigmoid mixed model structures based on logistic curves. We first describe a class of models with five main parameters, each of which may be a function of covariates and two of which have random effects. Focusing on a four-parameter subclass of the model in a single-group setting, we then use simulations to explore the statistical properties of our estimation technique over a variety of situations in which it might be used such as variable number of follow-ups, nonsmooth trajectories, and with normal and non-normal random errors. The technique is illustrated using data from two longitudinal cohort studies on deceased participants with at least 5 years of annual observations.

2 Sigmoid mixed model for longitudinal data

Let $i=1,\ldots, N$ denote subjects and $j=1,\ldots, n_i$ denote repeated observations per subject. Let t_{ij} denote the negative time between the jth observation and the terminal event (death). For example, if a subject has observations every other year, starting 9 years before death, then $j=1, 2, 3, 4, 5$, and $t_{ij}=-9, -7, -5, -3, -1$. Suppose that Y is a continuous normal outcome whose mean changes over time as a nonlinear function of, for simplicity, a single covariate X. Generally, a nonlinear mixed model[17] can be written in terms of a known nonlinear function g given by

$$y_{ij} = g(t_{ij}, \delta_i, \varphi) + e_{ij} \tag{1}$$

where the random errors $e_{ij} \sim N(0,\sigma^2)$, δ_i is a vector of person-specific parameters and φ is a vector of fixed parameters. Suppose that $\delta_i = (\theta_{1i}, \theta_{4i})$, $\varphi = (\theta_2, \theta_3, \theta_5)$, and we let the nonlinear trajectory of Y follow a five-parameter logistic structure,[12] such that

$$g\big(t_{ij}, \delta_i, \varphi\big) = \theta_{4i} - (\theta_{4i} - \theta_{1i})\frac{1}{\left(1 + (t_{ij}/\ \theta_2)^{\theta_3}\right)^{\theta_5}} \tag{2}$$

The parameter θ_{1i} represents the random intercept at $t_{ij}=0$. The parameter θ_{4i} represents the limiting random response before the onset of decline, since $\lim_{t_{ij}\to-\infty} y_{ij}$ equals θ_{4i}. Since time starts negative and progress to zero at the final event (e.g. death), θ_{1i} is the final level and θ_{4i} is the initial level before decline for person i. We can allow θ_{1i} to accommodate a fixed covariate and random effects by representing it by

$$\theta_{1i} = \alpha_1 + \beta_1 x + u_i \tag{3}$$

where $u_i \sim N(0,\sigma_u{}^2)$. Hence, in the cognitive decline example, if there is no covariate, α_1 represents the mean level of cognitive function proximate to death. Similarly, θ_{4i} can be represented by

$$\theta_{4i} = \alpha_4 + \beta_4 x + v_i \tag{4}$$

where $v_i \sim N(0,\sigma_v{}^2)$. Again, in the cognitive decline example, if there is no covariate, α_4 represents the mean level of cognitive function in early adulthood.

Note that equation (2) has two parts. The first part of the equation is the initial level θ_{4i}. The second part of the equation is the total decline $(\theta_{4i} - \theta_{1i})$ multiplied by a proportion (p) that is a nonlinear function of time $\left(p = \frac{1}{\left(1+(t_{ij}/\theta_2)^{\theta_3}\right)^{\theta_5}}\right)$. The interpretation of θ_2 and θ_3 are influenced by θ_5. The model can be simplified. If θ_5 is always one, we will have a four-parameter logistic. If, in addition, θ_{4i} is set to zero, we will have a three-parameter logistic. We provide interpretation of θ_2 and θ_3 for the three-parameters and four-parameters logistic models represented by the special case where θ_5 equals 1 $\left(p = \frac{1}{1+(t_{ij}/\theta_2)^{\theta_3}}\right)$. Later we discuss the more complex case when θ_5 does not equal 1.

Figure 1 illustrates the interpretation of the model based on four parameters. At each point in time a nonlinear proportion p of the total decline $\theta_{4i} - \theta_{1i}$ will be reduced from the initial level θ_{4i} until final level θ_{1i} was reached.

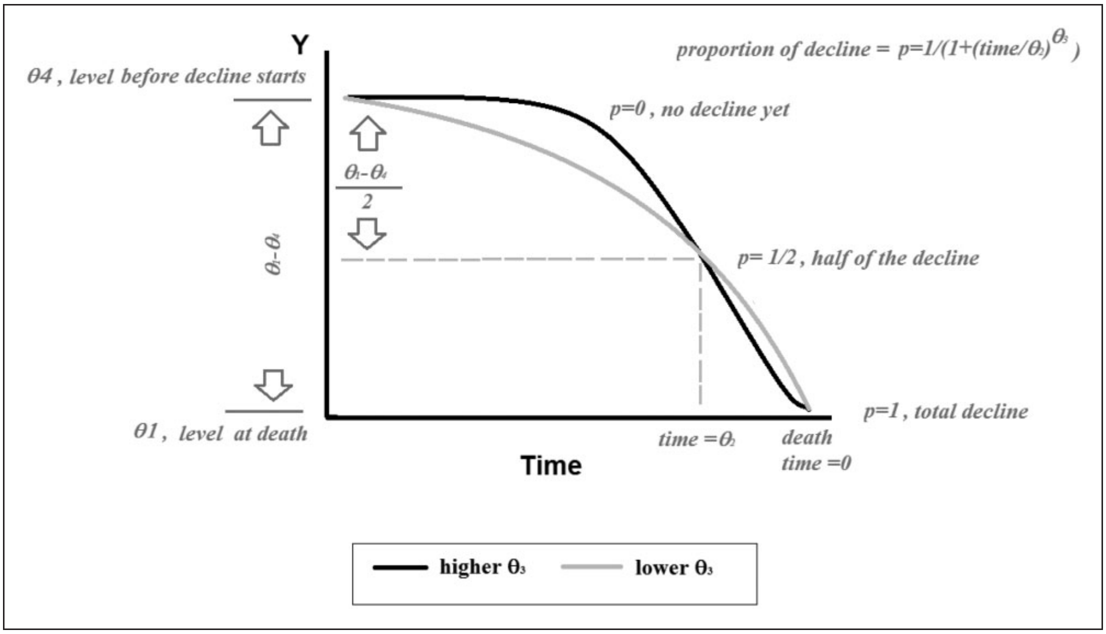

Figure 1. Interpretation of the sigmoidal model based on the four-parameter logistic.

