A latent model approach to define event onset time in the presence of measurement error

Peng Huang,
SKCCC Oncology Biostatistics Division, School of Medicine, Johns Hopkins University, 550 N. Broadway, STE 1103, Baltimore, MD 21205

Ming-Hui Chen, and
Department of Statistics, University of Connecticut, 215 Glenbrook Road, U-4120, Storrs, CT 06269-4120

Debajyoti Sinha
Department of Statistics, Florida State University, 117 N. Woodward Ave., Tallahassee, FL 32306

Peng Huang: phuang12@jhmi.edu

Abstract

For progressive diseases, it is often not so straightforward to define an onset time of certain disease condition due to disease fluctuation and clinical measurement variation. When a disease onset is claimed through the first presence of some clinical event which is subject to large measurement error, such onset time could be difficult to interpret if patients can often be seen to "recover" from the disease condition automatically. We generalize the traditional event onset time concept to control the recovery probability through the use of a stochastic process model. A simulation algorithm is provided to evaluate the recovery probability numerically. Bayesian latent residuals are developed for model assessment. This methodology is applied to define a new postural instability onset time measure using data from a Parkinson's disease clinical trial. We show that our latent model not only captures the essential clinical features of a postural instability process, but also outperforms independent probit model and random effects model. A table of estimated recovery probabilities is provided for patients under various baseline disease conditions. This table can help physicians to determine the new postural instability onset time when different thresholds of estimated recovery probability are used in clinical practice.

Keywords and phrases

Binary process; Brownian motion; Event onset time; Latent model; Probit model; Random effects model

1. INTRODUCTION

Many chronic diseases are progressive. For example, the progression of Parkinson's disease (PD) is caused by a continuous and irreversible brain neuronal cell death. The onset time to a milestone event, which is biologically difficult to recover without medical treatment, is often used by physicians to assess a patient's disease condition and to determine when certain intervention should be initiated. There are many such examples. The onset of a prespecified glomerular filtration rate for the end-stage renal disease is an indicator of kidney function failure and often requires the initiation of certain kidney treatment. In Parkinson's disease, the onset of postural instability is an indicator of impaired body balance and often requires special care to prevent patients from falling that can cause serious injury or death. If the presence or absence of an event is subject to large measurement errors, the

traditional onset time, defined as the first time when this event is seen, could become difficult to interpret and less valuable to physicians in medical decision making.

The onset of postural instability (PI) or impaired body balance and coordination in PD is one of the most serious indicators of disease progression that is associated with high risk of falling (resulting in fractured ankles, hips, shoulders, and skulls), dementia, and death (Kieburtz, 2003). Currently, PI status is primarily determined by the postural stability score which is evaluated through a *pull test*. The pull test is performed by pulling a patient's shoulders from behind briskly and scoring how well the patient could recover from falling. Although patients with PI are, in principle, difficult to recover to normal PI absence through PD medications (Bloem, 1992; Rogers, 1996), examiner's inadequate evaluation (even from a well trained examiner) makes more than 90% of the examinations incorrectly measured (Munhoz et al., 2004). Many factors contribute to measurement errors, for example, the examiner's force of pull and the patient's effort to recover body balance when pulled. As pointed out by Fahn (2006) and Hauser (2006), large measurement errors can still be seen even if examinations are performed by the same examiner, on the same patient, and at the same clinical visit. Large measurement errors lead to high false positive rate of PI presence, particularly for early PD patients. That is, patients without PI have a high chance to be misclassified as having PI. This is frequently seen in clinical trials and clinical practices. For example, in the multicenter controlled clinical trial of deprenyl and tocopherol antioxidative therapy of Parkinsonism (DATATOP, The Parkinson Study Group, 1989, 1990), more than 50% of treatment naive patients "recovered" to normal PI absence within 3 months of their previous PI presence without receiving any anti-PD medication. Although limitations of pull test were well documented in the literature, such a way of measuring PI onset time continues to serve as one of the most critical measures of PD disability because there is no other commonly accepted clinical measure that could replace it. As a consequence, an early PD patient not treated with PD medication can be seen to have frequently alternated PI presence and absence observations before the patient has become seriously paralyzed. This makes inference drawn from the traditional PI onset time difficult to interpret (Section 5).

We extend the traditional onset time measure to a more general setting when the presence of an event is subject to a high false positive rate. Let $\{Y(t), t \geq 0\}$ be a binary process taking value 0 or 1. For the above PD example, $Y(t) = 1$ (or $Y(t) = 0$) represents PI presence (or absence) at time t. Section 2 proposes a latent model that characterizes important clinical and biological features of longitudinal PI process. Whenever a $Y(t) = 1$ is observed at time $t = c$, we use the latent model to compute the recovery probability that an $Y(u) = 0$ is observed within time interval $u \in [c, c + d]$ for some clinically meaningful choice of interval width $d > 0$. The new PI onset time defined in Section 3 is the first time c such that the corresponding recovery probability is lower than some prespecified clinically meaningful threshold η. Theoretical properties of the new PI onset time are investigated. A simulation algorithm is provided to compute the recovery probability. Section 4 gives a model assessment method. Section 5 applies the proposed latent model and the new PI onset time measure to DATATOP trial data. A table of recovery probabilities for patients with different disease conditions is provided. Proofs of all relevant theorems are given in the Appendix.

2. A LATENT PROCESS MODEL

For a typical PI process $\{Y(t), t \geq 0\}$, the marginal variance of $Y(t)$ initially increases in time, representing a gradual increased variation in the function of body balance. Later, the variance decreases in time, representing the progression toward final paralyzation. A dynamic latent process model is a useful and popular tool to model a correlated binary process. Latent process models for independent binary response data proposed by Albert and Chib (1993, 1995), Chen, Dey, and Shao (1999) and others are not applicable to this

problem because they cannot account for within-subject association of $Y(t)$. Although Chen and Dey (1998) developed a unified approach to model correlated binary and ordinal response data using scale mixture of multivariate normal links, and Dunson (2003) proposed a dynamic model for multidimensional longitudinal data, their approaches essentially require observation times of $Y(t)$ to be evenly-spaced and equal for all patients. To ensure a good quality of the new PI onset time measure, this section presents a latent model that satisfactorily characterizes important clinical and biological features of PI process. Transitions between $Y(t) = 1$ status and $Y(t') = 0$ status are modeled longitudinally, accommodating within-subject association of the observed longitudinal process and potential confounding covariates.

