SHARED PARAMETER MODELS FOR THE JOINT ANALYSIS OF LONGITUDINAL DATA AND EVENT TIMES

Seventh Annual Winter Workshop
Longitudinal Data Analysis
University of Florida, Gainesville, Florida 2005
January 6-8, 2004

Edward F. Vonesh
(Ed.Vonesh@baxter.com)
Statistics, Epidemiology and Surveillance
Baxter Healthcare Corporation, Round Lake, IL 60073
Acknowledgments

• MDRD Study
  > Tom Greene and Mark Schluchter
  > Paper submitted for publication

• Head & Neck Cancer Study
  > F Rademaker, J Logemann, BR Pualoski, D Liu
  > Paper published in Head & Neck, 2003
  > NIH P01 CA40007 and NIH P50 DE/CA11921
1 Introduction

Longitudinal Studies:

- Longitudinal studies often gather joint information on
  - Serial Outcome Measures
    - Repeated Measurements, Growth Curve Data
  - Time to Some Event
    - Patient Survival, Time to Dropout

Questions of Interest

- Is primary interest in comparing event-free survival times?
  - How do we adjust for effects of serial outcome measures?
- Is primary interest in comparing serial trends over time?
  - How do we account for non-ignorable dropout?
Primary Applications

• Surrogate Markers of Outcomes
  - Justify a serial marker as a surrogate for patient outcome
    > Use of CD4 T-lymphocyte counts as a marker for patient survival in clinical trials of patients with AIDS
      - Tsiatis, DeGruttola and Wulfsohn (JASA, 1995)
      - Faucett and Thomas (Statistics in Medicine, 1996)
  
• Handling of Nonignorable Missing (NIM) Data Due to Dropout
  - Provide valid inference in presence of informative censoring
    > NIH-Sponsored Trials
    > Regulatory Submissions for New Drug Applications (NDA)

• Requires joint modeling of serial data and dropout
Example

- TRIAL DESIGN FOR OXYMORPHONE DRUG
  ESTABLISHED BY ENDO, FDA (News Release: July 20, 2004)

  “Endo Pharmaceuticals has reached an agreement with the FDA on the design of a new clinical trial to confirm the safety and efficacy of its experimental pain drug, oxymorphone extended-release (ER) tablets, the firm said.

  Last October, the FDA issued an approvable letter for Endo’s oxymorphone ER product, but requested the firm conduct additional trials. At a meeting with Endo in May, the FDA said it was concerned that the outcome of two of the three Phase III efficacy trials, which met the predefined primary endpoints, may have been favorably biased by the statistical handling of data from patients who did not complete the trials, Endo said. The additional 12-week clinical trial is intended to address this issue...“
2 Motivating Examples

MDRD Study (Schluchter et al, Stat in Med, 1992, 2001)
- Modification of Diet in Renal Disease (MDRD) Study B
  - Randomized trial of 255 patients with chronic kidney disease
  - Primary interest lies in comparing the effects of two different interventions on the progression of renal disease as measured by Glomerular Filtration Rate (GFR ml/min)

  > Modification of diet
    Diet L: Low protein intake,
    Diet K: Very low protein intake
  > Modification of blood pressure
    Usual blood pressure
    Low blood pressure
MDRD Study: Patient profiles for Diet K, Low BP Group

Diet K, Low BP Group

GFR (ml/min)

Months of Follow-up
MDRD Problems/Issues:

- Serial measurements of GFR planned at 0, 2, 4, 8 months and every 4 months thereafter
- Plots indicate a linear decline in GFR over time
- Patients with more severe renal impairment were enrolled
  - Initial GFR of 13-24 ml/min
- Due to staggered entry, planned follow-up was from $1 \frac{1}{2}$ to 4 years
- Patient drop out was 40% (101 of 255)
- Causes of dropout
  - dialysis (81 of 255 patients, 32%)
  - kidney transplant (11 of 255 patients, 4%)
  - death/other medical (9 of 255 patients, 4%)
MDRD Study: Evidence of Non-ignorable Dropout?