The parameter θ_2 represents the point in time where half of the total decline occurred, i.e. the midpoint of decline. That is, when time equals θ_2, then $t_{ij}/\ \theta_2$ equals one and p equals 0.5. It can accommodate a fixed covariate effect by representing it by

$$\theta_2 = \alpha_2 + \beta_2 x \tag{5}$$

The parameter space of θ_2 must be negative to allow t_{ij}/θ_2 to be positive, since t_{ij} starts at a negative number and increases towards zero.

The parameter θ_3 represents the rate of decline around the time θ_2. If $t_{ij}/\ \theta_2$ is above 1 (time before the half decline), high values of θ_3 will decrease p. For example, if $t_{ij}/\ \theta_2 = 1.25$, then a θ_3 of 1 will yield a p of $1/(1+1.25)$, while a θ_3 of 2 will yield a smaller p of $1/(1+1.25^2)$. Conversely, if $t_{ij}/\ \theta_2$ is less than 1 (time after the half decline, closer to death), high values of θ_3 will increase p. For example, if $t_{ij}/\ \theta_2 = 0.25$, then a θ_3 of 1 will yield a p of $1/(1+0.25)$, while a θ_3 of 2 will yield a larger p of $1/(1+0.25^2)$. The parameter θ_3 can be represented by

$$\theta_3 = \alpha_3 + \beta_3 x \tag{6}$$

The rate of decline θ_3 is a scale modification over time of the person specific curves, and influences the nonlinearity, or the steepness, of the curve. A positive effect on the rate of decline represents a decline that starts earlier versus a curve that starts flat and declines later. For three parameters logistic models where θ_{4i} is always zero the interpretation of the parameters θ_{1i}, θ_2, and θ_3 is the same as described above. Note that these three shape/scale parameters represented by θ_2 and θ_3 do not have random effects. Conceptually, random effects for scale parameters could be added to the model, but estimation may be impractical.

An additional rate/scale parameter can be added to allow for further flexibility. The parameter θ_5 can be represented by

$$\theta_5 = \alpha_5 + \beta_5 x \tag{7}$$

To understand θ_5, let's revisit the proportion p that now has the general form $\frac{1}{\left(1+(t_{ij}/\theta_2)^{\theta_3}\right)^{\theta_5}}$. The parameter θ_5, termed in the literature as "asymmetry" parameter,[12] will modify the curve given by the combination of θ_2 and θ_3, so that high values of the asymmetry θ_5 will postpone decline (including the half decline). The curve will look more asymmetric with a sharper decline close to death. While in the three- and four-parameter logistic θ_2 represents the

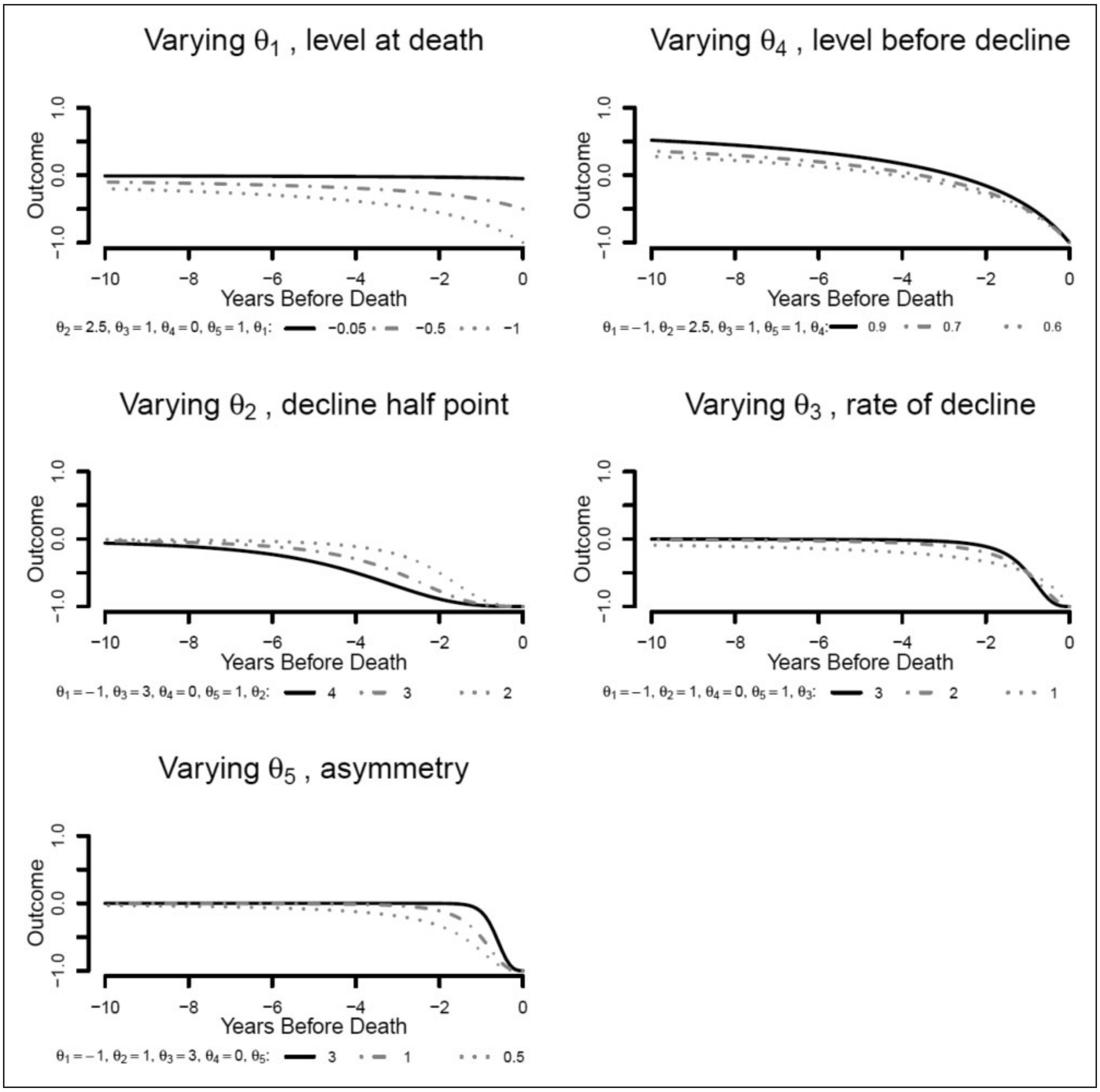

Figure 2. Theoretical illustration of possible trajectories of the sigmoidal model for three choices of each parameter.

half decline, in the five-parameter model that is no longer the case. When time equals θ_2, p will be $\frac{1}{2^{\theta_5}}$. For example, at time equals to θ_2, for a θ_5 equals to 1 the value of p will be ½, while for a θ_5 equals to 2 the value of p will still be only ¼. Figure 2 illustrates how changes in each parameter may change the trajectory.