Suppose a study has n subjects. Let $Y_i(t)$ be the observable binary PI status from subject i at time t measured by examiner δ_i, $\delta_i \in \{1, 2, \ldots, C\}$, $\mathcal{Y}_i(s) = \{Y_i(t), t \leq s\}$ be all available information up to time s from subject i ($i = 1, \ldots, n$), and $\mathcal{Y}(s) = (\mathcal{Y}_1(s), \ldots, \mathcal{Y}_n(s))'$. We consider the point process whose follow up times are independent of the treatment assignment and $\mathcal{Y}_i(t)$. The unobservable trajectory of disease biomarker $W_i(t)$ from the ith patient is modeled through a latent process

$$Y_i(t) = \begin{cases} 1 & \text{if } W_i(t) > 0, \\ 0 & \text{if } W_i(t) \leq 0, \end{cases}$$

(1)

where

$$W_i(t) = z_i'\beta_1 + h(t, \beta_2) + \sigma b_i + \tau b^{(\delta_i)} + \varepsilon_{i0} I[t=0] + \frac{1}{\sqrt{g(t)}} \varepsilon_i(t) I[t>0];$$

(2)

$b_i \sim iid\ N(0, 1)$; $b^{(\delta_i)} \sim iid\ N(0, 1)$; nonnegative continuous function $g(t)$ is used to capture the variation of the latent process; $\{\varepsilon_i\} = \{\varepsilon_i(t), t \geq 0\}$ is a standard Brownian motion with drift zero and origin $\varepsilon_i(0) \equiv 0$. Process $\{\varepsilon_i\}$, b_i, $b^{(\delta_i)}$, and ε_{i0} are independent. Since $var[W_i(t)] = \sigma^2 + \tau^2 + t/g(t)$ for $t > 0$ and $var[W_i(0)] = \sigma^2 + \tau^2 + 1$, it is natural to assume that $\lim_{t\to 0} t/g(t) = 1$ to ensure continuity of variance function at $t = 0$. Vector z_i contains all covariates from subject i. The $h(t, \beta_2)$ is a known continuous function of the follow up time t and unknown parameter β_2. Parameters β_1 and β_2 measure the fixed effects of the treatment and other risk factors on the patient-specific latent disease trajectory $W_i(t)$. Random variables b_i and $b^{(\delta_i)}$ capture patient-specific and examiner-specific random effects respectively. The indicator function $I[E]$ is defined as $I[E] = 1$ if event E is true and 0 otherwise. Random variable $\varepsilon_{i0} \sim N(0, 1)$ takes unit variance to ensure the identifiability of the latent process model. Variable ε_{i0} captures a possible lack of predictability of the baseline status $Y_i(0)$ for the status $Y_i(t)$ at time $t > 0$. The reason to include ε_{i0} is to adjust for the expected discontinuation of the latent disease trajectory at $t = 0$.

The motivation of this model is to capture the desired clinical and biological features of the postural instability process. First, the marginal expectation $E[W_i(t)|z_i] = h(t, \beta_2) + z_i'\beta_1$ depends on t only through $h(t, \beta_2)$. When the disease is progressive and the patients do not receive any treatment that could effectively stop the disease progression, $h(t, \beta_2)$ is a monotone increasing function in t. Second, since

$P\{Y_i(t) = 1 | \beta_1, \beta_2, \sigma, \tau, z_i\} = \Phi[\{z_i'\beta_1 + h(t, \beta_2)\} / \sqrt{\sigma^2 + \tau^2 + t/g(t)}]$, the marginal variation of $Y_i(t)$ is increasing in t when $z_i'\beta_1 + h(t, \beta_2) < 0$ and $h(t, \beta_2)$ is increasing in t. This is true for early PD

patients who are not on anti-PI treatment. However, after a patient has progressed to a certain advanced level such that $z_i'\beta_1 + h(t_{i0}, \beta_2) > 0$ for some $t_{i0} > 0$, the patient is more likely to have sustained $Y_i(t) = 1$ observations. This feature is further captured by the decreasing marginal variation of $Y_i(t)$ in t. Third, marginal probability $P\{Y_i(0) = 1\}$ is the same as the binary regression model with probit link. This property is important since the primary role of the latent Brownian process is to capture the dependency among repeated measures of $Y_i(t)$, and, at the same time, to preserve the marginal model of $Y_i(t)$ as a binary regression model under the probit link. Fourth, for any $0 \le t_1 < t_2$, correlation of two latent values from the same subject is

$$\mathrm{corr}\{(W_i(t_1), W_i(t_2))|b_i, b^{(\delta_i)}\} = \sqrt{\frac{t_1}{t_2}}.$$

It goes to zero as $(t_2 - t_1)$ increases to infinity while keeping t_1 fixed; and increases to 1 as both t_1 and t_2 go to infinity while keeping $(t_2 - t_1)$ fixed. This captures an essential PD feature that a PD patient will eventually completely lose the postural reflexes and becomes "frozen" if the PD progression was not stopped or reversed.

The proposed model does not require regularly spaced and common follow up visit times for all patients. Although $W_i(t)$ is Markovian, the resulting binary process $Y_i(t)$ is not. When the Markov assumption on the observable binary process $Y_i(t)$ is violated, our proposed latent process model is useful to model a dynamic binary process measured at irregular time points. In this sense, our proposed model is quite different from transition models discussed in Diggle, Liang, and Zeger (1994, Chapter 10).

Let $\mathbf{y}_{i0} = \{Y_i(t) : t = t_{i0} = 0, t_{i1}, \ldots, t_{im_i}\}$ be the collection of all observed PI values from patient i. Write $\mathbf{W} = (W_{i0}, \mathbf{W}_i', i = 1, 2, \ldots, n)'$ where $W_{i0} = W_i(0)$ and $\mathbf{W}_i = (W_i(t_{i1}), W_i(t_{i2}), \ldots, W_i(t_{im_i}))'$. Let $Y = (\mathbf{y}_{10}', \ldots, \mathbf{y}_{10}')'$, $\mathbf{z} = (z_1', z_2', \ldots, z_n')'$, and $T = (t_{ij}, j = 1, 2, \ldots, m_i, i = 1, 2, \ldots, n)$. We use $D = (\mathbf{Y}, \mathbf{z}, T, \mathbf{W})$ to denote the complete data and use $\theta = (\beta_1, \beta_2, \sigma^2, \tau^2)$ to denote the collection of all model parameters. The complete data likelihood function is given by

$$L(\theta|D) = \int \cdots \int \frac{1}{(2\pi)^{C/2}} \exp\left\{-\frac{1}{2}\sum_{k=1}^{C} b^{(k)2}\right\} \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}} \exp\left\{-\frac{1}{2}b_i^2\right\} \left[\prod_{j=0}^{m_i} \{Y_i(t_{ij})I[W_i(t_{ij}) > 0] + (1 - Y_i(t_{ij}))I[W_i(t_{ij}) \le 0]\}\right]$$

$$\times \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\{W_{i0} - z_i'\beta_1 - h(0, \beta_2) - \sigma b_i - \tau b^{(\delta_i)}\}^2\right]$$

$$\times \exp\left[-\frac{1}{2}\{W_i - (z_i'\beta_1 + \sigma b_i + \tau b^{(\delta_i)})J_{m_i} - H_i(\beta_2)\}' D_i V_i^{-1} D_i\{W_i - (z_i'\beta_1 + \sigma b_i + \tau b^{(\delta_i)})J_{m_i} - H_i(\beta_2)\}\right] \frac{|D_i||V_i|^{-1/2}}{(2\pi)^{m_i/2}} db_1 \ldots db_n db^{(1)}$$

where J_{m_i} is a $m_i \times 1$ vector whose components are all equal to 1, $t_{i0} = 0$, $H_i(\beta_2) = (h(t_{i1}, \beta_2), h(t_{i2}, \beta_2), \ldots, h(t_{im_i}), \beta_2))'$, $D_i = diag(\sqrt{g(t_{i1})}, \ldots, \sqrt{g(t_{im_i})})$, and

$$V_i = \begin{pmatrix} t_{i1} & t_{i1} & \cdots & t_{i1} \\ t_{i1} & t_{i2} & \cdots & t_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ t_{i1} & t_{i2} & \cdots & t_{im_i} \\ & \cdot & & \end{pmatrix}$$