In many cases, four parameters may provide adequate parsimony, fit and interpretation. A likelihood ratio test can be used to assess the need of the additional parameter in the five-parameter structure. Due to preliminary results from our dataset, we focus this work on the four-parameter logistic model. As described above, we assume that the between person random effects (u_i and v_i), as well as the within-person residuals (e_{ij}), follow normal distributions, where u_i and v_i may be correlated with each other but are assumed to be independent of the e_{ij}.

The model can be fit using SAS PROC NLMIXED. The model is fit by finding the maximum likelihood integrated over the random effects using adaptive Gaussian quadrature and quasi-Newton optimization. The initial parameter values required for the PROC NLMIXED can be obtained by examining the data trajectories. For example, an initial value for θ_1 can be the average of the observations close to death. Because we are describing a declining process, the initial for θ_4 should be smaller than the initial of θ_1. The linear slope close to death (i.e. from death to 3 years before death) versus the earlier linear slope (i.e. more than 3 years before death) can help define θ_2 and θ_3. If the ratio of the slopes is close to 1, we should use higher initial values for θ_2 and θ_3. In our experience, it is helpful to start with the empty model before adding covariates. We provide sample code for find the initial values for the empty model in the supplementary material, Appendix A. In Appendix A, we also provide sample code for fitting sigmoidal models with covariates. The estimates of the empty model can be used for the model with covariates.

3 Simulations

The models have previously been used in laboratory work where the curves are smooth and the number of observations is controlled. Epidemiological studies often have a varying number of follow-ups and trajectories are not necessarily smooth. The properties of the four-parameter longitudinal sigmoidal model were assessed via simulation over a range of situations. Parameters were selected based on preliminary data from studies described in section 4. A total of 1000 datasets were randomly generated under each parametric setting, with each dataset containing a sample of 100 participants and 20 years of annual follow-up (or 21 observations) per participant corresponding to, e.g. −0.01, −1.01,..., −20.01 years before death. We generated data assuming smooth nonlinear trajectories, linear trajectories, and nonsmooth trajectories with abrupt changes (piecewise linear trajectories). In addition, for a more challenging "wrong model" simulation, we generated data assuming a nonsmooth trajectory with non-normal random effects.

Data were first generated using the four-parameter logistic with parameter (equation (1) with $\theta_5=1$, and equations (2) to (7)) values selected to reproduce what was observed during application of the method to real data. We use two sets of parameters: data of type A were generated from $\alpha_1=-0.6111$, $\alpha_2=-1.8259$, $\alpha_3=0.7418$, $\alpha_4=0.5293$; data of type B were generated from $\alpha_1=-1.06544$, $\alpha_2=-1.9992$, $\alpha_3=1.10456$, $\alpha_4=0.82006$. Data A and B assume error variance 0.068 ($\sigma=0.2610$), variance of u equals 3.018 ($\sigma_u=1.7374$), variance of v equals 0.035 ($\sigma_v=0.1870$), and a correlation of zero between random effects.

To assess robustness, we simulated modeling under misspecification. Given that the four-parameter logistic characterizes curves that are nonlinear and smooth, we considered declines that are linear as well as declines that are not smooth. First, we considered an underlying linear decline (data generated assuming a linear trajectory). Data of type C were generated from one linear slope of −0.15 and intercept at death of −1.8. Second, we considered one or multiple abrupt changes to the underlying linear decline. The equations for the two abrupt changes (two-change-point linear decline) to the underlying linear decline were

$$y_{ij}=\begin{cases} b_{1i}+b_{2i}t_{ij}+e_{ij} & \textit{if } t_{ij}\le b_{3i} \\ b_{1i}+b_{2i}b_{3i}+b_{4i}(t_{ij}-b_{3i})+e_{ij} & \textit{if } b_{3i}<t_{ij}\le b_{5i} \\ b_{1i}+b_{2i}b_{3i}+b_{4i}(b_{5i}-b_{3i})+b_{6i}(t_{ij}-b_{5i})+e_{ij} & \textit{if } t_{ij}>b_{5i} \end{cases} \tag{8}$$

where $i=1,\ldots, N$ subjects, $j=1,\ldots, n_i$ repeated observations per subject and t_{ij} time between the jth observation and the terminal event. The parameters represent an intercept at the terminal event (b_{1i}), a terminal slope (b_{2i}), a change point from pre-terminal to terminal (b_{3i}), a pre-terminal slope (b_{4i}), a change point from early to pre-terminal (b_{5i}), and an early slope (b_{6i}). Note that the first and second lines (without b_{5i}) of equation (8) represent the equations used to generate the data for one abrupt change. The parameters b_{li}, where $l=1,\ldots, 6$ number of parameters, were randomly generated assuming a normal distribution such that

$$b_{li} \sim Normal(b_l, \sigma_l^2) \tag{9}$$

Data of type D were generated from an underlying one-change-point linear decline with the parameters: pre-terminal slope ($b_4=-0.1$, $\sigma_4=0.01$), change point from pre-terminal to terminal ($b_3=-4$, $\sigma_3=0.3$), terminal slope ($b_2=-0.4$, $\sigma_2=0.04$), and intercept at death ($b_1=-1.8$, $\sigma_1=0.14$). Data E were generated from an underlying two-change-point linear decline with the parameters: early slope ($b_6=-0.005$, $\sigma_6=0.01$), change point from early to pre-terminal ($b_5=-8$, $\sigma_5=0.3$), pre-terminal slope ($b_4=-0.1$, $\sigma_4=0.01$), change point from pre-terminal to terminal ($b_3=-4$, $\sigma_3=0.3$), terminal slope ($b_2=-0.4$, $\sigma_2=0.04$), and intercept at death ($b_1=-1.8$, $\sigma_1=0.14$).

Robustness under non-normal random effects was also evaluated. For an additional challenge, we assumed a nonsmooth decline. Data F were generated similarly to data E but assuming a departure from normality for the random effects. This was done by generating random errors ω following a gamma distribution with a shape parameter of 9 and a scale parameter of 0.5, and centering the distribution at its median (about 4.355). This introduced skewness in the random error distribution of about 0.7, such that

$$b_{li} = b_l + \omega\sigma_l \tag{10}$$

Using the data generated with an underlying four-parameter logistic we studied the effect of the number of observations per participant on the characterization of the curve using the sigmoidal mixed model. To be able to include the period of accelerated decline that occurs around 3 years before death (known as terminal decline), data of type A were also generated with 5, 7, and 10 years of annual follow-up before death. The sigmoidal mixed model was fit with the SAS PROC NLMIXED. Parameter bias and accuracy were evaluated. In addition for data of types A and B, the linear mixed model with a quadratic term for time before death was fit with the SAS PROC MIXED. Bias and accuracy of the estimated longitudinal mean were evaluated year by year, as these simulations were performed under misspecification (that is the true parameters used for data generation are not estimated).

Using the data generated from underlying linear or change-point linear declines (data C, D, and E), we assessed the sigmoidal mixed model robustness under misspecification. The sigmoidal mixed model was fit with the SAS PROC NLMIXED as described above. Bias and accuracy of the estimated longitudinal mean were evaluated year by year, as these simulations were performed under misspecification.

Parameters from the sigmoidal mixed model were adequately estimated using as few as 5 years of annual follow-up per participants (Supplementary Material Appendix B). All simulated parameters had coverage of above 93% and very low bias. The mean square error of the estimated parameters remained less than 0.035 with the decreasing number of observations.

Table 1 shows the results from simulations of sigmoidal mixed modeling and linear mixed modeling with a quadratic term using data A and B. Estimated longitudinal annual means had more than 10 times higher bias and mean square errors using linear models with a quadratic term compared to the sigmoidal mixed models.

Table 2 shows the analysis of robustness under misspecification. Underlying linear decline (data type C) generated a very large estimated decline midpoint, θ_2. Overall, the longitudinal mean estimated year by year using the sigmoidal mixed model was close to the true underlying longitudinal mean. Bias was low overall but relatively higher in estimations close to death.

Table 1. Simulation results of sigmoidal mixed modeling and linear mixed modeling with quadratic term.

		Simulation data A[a]					Simulation data B[b]			
		Sigmoidal model		Linear with quadratic term			Sigmoidal model		Linear with quadratic term	
Years before death	μ†	Bias	MSE	Bias	MSE	μ†	Bias	MSE	Bias	MSE
0	−0.61	−0.00	0.03	0.31	0.11	−1.07	−0.00	0.03	0.45	0.22
−1	−0.17	−0.00	0.01	−0.05	0.02	−0.47	−0.00	0.01	0.03	0.01
−2	−0.02	−0.00	0.01	−0.10	0.02	−0.12	−0.00	0.01	−0.13	0.03
−3	0.06	−0.00	0.01	−0.10	0.02	0.09	−0.00	0.01	−0.16	0.03
−4	0.12	−0.00	0.00	−0.06	0.01	0.22	−0.00	0.00	−0.11	0.02
−5	0.16	−0.00	0.00	−0.01	0.00	0.32	−0.00	0.00	−0.03	0.00
−6	0.20	−0.00	0.00	0.04	0.00	0.39	−0.00	0.00	0.08	0.01
−7	0.22	−0.00	0.00	0.11	0.01	0.44	−0.00	0.00	0.21	0.04
−8	0.24	−0.00	0.00	0.18	0.03	0.48	−0.00	0.00	0.35	0.12
−9	0.26	−0.00	0.00	0.25	0.06	0.52	−0.00	0.00	0.49	0.25
−10	0.28	−0.00	0.00	0.32	0.11	0.55	−0.00	0.00	0.65	0.42
−11	0.29	−0.00	0.00	0.40	0.16	0.57	−0.00	0.00	0.81	0.65
−12	0.30	−0.00	0.00	0.48	0.23	0.59	−0.00	0.00	0.97	0.94
−13	0.31	−0.00	0.00	0.56	0.32	0.61	−0.00	0.00	1.13	1.29
−14	0.32	−0.00	0.00	0.64	0.42	0.62	−0.00	0.00	1.30	1.70
−15	0.33	−0.00	0.00	0.72	0.53	0.64	−0.00	0.00	1.47	2.16
−16	0.34	−0.00	0.00	0.81	0.66	0.65	−0.00	0.00	1.64	2.70
−17	0.35	−0.00	0.00	0.89	0.81	0.66	−0.00	0.00	1.81	3.29
−18	0.35	−0.00	0.00	0.97	0.97	0.67	−0.00	0.00	1.98	3.95
−19	0.36	−0.00	0.00	1.06	1.14	0.68	−0.00	0.00	2.15	4.67
−20	0.36	−0.00	0.00	1.14	1.33	0.68	−0.00	0.00	2.33	5.46

MSE: mean square error.
[a] $\alpha_1 = -0.6111$, $\alpha_4 = 0.5293$, $\alpha_3 = 0.7418$, $\alpha_2 = -1.8259$, $\sigma_{uv} = 0$, $\sigma = 0.2610$, $\sigma_u = 1.7374$, $\sigma_v = 0.1870$.
[b] $\alpha_1 = -1.06544$, $\alpha_4 = 0.82006$, $\alpha_3 = 1.10456$, $\alpha_2 = -1.9992$, $\sigma_{uv} = 0$, $\sigma = 0.2610$, $\sigma_u = 1.7374$, $\sigma_v = 0.1870$.
μ= true mean.

Table 2. Simulation results of sigmoidal model for misspecification underlying processes.