(3)

The following theorem can be used to simplify expression

$\{W_i - (z_i'\beta_1 + \sigma b_i + \tau b^{(\delta_i)}) J_{m_i} - H_i(\beta_2)\}' D_i V_i^{-1} D_i \{W_i - (z_i'\beta_1 + \sigma b_i + \tau b^{(\delta_i)}) J_{m_i} - H_i(\beta_2)\}$ in likelihood function $L(\theta \mid D)$:

Theorem 1

For any vector $\mathbf{x} = (x_1, \ldots, x_m)'$, the following identity holds

$$\mathbf{x}' V_i^{-1} \mathbf{x} = \frac{x_1^2}{t_{i1}} + \sum_{j=2}^{m_i} \frac{(x_j - x_{j-1})^2}{t_{ij} - t_{i,j-1}}.$$

To estimate parameter θ, one can use a noninformative prior $\pi(\theta) = \pi_1(\beta_1) \pi_2(\beta_2) \pi_3(\sigma) \pi_4(\tau) \propto \pi_3(\sigma)\pi_4(\tau)$ for θ, where distributions of $\pi_3(\sigma)$ and $\pi_4(\tau)$ are obtained by truncating *Normal*$(0, \lambda^2)$ (for large λ) distributions to the range of $(0, +\infty)$. Given $\mathbf{W}$, b_i's, $b^{(k)}$'s, β_1, β_2, and τ, the conditional posterior distribution of σ is a truncated normal. Similarly, given $\mathbf{W}$, b_i's, $b^{(k)}$'s, β_1, β_2, and σ, the conditional posterior distribution of τ is also a truncated normal. Gibbs sampling algorithm can be used to sample $(\theta, \mathbf{W})$ from the joint posterior distribution $\pi(\theta, \mathbf{W}|D_{obs})$. The implementation of the Gibbs algorithm is straightforward due to the result established in Theorem 1 as well as the conjugate priors for σ and τ.

3. A NEW MEASURE OF PI ONSET TIME

Applying the latent process model described in Section 2, this section develops a new measure of PI onset time that is more consistent with characteristics of PD and also incorporates measurement errors. When a PI is observed at time c, we compute the conditional probability $P \{\min_{(c \leq t \leq c+d)} Y(t) = 0 \mid Y(c) = 1, \theta, \tau, z\}$, where d is a pre-specified time interval width, and $\theta = (\beta_1, \beta_2, \sigma^2, \tau^2)$. More specifically, for a chosen critical value η ($0 < \eta < 0.5$) and a pre-specified width d, we define the PI onset time of a patient with covariate vector z_i by $\inf\{c > 0, P [\min_{(c \leq t \leq c+d)} Y_i(t) = 0 \mid Y_i(c) = 1, \theta, Z_i] \leq \eta\}$. In other words, our new PI onset time is defined as the first time when a patient has little chance to recover from PI in a given length of time interval. Since PD is progressive and the loss of postural reflexes is difficult to recover, the new PI onset time measure is more robust than the traditional one in the sense that it not only reflects the steadily degenerative feature of PD, but also incorporates the variation from the examiner. The use of a threshold value η to the conditional probability $P \{\min_{(c \leq t \leq c+d)} Y(t) = 0 \mid Y(c) = 1, \theta, \tau, z\}$ makes the new PI onset time more clinically meaningful to physicians in defining patient's disease status.

Write $W_i(t)$ as $W_i(t) = a_{i0}(t) + \varepsilon_i(t)I[t>0]/\sqrt{g(t)} + \varepsilon_{i0}I[t=0]$, where $\{\varepsilon_i(t), t \geq 0\}$ is a standard Brownian motion with zero drift, $\varepsilon_i(0) = 0$, and $a_{i0}(t) = z_i'\beta_1 + h(t, \beta_2) + \sigma b_i + \tau b^{(\delta_i)}$. In the following discussion about the computation of $P \{\min_{(c \leq t \leq c+d)} Y_i(t) = 0 \mid Y_i(c) = 1, \theta, Z_i, b_i, b^{(\delta_i)}\}$, we focus on the computation for a fixed subject. We thus omit subscript i to simplify our notation and use b to denote $(b_i, b^{(\delta_i)})$. The density function of $W(c)$ at $W(c) = w$, given

(θ, Z, b), is $f(c, w) = \varphi(a_0(c) - w)$, where $\varphi(x) = \exp(-x^2/2)/\sqrt{2\pi}$ is the density function of the standard normal distribution. For $w > 0$, define continuous function

$a(t, c, w) = a_0(t+c)\sqrt{g(t+c)} - \{a_0(c) - w\}\sqrt{g(c)}$ and the stopping time $T = \inf\{t : t \geq 0, \varepsilon(t) = a(t, c, w)\}$. We are led to the following Theorem.

Theorem 2

Suppose $a(t, c, w)$ is continuous and left differentiable in $0 < t \leq d$, $P\{W(c) > 0 | \theta, b, Z\} = p_c$, and $f(c, w) = \varphi\{a_0(c) - w\}$. For the process $\{Y(t)\}$ defined above and two positive constants c and d, we have $P\{\min_{(c \leq t \leq c+d)} Y(t) = 0 | W(c) = w, \theta, b, Z\} = P\{T \leq d | \theta, b, z\}$. Furthermore,

$$
\begin{aligned}
P\{\min_{(c \leq t \leq c+d)} Y(t) \\
= 0 | Y(c) \\
= 1, \theta, Z, b\} &= \frac{1}{p_c} \int_0^{+\infty} f(c, w) dw \int_0^d \frac{1}{\sqrt{t}} \varphi\left(\frac{a(t, c, w)}{\sqrt{t}}\right) \\
&\times \lim_{s \uparrow t} \frac{1}{t - s} E[I\{T > s\}\{a(s, c, w) - \varepsilon(s)\} | \varepsilon(t) = a(t, c, w), \theta, b, Z] dt.
\end{aligned}
$$

This theorem shows that the new PI onset time can be computed through the integration of a weighted boundary cross probability of a standard Brownian motion to hit stopping boundary $a(t, c, w)$ in a given time interval. The computation can be quite involved. However, if we note that $P\{T \leq d | \theta, b, z\} = P\{\varepsilon(t)$ crossing over $a(t, c, w)$ in $0 \leq t \leq d | \theta, b, z\}$, we can use the following simulation algorithm to estimate probabilities $P\{\min_{(c \leq t \leq c+d)} Y(t) = 0 | W(c) = w, \theta, Z, b\}$ and $P\{\min_{(c \leq t \leq c+d)} Y(t) = 0 | Y(c) = 1, \theta, Z, b\}$:

Step 1 Generate a random value of b from the standard normal distribution $N(0, 1)$;

Step 2 Compute $a_0(t) = z'\beta_1 + h(t, \beta_2) + b\sqrt{\sigma^2 + \tau^2}$;

Step 3 Generate w from a truncated normal distribution $N(a_0(c), 1)$ with $w \geq 0$. Let

$a(t, c, w) = a_0(t+c)\sqrt{g(t+c)} - \{a_0(c) - w\}\sqrt{g(c)}$;

Step 4 Choose an integer n and generate a random sample $x_1, \ldots, x_n$ from the normal distribution $N(0, d/n)$. Compute the cumulative sum $s(i) = x_1 + \cdots + x_i$, $i = 1, \ldots, n$ for $i = 1, \ldots n$;

Step 5 Compute $\delta = \max_{1 \leq i \leq n}\{s(i) - a(id/n, c, w)\}$;

Step 6 Repeat Steps 4 and 5 for N_1 times and record the proportion of times when a $\delta > 0$ is observed. Denote this proportion by $p(w, \theta, b)$. This gives a Monte Carlo estimate of $P\{\min_{(c \leq t \leq c+d)} Y(t) = 0 | W(c) = w, \theta, Z, b\}$;

Step 7 Repeat Steps 3 through 6 for N_2 times and average $p(w, \theta, b)$ over w. This gives a Monte Carlo estimate of $P[\min_{(c \leq t \leq c+d)} Y(t) = 0 | Y(c) = 1, \theta, Z, b]$;

Step 8 Repeat Steps 1 through 7 for N_3 times and average $P\{\min_{(c \leq t \leq c+d)} Y(t) = 0 | Y(c) = 1, \theta, Z, b\}$ over b. This gives a Monte Carlo estimate of $P\{\min_{(c \leq t \leq c+d)} Y(t) = 0 | Y(c) = 1, \theta, Z\}$.

When a patient's disease condition is worsening over time, $h(t, \beta_2)$ will be a monotone function in t. This is the case when the onset time of a milestone event onset is informative to physicians in determining treatment options. The monotonicity established in the following Theorem 3 provides a theoretical support that the new PI onset time measure not

only provides a meaningful measure of disease status but also reflects the stable degenerating nature of the disease as well. Time c discussed above represents the potential time that PI can occur. Let $c_0(> 0)$ be a lower bound of c. The value of c_0 depends on patient population and is generally known to physicians. For example, $c_0 \geq 1$ day is always true for all PD patients who does not have PI at baseline ($t = 0$) since PD is a slow progressive disease. We now establish some sufficient conditions when the probability function $P\{\min_{(c \leq t \leq c+d)} W(t) \leq 0 \mid W(c) = w, \theta, Z, b\}$ is monotone in c ($\geq c_0$) and w (> 0).

Theorem 3

(i) For any fixed c, θ, and Z, the probability function $P[\min_{(c \leq t \leq c+d)} Y(t) = 0 \mid Y(c) = 1, \theta, Z, b]$ is monotone increasing in d for $d > 0$. (ii) If $g(t) = t$ and function

$a_0(t+c)\sqrt{g(t+c)} - a_0(c)\sqrt{g(c)}$ is monotone increasing in c, then $P[\min_{(c \leq t \leq c+d)} W(t) \leq 0 \mid W(c) = w, \theta, Z, b]$ is monotone decreasing in c. (iii) If $a_0(t)$ is linear in t such as $a_0(t) = \gamma_0 + \gamma_1 t$ and $g(t) = t$, then the sufficient conditions that $P[\min_{(c \leq t \leq c+d)} W(t) \leq 0 \mid W(c) = w, \theta, Z, b]$ is monotone decreasing in c in interval $c \geq c_0$ are $\gamma_1 > 0$ and $\gamma_0 \leq 3c_0\gamma_1$.

The sufficient conditions (iii) in Theorem 3 have a nice clinical interpretation. When $h(t, \beta_2)$ is monotone increasing in t, we must have $\gamma_1 > 0$. Quantities γ_1 and γ_0 characterize the disease progression rate and patient's baseline disease condition. If the patient's condition is worsening over time, and the disease progression rate is fast enough compared to the patient's baseline condition (malignant PD) such that $\gamma_1 > 0$ and $\gamma_0 \leq 3c_0\gamma_1$ hold, then the patient's probability to become PI in a time interval with fixed interval width is increasing over time.

4. MODEL ASSESSMENT

Since repeated measures from binary process $Y_i(t)$ are correlated, classical residuals such as Pearson residuals and deviance residuals (McCullagh and Nelder, 1989) are difficult to calibrate because they have unknown sampling distributions. We propose to use Bayesian residuals with continuous posterior distributions. These residuals can be plotted to check model fitting and to identify outliers. Albert and Chib (1995) proposed Bayesian latent residuals for independent binary response regression models. Chen and Dey (2000) generalized the univariate Bayesian residuals of Albert and Chib (1995) for correlated ordinal data.

For the latent model defined by (1) and (2), we propose a simple approach based on subject-specific latent residuals for assessing the goodness-of-fit of the proposed model. The subject-specific latent residuals are defined as

$$\xi_i = V_i^{-1/2} D_i \{W_i - (z_i'\beta_1 + \sigma b_i + \tau b^{(\delta_i)}) J_{m_i} - H_i(\beta_2)\}$$

where ξ_i is an m_i-dimensional latent vector. It is easy to see that $\xi_i \sim N(0, I_{m_i})$ *apriori*. Also, $\|\xi_i\|^2 = \xi_i'\xi_i \sim \chi_{m_i}^2$ *apriori*. One common approach of using latent residuals in model assessment is to compare the posterior distributions of ξ_i's. However, this approach does not work here because m_i's are subject-dependent and posterior distributions of ξ_i's are not directly comparable between any two subjects with different m_i's. To circumvent this problem, we calculate the posterior probability $p_i = P(\|\xi_i\|^2 \geq \xi_{m_i,\alpha}^2 \mid D_{obs})$ and plot p_i versus $E(\|\xi_i\|^2 \mid D_{obs})$, where $\xi_{m_i,\alpha}^2$ denotes the $(1 - \alpha)$th quantile of the $\chi_{m_i}^2$ distribution, and the expectation is taken with respect to the posterior distribution $\pi(\theta, W \mid D_{obs})$. As $\|\xi_i\|^2 \sim \chi_{m_i}^2$,

apriori, it is expected that p_i's should be roughly around α. Many large p_i's cast doubt upon the model. In addition, an extreme large value of p_i reveals a possible outlier. For our application, $\alpha = 0.05$ may be sufficient although some other values of α can be tried. The subject-specific residuals are quite attractive as they account for the dependency of the repeated longitudinal measurements of $Y_i(t)$. All posterior quantities involved in the subject-specific latent residuals can be easily computed using Gibbs outputs. Thus, our method of using Bayesian latent residuals clearly has an advantage of computational simplicity.