							Nonsmooth piecewise linear data					
	Linear Data C			Data D[a]			Data E[b]			Data F[c]		
Years before death	True mean	Bias	MSE	True mean	Bias	MSE	True mean	Bias	MSE	True mean	Bias	MSE
0	−1.80	0.20	0.15	−1.80	0.55	0.31	−1.80	0.31	0.10	−1.80	0.34	0.11
−1	−1.65	0.20	0.14	−1.40	0.49	0.24	−1.40	0.43	0.18	−1.40	0.46	0.21
−2	−1.50	0.20	0.13	−1.00	0.35	0.12	−1.00	0.45	0.21	−1.00	0.49	0.24
−3	−1.35	0.19	0.12	−0.60	0.17	0.03	−0.60	0.33	0.11	−0.60	0.37	0.14
−4	−1.20	0.19	0.11	−0.20	−0.04	0.00	−0.20	0.10	0.01	−0.20	0.15	0.02
−5	−1.05	0.19	0.10	−0.10	0.03	0.00	−0.10	0.11	0.01	−0.10	0.16	0.03
−6	−0.90	0.18	0.09	0.00	0.08	0.01	0.00	0.08	0.01	0.00	0.14	0.02
−7	−0.75	0.18	0.08	0.10	0.12	0.01	0.10	0.03	0.00	0.10	0.09	0.01
−8	−0.60	0.17	0.08	0.20	0.14	0.02	0.20	−0.04	0.00	0.20	0.03	0.00
−9	−0.45	0.17	0.07	0.30	0.16	0.03	0.21	−0.02	0.00	0.21	0.05	0.00
−10	−0.30	0.16	0.06	0.40	0.17	0.03	0.21	0.00	0.00	0.21	0.07	0.01
−11	−0.15	0.16	0.06	0.50	0.17	0.03	0.22	0.01	0.00	0.22	0.08	0.01
−12	0.00	0.15	0.06	0.60	0.16	0.03	0.22	0.01	0.00	0.22	0.09	0.01
−13	0.15	0.15	0.05	0.70	0.15	0.02	0.23	0.02	0.00	0.23	0.09	0.01
−14	0.30	0.14	0.05	0.80	0.14	0.02	0.23	0.02	0.00	0.23	0.10	0.01
−15	0.45	0.13	0.05	0.90	0.12	0.01	0.24	0.02	0.00	0.24	0.10	0.01
−16	0.60	0.12	0.05	1.00	0.09	0.01	0.24	0.02	0.00	0.24	0.10	0.01
−17	0.75	0.11	0.05	1.10	0.07	0.01	0.25	0.02	0.00	0.25	0.10	0.01
−18	0.90	0.10	0.06	1.20	0.04	0.00	0.25	0.02	0.00	0.25	0.10	0.01
−19	1.05	0.09	0.06	1.30	0.00	0.00	0.26	0.02	0.00	0.26	0.10	0.01
−20	1.20	0.08	0.07	1.40	−0.03	0.00	0.26	0.01	0.00	0.26	0.10	0.01

[a]Piecewise linear with one point change and normal random effects.
[b]Piecewise linear with two points of change and normal random effects.
[c]Piecewise linear with two points of change and non-normal random effects.

Using a sigmoidal model was also relatively insensitive to data with nonsmooth declines and with non-normal random effects. Overall, the longitudinal means estimated year by year using the sigmoidal mixed model were still close to the true underlying longitudinal mean. However, there was a noticeable increase in bias close to death and in the earlier years. We also observed a small increase in the mean square error with moderate departures from normality of the random effects (data F versus data E).

4 Analyses of late-life body mass index and cognitive decline

Analyses of body mass index (BMI) and late-life cognitive decline showed mixed results. Some studies show that low or decreasing BMI is associated with cognitive decline, but others show no association.[18–20] Few of these studies, however, followed both cognition and BMI for many years until death. Annual assessment of cognition, weight, and height from the Rush Religious Order Study and the Rush Memory and Aging Project were used to model change in global cognition and BMI. These analyses indicate that departures from stable body mass are associated with earlier cognitive decline.

The Religious Orders Study began in 1994. Participants are older Catholic nuns, priests, and monks recruited from more than 40 groups across the United States.[21] The Rush Memory and Aging Project began in 1997. It involves older lay persons recruited from retirement communities, subsidized housing facilities, churches, and social service agencies in the Chicago metropolitan area.[22] Studies were approved by the institutional review board of Rush University Medical Center. Details on the individual tests and the derivation of the composite measure of global cognition have been previously published.[6,23,24] Years of formal education is obtained at baseline. For each individual, BMI was calculated as weight at visit in kilograms over person-specific median height in meters squared (kg/m^2). Range of person-specific BMI was calculated as the average of the two largest values minus the average of the two lowest values

Data from 512 deceased participants with 5 to 20 years of annual cognitive testing (4982 observations) were included in the analyses. Among these individuals, 436 had BMI information close to death, obtained in the same visit of the last cognitive evaluation or visit before that. A random sample of person-specific observed BMI and global cognition is available as Supplementary Material (Appendix C). These 436 participants had a mean of 78.6 years of age at baseline ($sd = 6.5$), 88.4 years of age at death ($sd = 6.2$), and 16.7 years of formal education ($sd = 3.6$). They had a mean BMI of 27 at baseline and 25 at death. The median range of person-specific BMI was 3.1 (min = 0 and max = 14.8).

In the preliminary analyses, we compared the sigmoidal model with a more established method, the mixed-effects change point model.[10] The change point model is a piecewise linear model such that

$$y_{ij} = \begin{cases} b_{1i} + b_{2i}t_{ij} + e_{ij} & \text{if } t_{ij} \leq b_{3i} \\ b_{1i} + b_{2i}b_{3i} + b_{4i}(t_{ij} - b_{3i}) + e_{ij} & \text{if } t_{ij} > b_{3i} \end{cases} \tag{11}$$

where $i = 1, \ldots, N$ subjects, $j = 1, \ldots, n_i$ repeated observations per subject and t_{ij} time between the jth observation and the terminal event (death). The random effects for cognition at death (b_{1i}), terminal slope (b_{2i}), change point (b_{3i}), and pre-terminal slope (b_{4i}) are assumed to be multivariate normally distributed and allowed to be correlated. We modeled the nonlinear trajectory of global cognition as a function of education only (data not shown). To fit the mixed-effects change point model, we used a Bayesian Monte Carlo Markov Chain Approach.[11,25] Modeling was performed in Openbugs[26] using 10,000 iterations, 3000 burn-ins and 100 as a thinning parameter. Figure 3, which shows person-specific observed and predicted trajectories for a random sample of 15 cases, suggests good agreement between the two modeling approaches.

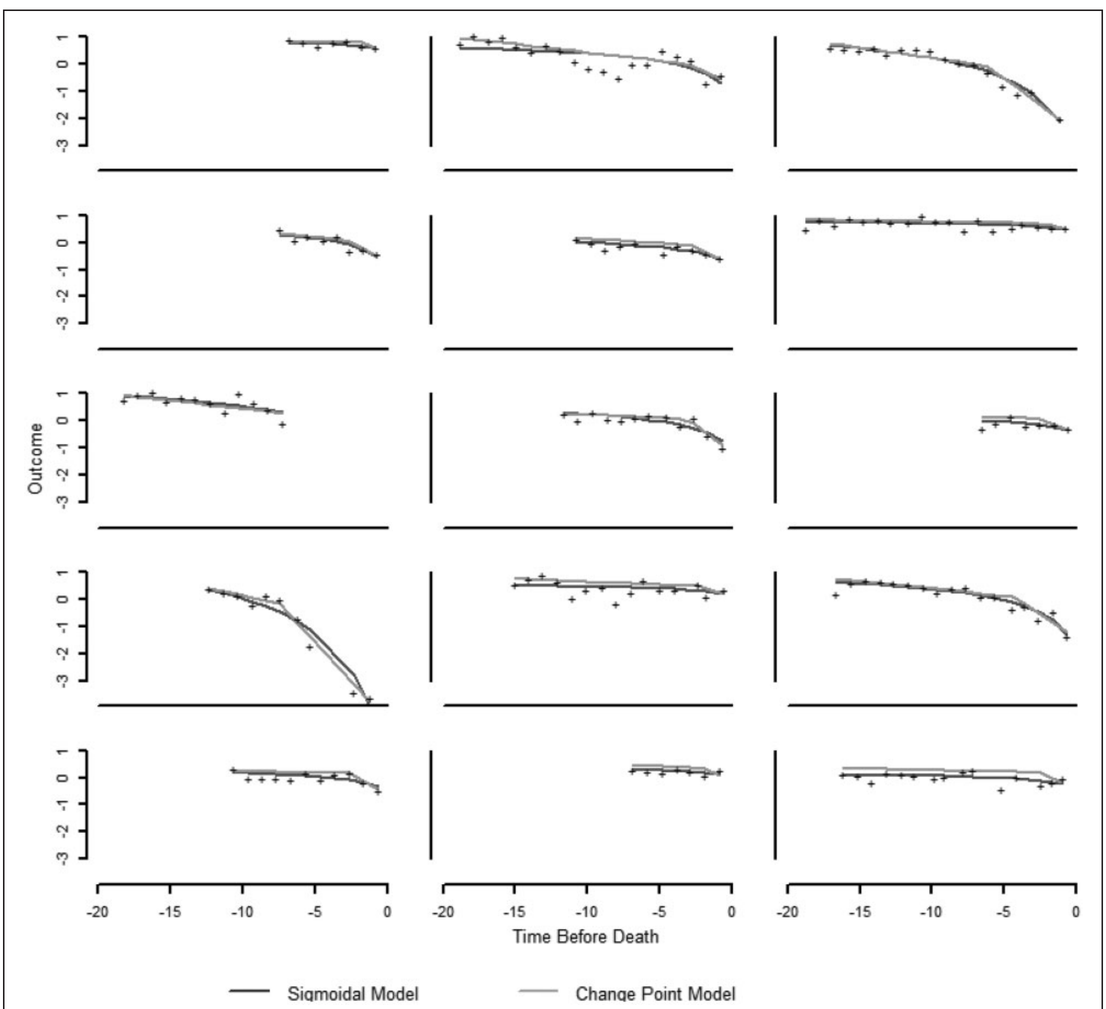

Figure 3. A random sample of person-specific observed and predicted global cognition.

In the main analyses, we modeled the nonlinear trajectory of global cognition as a function of age, BMI and education. Age and education were centered close to their means to improve interpretation of results. It is important to adjust for the fact that participants that enter the study at earlier age should have higher initial cognition, as well as participants that die younger may end up with a higher cognition. For this reason, we entered a term for age at baseline for early levels of cognition (θ_4) and a term for age at death for level of cognition at death (θ_1). The initial level of global cognition (θ_4) also included terms for education and BMI at baseline. Conversely, the level of global cognition at death (θ_1) also included terms for education and the last BMI available before death. The decline midpoint (θ_2) and the rate of decline (θ_3) included terms for education, age at baseline and death, and range of person-specific BMI. Modeling was performed in SAS using the PROC NLMIXED.

Table 3 presents results from the main model. Consider a typical person with 17 years of formal education that, 79 years of age at baseline and 89 years of age at death, and constant BMI of 26. Initial (α_4) and final (α_1) level of cognition were estimated to be 0.59 and −0.90, respectively. Half of the decline in global cognition (α_2) was estimated to occur 1.4 years before death. That is, half of the decline was estimated to occur during the years before the last 1.4 years of life, and half of the decline was estimated to occur the last 1.4 years of life. That is because the nonlinear curve is much steeper proximate to death. Education was only related to the initial (θ_4) global cognitive score, with each additional year of formal education increasing initial (θ_4) global cognitive score by 0.018. Age at baseline was related to initial (θ_4) but not final (θ_1) level of cognition. Age at death was related to final (θ_1) but not initial level (θ_4) of cognition. Participants who were younger at baseline had a higher initial (θ_4) level of global cognition. Participants who were younger at death had a higher final (θ_1) level of global cognition. The effect of age at death in the midpoint of decline (θ_2) was larger than the effect of age at baseline.

Range of person-specific BMI over time was significantly related to cognitive decline. BMI at baseline and the last BMI available before death did not have a significant effect on either initial (θ_4) or final (θ_1) level of global

Table 3. Results from sigmoidal modeling on global cognition.

Parameters	Estimate	Standard error	Pr > |t|
Initial level			
α_4	0.587	0.057	<0.0001
Education	0.018	0.009	0.045
Age at baseline	−0.016	0.005	0.000
BMI at baseline	0.002	0.005	0.725
Decline midpoint			
α_2	−1.402	0.172	<0.0001
Education	0.001	0.027	0.981
Age at baseline	0.169	0.029	<0.0001
Age at death	−0.216	0.025	<0.0001
Range of BMI	−0.265	0.048	<0.0001
Rate of decline			
α_3	0.830	0.091	<0.0001
Education	0.027	0.009	0.003
Age at baseline	−0.033	0.009	0.001
Age at death	0.067	0.010	<0.0001
Range of BMI	0.037	0.014	0.009
Level at death			
α_1	−0.902	0.092	<0.0001
Education	0.016	0.022	0.465
Age at death	−0.037	0.012	0.003
Last BMI before death	−0.016	0.014	0.237
Variance and covariance			
Error	0.246	0.003	<0.0001
U	2.120	0.227	<0.0001
V	0.174	0.034	<0.0001
u,v	−0.284	0.100	0.005

Education represents years of formal education centered at 17 years. Age at baseline and at death is centered respectively at 79 and 89 years. Body mass index (BMI) at baseline and the last before death were centered at 26.