5. APPLICATION IN PARKINSON'S DISEASE CLINICAL TRIAL

5.1 Description of DATATOP Data

The multicenter controlled clinical trial of deprenyl and tocopherol antioxidative therapy of Parkinsonism (DATATOP) was carried out in 1987–1989 (The Parkinson Study Group, 1989, 1990). Eight hundred early PD patients were evaluated by 34 investigators (examiners) from 28 clinical centers. We adopted the PI definition from Elm *et al.* (2005) to define $Y(t) = 1$ if any one of the following three conditions holds at time t: postural stability ≥ 1; or falling ≥ 1; or freezing ≥ 3, and defines $Y(t) = 0$ otherwise. In the DATATOP trial, although each subject was examined by the same investigator throughout the study, the high PI recovery probability makes the traditional PI onset time measure based on the first transition time to $Y(t) = 1$ difficult to interpret and even controversial, as shown below.

Since no tocopherol effect was found in the DATATOP trial, we combined all patients who received deprenyl in DATATOP to form a new treatment group and combined all other patients to form a new placebo group. Let T_{1t} and T_{1p} be the times from baseline with $Y(0) = 0$ to the first times of PI presence with $Y(t) = 1$ in deprenyl and placebo groups, respectively, and let $T_{2t} = \min\{t - T_{1t} : t > T_{1t}, Y(t) = 0\}$ and $T_{2p} = \min\{t - T_{1p} : t > T_{1p}, Y(t) = 0\}$ be the corresponding first recovery times from PI presence status to PI absence status in deprenyl treatment group and placebo group respectively. Using all subjects with $Y(0) = 0$ at baseline (301 subjects in placebo and 302 subjects in deprenyl treatment), and comparing Kaplan-Meier curves of T_{1t} and T_{1p} using the logrank test, we find a strong evidence of deprenyl treatment superiority over the placebo (logrank test p-value = 0.0239). After fitting a proportional hazards model with baseline covariates gender, age, disease duration (duration0 and du2=duration0^2), total UPDRS (total0), Schwab & England activities of daily living scale (seadl0), and treatment (indicator of receiving deprenyl), the estimated hazard ratio is $\exp(-0.324) = 0.723$ for deprenyl with p-value equal to 0.016 (Table 1). These results suggest that deprenyl treatment improves the PI measure and possibly even decreases the progression rate of PD. However, by comparing T_{2t} and T_{2p} using proportional hazards model adjusting for covariates measured at times T_{1t} and T_{1p} respectively, the hazard ratio was $\exp(-0.478) = 0.620$ for deprenyl with p-value equal to 0.006 (Table 1). This result, contrasting to the previous result, show that deprenyl worsens the progression of PD. Such conflicting results not only make the assessment of treatment effect on PD progression difficult, but also make the PI onset time measure not useful to assess a patient's disease condition.

5.2 Posterior inference

There were 34 investigators (examiners) from all 28 centers in the DATATOP trial. Each patient was evaluated by the same investigator throughout the study. Each investigator examined 19 to 44 patients. We use $b^{(\delta_i)}$ to denote the random effect from investigator δ_i and set $C = 34$. To apply the latent process model described in Section 2, we choose $h(t, \beta_2) = \beta_{20} + \beta_{21}t + \beta_{22}T rt + \beta_{23}t \times T rt$ because a linear fit in time is adequate for such a short two-year study. The $T rt$ is the binary indicator of the deprenyl treatment assignment. Covariates in z_i include baseline measures of gender (indicator of female), age (age0), disease duration

(duration0 and du20=duration0^2), total UPDRS (total0), and Schwab & England activities of daily living scale (seadl0). The follow up time t is measured in quarters (= 3 ×30.4375 days). For a comparison, we fit data using three models: the latent process model (LPM) proposed in Section 2, the independent probit model (IPM) that ignores dependency among repeated measures with $P(Y_i(t)=1|\beta_1,\beta_2,\sigma,z_i)=\Phi(z_i'\beta_1+h(t,\beta_2))$, and the random effects probit model (RPM) with $W_i(t)=z_i'\beta_1+h(t,\beta_2)+\sigma b_i+\tau b^{(\delta_i)}+w_{it}$ and $w_{it} \sim N(0, 1)$, where $\{w_{it}, t \geq 0\}$, b_i, and b^{δ_i} are independent.

Gibbs sampling algorithm is used to sample from the posterior distribution and 50,000 Gibbs samples after a burn-in of 1,000 iterations were used to obtain all posterior estimates. The Gibbs sampling algorithm performs quite well. The autocorrelations for all model parameters disappear at lag 10, and the Gibbs sampler converges much earlier before 1,000 iterations. The resulting posterior estimates from the 3 models are given in Table 2. All 95% highest posterior density (HPD) intervals are computed via the Monte Carlo method proposed by Chen and Shao (1999).

Large measurement error in PI measures was well recognized among neurologists. Variability of the force of the pull test leads to inconsistence among examiners and even by the same examiner (Fahn, 2006). This is described vividly by Dr. Robert Hauser: "Sometimes the patient doesn't quite understand the instructions, sometimes he is a little better prepared, sometimes I may pull a little harder, sometimes he'll lean forward just as I pull because he knows what's coming, sometimes he might not try as hard" (Hauser, 2006). Such large measurement error was confirmed by the fitted latent process model. From Table 2, we see that both LPM and RPM found significant effects from patient and examiner as both τ and σ are significantly greater than zero. IPM indicates significant effects of both treatment and treatment-by-time interaction. It suggests that treatment has a short-term benefit at baseline and then it accelerates disease progression during the follow up. The positive sign of the interaction coefficient in RPM also suggests that treatment accelerates, though not significantly, disease progression. Contrasting to IPM and RPM, the negative sign of the treatment and the magnitude of its interaction with time in LPM suggest that treatment slows down, though not significantly, disease progression. Comparing these three models, results from LPM are more consistent with findings from DATATOP investigators who reported beneficial effect of dyprenyl on PD (The Parkinson Study Group, 1989, 1990) and also pointed out that such an effect was only symptomatic but not neuroprotective (Shoulson, 1998).

All three models give similar conclusions regarding to effects of gender, age0, total UPDRS, and seadl0. Both LPM and RPM suggest that patient's PI condition was worsening over time when other covariate levels are held fixed. But IPM fails to detect this PI progression. The posterior mean of σ under RPM is larger than the one under LPM. This is intuitive since σ is the only parameter in RPM to capture the longitudinal dependency. Both LPM and RPM have identified a significant investigator (examiner) effect.