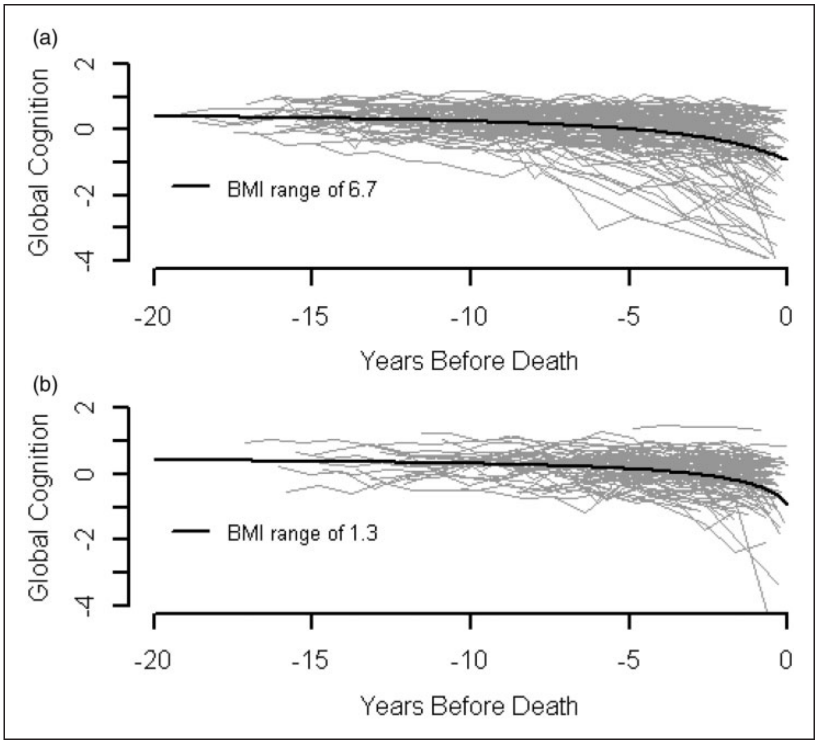

Figure 4. Spaghetti plot of global cognition among participants with upper quartile of range of person-specific body mass index (BMI) over time (plot A, BMI range $\geq$ 4.7) and lower quartile of range of person-specific BMI over time (plot B, BMI range $\leq$ 1.9), and predicted line based on sigmoidal model for a specific range of person-specific BMI (plot A, BMI range = 6.7; plot B, BMI range = 1.3).

cognition. However, for every additional 4 points of range of person-specific BMI, the midpoint of decline (θ_2) occurred about 1 year earlier (or -0.265 years for an increment of 1 in the range of BMI). That is, half of the decline in global cognition (θ_2) was estimated to occur 2.4 years before death for a typical person if range of person-specific BMI equalled 4. Figure 4 shows the predicted curves for a person with a range of person-specific BMI of 1.3 versus 6.7.

In secondary analyses, we estimated two additional models. In one model, the term for the last BMI before death was replaced by a term for range of person-specific BMI; in a second model, the terms for range of person-specific BMI and BMI proximate to death were replaced by terms for BMI at baseline. These analyses did not present additional significant terms (data not shown).

5 Discussion

Sigmoidal models based on the logistic family can accommodate longitudinal decline that includes a variety of monotone trajectories. For example, depending on risk and protective factors, some participants may never decline or may decline only shortly before death, while others may begin to decline early but decline at a gradual rate, and others may decline more precipitously but differ in when decline accelerates.

With the sigmoidal model we were able to identify risk/protective factors for initial levels of cognition and risk/protective factors for final levels of cognition. The model offers a possibility of including some correlation between u and v that is demonstrated in the example presented in this paper. Results from the example are consistent with previous research on education and terminal decline.[7,9] Person-specific BMI range is included as one of the covariates in the analysis. We could detect whether changes in BMI can modify the trajectory curves and anticipate decline. Here, we test change in body mass that is not necessarily a trend. Although decline in BMI is observed in general among the elderly, BMI range also captures drops in BMI even when followed by complete recovery. Most of the contribution of BMI range is on the midpoint of decline. In general the effects on rate of decline θ_3, although significant, were small.

The model showed adequate fit to the data. Comparing the person-specific predicted trajectories obtained with the sigmoidal model to the change-point model was, therefore, an important exercise. Although both models were informative and person-specific fit was in general very close, fitting the sigmoidal model took less time. While convergence with change-point models can take hours to days, depending on the complexity of the model, the sigmoidal model often takes minutes. We observed that the trajectories predicted by the sigmoidal model are sometimes less sensitive than the change-point model to the influence of "bad days" (some outlying observations below the overall trajectory) as shown in the person-specific plots. The models share a limitation that inference only reflects the observed period. For example, we cannot ascertain a change point or fully characterize a nonlinear curve when the period that precedes the terminal acceleration was not observed.

Model adequacy aside, formulating hypotheses for certain research questions using a change point model can be difficult. While the sigmoidal model with four parameters allows for risk factors to modify initial and/or final level (i.e. level of cognition earlier in life and level of cognition proximate to death), the single change-point model has only one parameter directly determining level (i.e. intercept can be at baseline or at death, but not both). For example we show that education is important for the initial level but not for the final level of global cognition. In addition multiple change points would be necessary for certain trajectories (e.g. Figure 1, lower θ_3). Hypotheses concerning the location of the half decline will be more mathematically tractable than hypotheses concerning a single change point.

Based on results from the simulations, we used 5 or more years of annual follow-up to characterize the nonlinear curve. In future work, it would be of interest to determine whether including very brief periods of follow-up for some individuals would improve or degrade the estimation. Another limitation is that the alignment of person-specific trajectories is fundamental. For example, in this work, all curves are aligned at death. Alignment at the diagnoses of dementia has being used in other studies that employed change-point methods.[10] Regardless of the biological event used, misalignment of trajectories may create a false idea that some are declining before others.

In addition to the simulations presented herein, we explored other situations. For example, we regenerated data A with a covariance between u and v of -0.2 instead of assuming no covariance. Coverage above 90% for all four parameters was obtained with 5 years of annual follow-up per participant and a sample of 500. We observed that the covariance of u and v has an impact on the coverage of θ_2 and θ_3. We also regenerated data A with increasing standard deviation σ_u of 3 and σ_v of 0.2 (variance of 9 and 0.04). Simulation was repeated using 5 years of annual follow-up per participant. The estimated parameters had coverage above 93%, bias below 0.02, the mean square error of 0.08 or below.