5.3 Model fit assessment

We compare the goodness-of-fit of the proposed latent structural model with the independent probit model and the random effects probit model through the quantity $p_i=P(\|\xi_i\|^2 \geq \xi_{m_i,\alpha}^2|D_{obs})$. For LPM, $\xi_i=V_i^{-1/2}D_i\{W_i - (z_i'\beta_1+\sigma b_i+\tau b^{(\delta_i)})J_{m_i} - H_i(\beta_2)\}$ which is the subject-wise latent residual defined in Section 4; for IPM, $\xi_i=W_i - z_i'\beta_1 J_{m_i} - H_i(\beta_2)$; and for RPM, $\xi_i=W_i - (z_i'\beta_1+\sigma b_i+\tau b^{(\delta_i)})J_{m_i} - H_i(\beta_2)$. Frequencies of the p_is in different ranges ($p_i \leq 0.05$, $0.05 < p_i \leq 0.1$, $0.1 < p_i \leq 0.3$, $0.3 < p_i \leq 0.5$, $0.5 < p_i \leq 0.8$ and $0.8 < p_i \leq 1$) are tabulated in Table 3. Although IPM has the highest proportion of p_is below 0.05, it has

16.79% of $p_i s$ greater than 0.1 with 2.88% of them even greater than 0.5. The LPM is slightly better than RPM with a higher proportion of $p_i s$ below 0.05. However, LPM provides many more attractive features than the other two models and is more suitable for modeling the PI process.

The fit from latent process model (LPM) is further compared with the observed proportion of $Y(t) = 1$ in different homogeneous patient clusters. We partition all patients into 4 mutually exclusive groups: male/treatment, male/placebo, female/treatment, and female/placebo. In each of these 4 groups, we standardize patients' baseline measures of seadl0, total0, age0, and duration0 by dividing them by their sample standard deviations. We use 4-dimensional Euclidean distance in standardized (seadl0, total0, age0, duration0) measure to further partition patients into two clusters. For each cluster, we compute the observed proportion of $\{Y(t) = 1\}$ at each time interval (with width of 3 months), the posterior estimate, $E[P\{Y(t) = 1|\bar{z}, \theta\}|D_{obs}]$, and the corresponding 95% HPD interval of $P\{Y(t) = 1|\bar{z}, \theta\}$. The observed proportion of $\{Y(t) = 1\}$ is computed using the sample mean of all available observations in $Y(t)$ within that time window no matter whether it was from the same subject or not. The expectation $E[P\{Y(t) = 1|\bar{z}, \theta\}|D_{obs}]$ is taken with respect to the posterior distribution of $\theta = (\beta_1, \beta_2, \sigma^2)$ and $\bar{z}$ is the average baseline covariates in that cluster. These results show that the estimated values from LPM are quite close to the observed values across all clusters. Figure 1 displays plots from two chosen clusters.

5.4 New measure of PI onset time

Let $v(\theta|c, d, Z) = P\{\min_{(c \le t \le c+d)} Y(t) = 0 \mid Y(c) = 1, \theta, Z\}$. To predict DATATOP patient's PI onset time using the new PI onset time measure, we compute the probability $v(\hat{\theta}|c, d, Z)$ using some typical combinations of covariate values Z, where $\hat{\theta}$ is the posterior mean of θ. In DATATOP baseline visit, 75% patients have disease duration less than 1.66 years, and the first three quartiles of seadl0 are 90, 90, and 95, respectively. We thus fix duration0 = 1 year and seadl0 = 90. Using the algorithm given in Section 3 with $N_1 = 500$, $N_2 = 500$, $N_3 = 2000$, and $n = 20$, we compute Monte Carlo estimates of probability $v(\hat{\theta}|c, d, Z)$ under different combinations of baseline covariates Z and set $c = 1, 1.5$ years. The results are given in Table 4 where duration0 = 1, seadl0 = 90, and d = 3 months and 6 months respectively. It is seen that older female subjects with higher baseline total UPDRS score tend to have lower probabilities to recover from PI in both $d = 3$ months and $d = 6$ months cases. As c increases, the probability to recover from $Y(t) = 1$ decreases, representing a decreased chance to observe $Y(t) = 0$ and thus an increased chance to become PI. The simulated recovery probabilities $v(\hat{\theta}|c, d, Z)$, across all covariate combinations in Table 4, are much smaller than the corresponding observations from T_{2p} and T_{2t} defined in Section 5.1. In general, the threshold value η for the recovery probability $v(\hat{\theta}|c, d, Z)$ should be obtained from clinicians or physicians, and it often depends on physician's objective. For example, if the goal is to detect mild impairment in postural reflexes, η can be set at a relative larger value such as $\eta > 30\%$. On the other hand, if the goal is to detect a severe postural impairment, one can set $\eta < 5\%$. In any case, providing $v(\hat{\theta}|c, d, Z)$ with different combinations of c, d, and Z will help investigators and physicians in assessing a patient's disease condition and designing future Parkinson's disease studies.

We close this section with a remark that Monte Carlo estimates of probability $v(\hat{\theta}|c, d, z)$ evaluated at the posterior mean of θ are very close to the Monte Carlo estimates of probability $E[v(\theta|c, d, z) \mid D_{obs}]$, where the expectation is taken with respect to the posterior distribution of θ given the observed data D_{obs}. For example, for all estimates with $d = 3$ listed in Table 4, the mean difference $\{v(\hat{\theta}|c, d, z) - E[v(\theta|c, d, z) \mid D_{obs}]\}$ is −0.011 with a range from −0.028 to 0.013. Similar results are obtained when $d = 6$. This finding suggests that $v(\hat{\theta}|c, d, z)$ provides a good approximation to $E[v(\theta|c, d, z) \mid D_{obs}]$. In addition, the use

of $v(\hat{\theta}|c, d, z)$ is more computationally attractive as computing $E[v(\theta|c, d, z) \mid D_{obs}]$ takes much longer time.

6. DISCUSSION

A good clinical outcome should be clinically relevant and accurate. Traditional onset time can be inaccurate and clinically less useful when the event indicator is subject to large measurement errors. For progressive diseases, an event onset time defined through a latent model can incorporate measurement errors from both patients and examiners, and thus provide a more clinically meaningful measure. Because a latent model can borrow strength from different subjects and repeated measures, and can adjust for effects from other confounding risk factors, it is nature to expect that the new onset time defined through a latent mode will be more accurate provided the model is reasonable.

The use of η and d makes it flexible in defining an event onset time. The time interval width d relates to the robustness of the new event onset measure, and threshold value η relates to the severity of the disease condition. When $d = \eta = 0$, our new onset time reduces to the traditional onset time. The algorithm in Section 3 makes the new onset time computation feasible. This paper only illustrates its application in PI process for Parkinson's disease, the same methodology can be applied to develop other event onset time measures with similar features. For example, the time to certain stage of glomerular filtration rate in end-stage renal disease, and the time to dementia in many neurodegenerated diseases.

Acknowledgments

References

1. Albert JH, Chib S. Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 1993;88:669–679.

2. Albert JH, Chib S. Bayesian residual analysis for binary response regression models. Biometrika 1995;82:747–759.

3. Bloem BR. Postural instability in Parkinson's disease. Clinical Neurology and Neurosurgery 1992;94(Suppl):S41–S45. [PubMed: 1320515]

4. Chen M-H, Dey DK. Bayesian modeling of correlated binary responses via scale mixture of multivariate normal link functions. Sankhyâ, Series A 1998;60:322–343.

5. Chen M-H, Dey DK. A unified Bayesian analysis for correlated ordinal data models. Brazilian Journal of Probability and Statistics 2000;14:87–111.

6. Chen M-H, Dey DK, Shao Q-M. A new skewed link model for dichotomous quantal response data. Journal of the American Statistical Association 1999;94:1172–1186.

7. Chen M-H, Shao Q-M. Monte Carlo estimation of Bayesian credible and HPD intervals. Journal of Computational and Graphical Statistics 1999;8:69–92.

8. Diggle, P.; Liang, KY.; Zeger, SL. Analysis of Longitudinal Data. Oxford: Oxford University Press; 1994.

9. Dunson DB. Dynamic latent trait models for multidimensional longitudinal data. Journal of the American Statistical Association 2003;98:555–563.

10. Elm, et al. A responsive outcome for Parkinson's Disease Neuroprotection Futility Studies. Annals of Neurology 2005;57:197–203. [PubMed: 15668964]

11. Fahn, S. Personal conversation. 2006.

12. Hauser, R. Personal conversation. 2006.

13. Kieburtz K. Designing neuroprotection trials in Parkinson's disease (with discussions). Annals of Neurology 2003;53:S100–109. [PubMed: 12666102]

14. Landau W. Pyramid sale in the bucket shop: DATATOP bottoms out. Neurology 1990;40:1337–1339. [PubMed: 2118239]

15. Munhoz, et al. Evaluation of the pull test technique in assessing postural instability in Parkinson's disease. Neurology 2004;62:125–127. [PubMed: 14718714]

16. Rogers MW. Disorders of posture, balance, and gait in Parkinson's disease. Clinics in Geriatric Medicine 1996;12(4):825–45. [PubMed: 8890118]

17. Shoulson I. Parkinson Study Group. DATATOP: a decade of neuroprotective inquiry. Annals of Neurology 1998;44(Suppl 1):S160–S166. [PubMed: 9749589]

18. The Parkinson Study Group. DATATOP: a multiceter controlled clinical trial in early Parkinson's disease. Archives of Neurology 1989;46:1052–1060. [PubMed: 2508608]

19. The Parkinson Study Group. Effect of deprenyl on the progression of disability in early Parkinson's disease. New England Journal of Medicine 1990;332(21):1526–1528.

APPENDIX: PROOFS OF THEOREMS

Proof of Theorem 1

Let A be a $m_i \times m_i$ matrix whose (i, j)-element a_{ij} satisfies $a_{ii} = 1$, $a_{i,i-1} = -1$, $a_{ij} = 0$ for all $j < i - 1$ and $j > i$, i.e.,

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \cdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & -1 & 1 \end{pmatrix}.$$

Then $AV_iA' = D_i = \mathrm{diag}(t_{i1}, t_{i2} - t_{i1}, t_{i3} - t_{i2}, \ldots, t_{im} - t_{i,m_i-1})$ and $V_i^{-1} = A'D_i^{-1}A$. For any vector $\mathbf{x} = (x_1, \ldots, x_m)^T$, we have

$$\mathbf{x}'V_i^{-1}\mathbf{x} = (A\mathbf{x})'V_i^{-1}(A\mathbf{x}) = \frac{x_1^2}{t_{i1}} + \sum_{j=2}^{m_i} \frac{(x_j - x_{j-1})^2}{t_{ij} - t_{i,j-1}}.$$

Proof of Theorem 2

We note that

$$P\left(\min_{(c \leq t \leq c+d)} Y(t) = 0 \mid W(c) = w\right) = P\left(\varepsilon(t) \text{ up crossing } a(t, c, w), 0 \leq t \leq d\right). \tag{A.1}$$

Define stopping time $T = \inf\{t : t \geq 0, \varepsilon(t) = a(t, c, w)\}$. Applying Durbin's formula (1985), we obtain

$$P(\min_{c \leq t \leq c+d} Y(t)=0|Y(c)=1)P(W(c)>0)$$

$$=\int_0^{+\infty} P(\min_{c \leq t \leq c+d} W(t) \leq 0|W(c)=w)f(c,w)dw$$

$$=\int_0^{+\infty} (P(\varepsilon(t) \text{ up crossing } a(t,c,w), 0 \leq t \leq d) f(c,w)dw,$$

$$=\int_0^{+\infty} f(c,w)dw \int_0^d \frac{1}{\sqrt{t}}\varphi\left(\frac{a(t,c,w)}{\sqrt{t}}\right) \times \lim_{s \uparrow t}\frac{1}{t-s}E\left\{1_{(T>s)}(a(t,c,w) - \varepsilon(s))\Big|\varepsilon(t)=(t,c,w)\right\} dt.$$

Proof of Theorem 3

Claim in (i) is obvious.

(ii). Based on the identity (A.1) derived in the proof of Theorem 2 and the fact that the boundary function $a(t, c, w)$ is continuous in $t \geq 0$ with $a(0,c,w)=w\sqrt{c}>0$, it suffices to show that $a(t, c, w)$ is monotone increasing in c. In fact, that $a_0(t+c)\sqrt{g(t+c)} - a_0(c)\sqrt{g(c)}$ is monotone in-creasing in c implies that $a(t,c,w)=a_0(t+c)\sqrt{g(t+c)} - a_0(c)\sqrt{g(c)}+w\sqrt{g(c)}$ is also monotone increasing in c for any $w > 0$.

(iii). When $a_0(t) = \gamma_0 + \gamma_1 t$, we have

$$\frac{\partial a(t,c,w)}{\partial c}=\frac{1}{2}(\sqrt{t+c} - \sqrt{c})\frac{3\gamma_1\sqrt{c(t+c)}-\gamma_0}{\sqrt{c(t+c)}}+\frac{w}{2\sqrt{c}}$$

$$\geq \frac{1}{2}(\sqrt{t+c} - \sqrt{c})\frac{3c_0\gamma_1-\gamma_0}{\sqrt{c(t+c)}}+\frac{w}{2\sqrt{c}}>0.$$

This implies that $a(t, c, w)$ is monotone increasing in c for any fixed $t \geq 0$.

Figure 1.
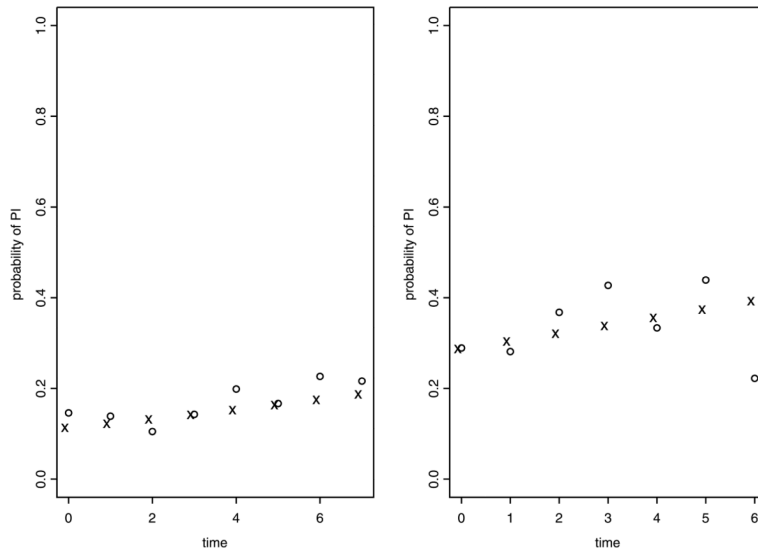Comparison between observed proportion of Y (t) = 1 (symbol ∘) and posterior estimate E[P {Y (t) = 1|z̄, θ}|D_{obs}] (symbol ×)