Simulations indicated that the sigmoidal model was robust under misspecification. The bias and mean square error were relatively low even for data with nonsmooth trajectories and moderate departure from normality of random effects. The model could fit linear trajectories but underlying linear declines tend to be translated into high θ_2. That is, if θ_2 is outside the range of the studied time, it is important to consider whether trajectories are linear.

A linear mixed model with a fixed quadratic term for time is a common approach for nonlinear trajectories. However, these models are more adequate for trajectories with a peak or a valley. Simulations showed much lower bias of the sigmoidal model compared to the linear mixed model with a fixed quadratic term for time. That is not surprising given that data was simulated following an underlying sigmoidal curve. However, if bias were small, there would be no justification for the additional complexity of the sigmoidal model.

We believe that these innovative yet practical models will help researchers in a wide array of fields where longitudinal trajectories are nonlinear. In particular, neurodegenerative diseases have a chronic course that involves very gradual changes in behavior over time, eventually resulting in rapid change in the last years of life. The current clinical understanding of the most common neurodegenerative disease, Alzheimer's disease, describes phases with different patterns of decline, each lasting many years, so that longitudinal studies with 5 to 10 years of annual follow-up may include people with multiple observation points in different phases.[27] Since the neurodegenerative changes progress over time, it is thought that the influences of environmental, genetic, and behavioral factors may differ across the time-course of the disease. Statistical models that accurately capture variation are valuable tools to identify factors that are more important early and others that are more important only later. Understanding how environmental and genetic factors influence these nonlinear trajectories may, therefore, be enhanced by using sigmoidal mixed models.

6 Supplementary materials

The SAS code and the web appendices referenced in section 3 and 4 are available in the online supplementary document.

Acknowledgments

The authors thank the participants of the Religious Orders Study, and Rush Memory and Aging Project, and the staff of the Rush Alzheimer's Disease Center.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the State of Illinois, and National Institutes of Health [grants P30AG10161, R01AG17917, R01AG15819, and R01AG34374].

References

1. Singh-Manoux A, Kivimaki M, Glymour MM, et al. Timing of onset of cognitive decline: Results from Whitehall II prospective cohort study. *BMJ* 2012; **344**: d7622.
2. Buracchio T, Dodge HH, Howieson D, et al. The trajectory of gait speed preceding mild cognitive impairment. *Arch Neurol* 2010; **67**: 980–986.
3. Lindenberger U and Ghisletta P. Cognitive and sensory declines in old age: Gauging the evidence for a common cause. *Psychol Aging* 2009; **24**: 1–16.
4. Gerstorf D, Ram N, Mayraz G, et al. Late-life decline in well-being across adulthood in Germany, the United Kingdom, and the United States: Something is seriously wrong at the end of life. *Psychol Aging* 2010; **25**: 477–485.
5. Amieva H, Le Goff M, Millet X, et al. Prodromal Alzheimer's disease: Successive emergence of the clinical symptoms. *Ann Neurol* 2008; **64**: 492–498.
6. Wilson R, Barnes L and Bennett D. Assessment of lifetime participation in cognitively stimulating activities. *J Clin Exp Neuropsychol* 2003; **25**: 634–642.
7. Batterham PJ, Mackinnon AJ and Christensen H. The effect of education on the onset and rate of terminal decline. *Psychol Aging* 2011; **26**: 339–350.
8. Wilson RS, Segawa E, Buchman AS, et al. Terminal decline in motor function. *Psychol Aging* 2012; **27**: 998–1007.
9. Terrera GM, Minett T, Brayne C, et al. Education associated with a delayed onset of terminal decline. *Age Ageing* 2014; **43**: 26–31.
10. Hall CB, Lipton RB, Sliwinski M, et al. A change point model for estimating the onset of cognitive decline in preclinical Alzheimer's disease. *Stat Med* 2000; **19**: 1555–1566.
11. Hall CB, Ying J, Kuo L, et al. Bayesian and profile likelihood change point methods for modeling cognitive function over time. *Comput Stat Data Anal* 2003; **42**: 91–109.
12. Gottschalk PG and Dunn JR. The five-parameter logistic: A characterization and comparison with the four-parameter logistic. *Anal Biochem* 2005; **343**: 54–65.
13. Schabenberger O and Pierce FJ. *Contemporary statistical models for the plant and soil sciences*. Boca Raton, FL: CRC Press, 2002.
14. Martins CA, Oulhaj A, de Jager CA, et al. APOE alleles predict the rate of cognitive decline in Alzheimer disease: A nonlinear model. *Neurology* 2005; **65**: 1888–1893.
15. Grimm KJ and Ram N. Non-linear growth models in M and SAS. *Struct Equ Model* 2009; **16**: 676–701.
16. Davidian M and Giltinan DM. *Nonlinear models for repeated measurement data*. London: Chapman and Hall, 1996.
17. Wu L. *Mixed effects models for complex data*. Boca Raton, FL: CRC Press, 2010.
18. Lo Giudice D, Smith K, Fenner S, et al. Incidence and predictors of cognitive impairment and dementia in Aboriginal Australians: A follow-up study of 5 years. *Alzheimers Dement* 2016; **2**: 252–261.
19. Kim HJ, Oh ES, Lee JH, et al. Relationship between changes of body mass index (BMI) and cognitive decline in Parkinson's disease (PD). *Arch Gerontol Geriatr* 2012; **55**: 70–72.
20. Aslan AK, Starr JM, Pattie A, et al. Cognitive consequences of overweight and obesity in the ninth decade of life? *Age Ageing* 2015; **44**: 59–65.
21. Bennett DA, Schneider JA, Arvanitakis Z, et al. Overview and findings from the religious orders study. *Curr Alzheimer Res* 2012; **9**: 628–645.
22. Bennett DA, Schneider JA, Buchman AS, et al. Overview and findings from the rush Memory and Aging Project. *Curr Alzheimer Res* 2012; **9**: 646–663.
23. Wilson RS, Beckett LA, Barnes LL, et al. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging* 2002; **17**: 179–193.

24. Wilson RS, Boyle PA, Yu L, et al. Conscientiousness, dementia related pathology, and trajectories of cognitive aging. *Psychol Aging* 2015; **30**: 74–82.
25. Laird NM and Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982; **38**: 963–974.
26. Lunn D, Spiegelhalter D, Thomas A, et al. The BUGS project: Evolution, critique and future directions. *Stat Med* 2009; **28**: 3049–3067.
27. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011; **7**: 280–292.