Table 1

Maximum likelihood estimate of parameters using a proportional hazard model. Covariates for T_1 are measures at baseline visit while covariates for T_2 are measures at time T_1

	Variable	Estimate	Standard Error	P-value	Hazard Ratio
T_1 = time from baseline to the first onset time of $Y = 1$	treatment	−0.324	0.134	0.016	0.723
	age0	0.029	0.008	0.0001	1.030
	total0	0.045	0.006	<.0001	1.046
	seadl0	−0.006	0.010	0.561	0.994
	duration0	0.043	0.168	0.797	1.044
	du20	0.019	0.037	0.598	1.020
	gender	0.610	0.139	<.0001	1.841
T_2 = time from T_1 to the first recovery time of $Y = 0$	treatment	−0.478	0.174	0.006	0.620
	age	0.021	0.010	0.050	1.021
	total	0.004	0.009	0.621	1.004
	seadl	0.002	0.016	0.887	1.002
	duration	−0.340	0.244	0.155	0.706
	du2	0.092	0.045	0.041	1.097
	gender	−0.091	0.178	0.608	0.913

Table 2

Posterior Estimates for DATATOP Trial Data

Variable	model	Posterior Mean	Posterior Standard Error	95% HPD Interval
intercept	LPM	−1.382	0.717	(−2.827, − 0.023)
	IPM	−2.037	0.331	(−2.689, − 1.396)
	RPM	−1.912	1.031	(−3.863, 0.183)
gender	LPM	0.313	0.072	(0.174, 0.458)
	IPM	0.363	0.034	(0.296, 0.428)
	RPM	0.492	0.100	(0.301, 0.690)
age0	LPM	0.015	0.004	(0.007, 0.022)
	IPM	0.013	0.002	(0.010, 0.017)
	RPM	0.023	0.005	(0.013, 0.033)
duration0	LPM	0.085	0.088	(−0.090, 0.255)
	IPM	0.240	0.044	(0.155, 0.328)
	RPM	0.128	0.125	(−0.117, 0.370)
du2	LPM	−0.020	0.020	(−0.059, 0.021)
	IPM	−0.056	0.011	(−0.077, − 0.035)
	RPM	−0.030	0.029	(−0.086, 0.027)
total0	LPM	0.038	0.004	(0.030, 0.047)
	IPM	0.035	0.002	(0.031, 0.039)
	RPM	0.055	0.006	(0.043, 0.067)
seadl0	LPM	−0.020	0.007	(−0.033, − 0.008)
	IPM	−0.012	0.003	(−0.018, − 0.006)
	RPM	−0.032	0.010	(−0.050, − 0.013)
$T\,rt$	LPM	−0.113	0.075	(−0.259, 0.034)
	IPM	−0.130	0.053	(−0.233, − 0.027)
	RPM	−0.176	0.105	(−0.382, 0.031)
t	LPM	0.064	0.010	(0.045, 0.082)
	IPM	0.008	0.007	(−0.005, 0.021)
	RPM	0.057	0.010	(0.038, 0.076)
$t \times T\,rt$	LPM	0.001	0.012	(−0.023, 0.026)
	IPM	0.024	0.009	(0.006, 0.041)
	RPM	0.006	0.012	(−0.018, 0.031)
σ	LPM	0.346	0.109	(0.111, 0.546)
	RPM	1.039	0.064	(0.920, 1.169)
τ	LPM	0.414	0.069	(0.287, 0.553)
	RPM	0.548	0.092	(0.380, 0.734)

Table 3

Frequency and proportion of residuals $p_i = P(\|\xi_i\|^2 \geq \xi^2_{m_i,\alpha}|D_{obs})$ in different ranges

Range	LPM	IPM	RPM
$p_i \leq 0.05$	490 (61.40%)	527 (66.04%)	422 (52.88%)
$0.05 < p_i \leq 0.1$	277 (34.71%)	137 (17.17%)	350 (43.86%)
$0.1 < p_i \leq 0.3$	31 (3.88%)	86 (10.78%)	26 (3.26%)
$0.3 < p_i \leq 0.5$	0 (0%)	25 (3.13%)	0 (0%)
$0.5 < p_i \leq 0.8$	0 (0%)	9 (1.13%)	0 (0%)
$0.8 < p_i \leq 1$	0 (0%)	14 (1.75%)	0 (0%)
Total	798 (100%)	798 (100%)	798 (100%)

Table 4

Monte Carlo Estimates of Probabilities $v(\hat{\theta}|c, d, z)$ with duration0= 1, seadl0 = 90, and d=3 months and 6 months for DATATOP Trial Data

c in years	treatment	total0	age0	gender	$v(\hat{\theta}|c, d = 3, z)$	$v(\hat{\theta}|c, d = 6, z)$
1.0	Placebo	16	55	Female	0.378	0.469
1.0	Placebo	16	55	Male	0.425	0.524
1.0	Placebo	16	69	Female	0.342	0.428
1.0	Placebo	16	69	Male	0.399	0.494
1.0	Placebo	25	55	Female	0.323	0.404
1.0	Placebo	25	55	Male	0.372	0.462
1.0	Placebo	25	69	Female	0.279	0.352
1.0	Placebo	25	69	Male	0.343	0.428
1.0	Deprenyl	16	55	Female	0.395	0.489
1.0	Deprenyl	16	55	Male	0.444	0.546
1.0	Deprenyl	16	69	Female	0.357	0.444
1.0	Deprenyl	16	69	Male	0.409	0.505
1.0	Deprenyl	25	55	Female	0.328	0.410
1.0	Deprenyl	25	55	Male	0.385	0.478
1.0	Deprenyl	25	69	Female	0.297	0.374
1.0	Deprenyl	25	69	Male	0.349	0.435
1.5	Placebo	16	55	Female	0.296	0.374
1.5	Placebo	16	55	Male	0.340	0.427
1.5	Placebo	16	69	Female	0.268	0.340
1.5	Placebo	16	69	Male	0.305	0.386
1.5	Placebo	25	55	Female	0.241	0.307
1.5	Placebo	25	55	Male	0.286	0.362
1.5	Placebo	25	69	Female	0.220	0.282
1.5	Placebo	25	69	Male	0.254	0.323
1.5	Deprenyl	16	55	Female	0.305	0.384
1.5	Deprenyl	16	55	Male	0.353	0.442
1.5	Deprenyl	16	69	Female	0.280	0.354
1.5	Deprenyl	16	69	Male	0.330	0.415

c in years	treatment	total0	age0	gender	$v(\theta|c, d = 3, z)$	$v(\theta|c, d = 6, z)$
1.5	Deprenyl	25	55	Female	0.260	0.329
1.5	Deprenyl	25	55	Male	0.304	0.383
1.5	Deprenyl	25	69	Female	0.229	0.292
1.5	Deprenyl	25	69	Male	0.280	0.354