**EBLUP's of GFR Slope Assuming Missing Data are MAR**

- **GFR Slope (mL/min/month)**
- **Months of Follow-up**

**Dropout Status:**
- • • • Completed Study
- • • • Terminated Study
ADEMEX Study (Paniagua et al., JASN, 2002)

- Adequacy of Peritoneal Dialysis in Mexico (ADEMEX)
  - Randomized trial of 965 patients with end-stage renal disease
  - Primary interest lies in comparing the effects of a single intervention on patient survival
    > Dose of Dialysis
      - Control: Standard dose of dialysis (N=484)
      - Treated: High dose dialysis (N=481)
  - Secondary interest lies in comparing the decline in Glomerular Filtration Rate (GFR) over time
    > Does the dose of dialysis affect the decline in GFR?
ADEMEX: SS and PA profiles assuming data are MCAR
ADEMEX Problems/Issues

ADEMEX Patient status at end of trial

<table>
<thead>
<tr>
<th>Patient status</th>
<th>Control</th>
<th>Treated</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>-Lost to followup</td>
<td>48</td>
<td>10%</td>
<td>42</td>
</tr>
<tr>
<td>-Completed</td>
<td>191</td>
<td>39%</td>
<td>201</td>
</tr>
<tr>
<td>-Died</td>
<td>112</td>
<td>23%</td>
<td>103</td>
</tr>
<tr>
<td>-Drop to HD</td>
<td>48</td>
<td>10%</td>
<td>63</td>
</tr>
<tr>
<td>-Drop to PD</td>
<td>47</td>
<td>10%</td>
<td>45</td>
</tr>
<tr>
<td>-Return of Kidney</td>
<td>1</td>
<td>0.2%</td>
<td>1</td>
</tr>
<tr>
<td>-Transplant</td>
<td>37</td>
<td>8%</td>
<td>26</td>
</tr>
<tr>
<td>Total</td>
<td>484</td>
<td>100%</td>
<td>481</td>
</tr>
</tbody>
</table>

15
Head & Neck Cancer Study (Rademaker et al., *Head and Neck*, 2003)

- NIH sponsored study of head and neck cancer patients
  - Prospective one-year longitudinal study of 255 cancer patients
  - Primary interest lies in characterizing the functional ability to eat over a 12 month period following chemoradiation treatment in patients with head and neck cancer

  > Outcome measures (binary)
  
  Percent of patients with % oral intake < 50%
  
  Percent of patients who could not eat a normal diet

  > Covariates
  
  Age, Gender, Race, Site of cancer, Time to dropout
Head & Neck Cancer Study - Problems/Issues

- Longitudinal study of binary outcome measures
- Approximate time to dropout
  - midpoint between last visit and date of last contact
  - discrete time survival based on planned visits
- Number of patients still in study

<table>
<thead>
<tr>
<th>Evaluation Point</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Pretreatment (tx)</td>
<td>255</td>
<td>100%</td>
</tr>
<tr>
<td>- 1 months post-tx</td>
<td>186</td>
<td>73%</td>
</tr>
<tr>
<td>- 3 months post-tx</td>
<td>148</td>
<td>58%</td>
</tr>
<tr>
<td>- 6 months post-tx</td>
<td>126</td>
<td>49%</td>
</tr>
<tr>
<td>- 12 months post-tx</td>
<td>90</td>
<td>35%</td>
</tr>
</tbody>
</table>
\section{Shared Parameter (SP) Models}

- Let \( \mathbf{y}_i = (\mathbf{y}_i^o, \mathbf{y}_i^m)' = (y_{i1}, \ldots, y_{ip})' \) be the “complete-data” vector for the \( i^{th} \) subject \((i = 1, \ldots, n)\) with \( \mathbf{y}_i^o = (y_{i1}, \ldots, y_{ip})' \) representing the observed data and \( \mathbf{y}_i^m \) unobserved or missing data.
- Alternatively, set the “complete-data” \( \mathbf{y}_i = \mathbf{y}_i(t) \) to be a continuous time process with \( \mathbf{y}_i^o = \{y_i(t) : t = 1, \ldots, p_i\} \)
- Let \( T_i = \) time to some event
- Let \( \mathbf{b}_i \) be a shared random-effects vector
- The Shared Parameter Model for Jointly Modeling \( (\mathbf{y}_i, T_i) \)

\[
\pi(\mathbf{y}_i, T_i) = \int_{\mathbf{b}} \pi(\mathbf{y}_i, T_i|\mathbf{b}_i) \pi(\mathbf{b}_i) d\mathbf{b}_i
\]

\[
= \int_{\mathbf{b}} \pi(\mathbf{y}_i|\mathbf{b}_i) \pi(T_i|\mathbf{b}_i) \pi(\mathbf{b}_i) d\mathbf{b}_i
\]

Observe that \( \mathbf{y}_i|\mathbf{b} \) and \( T_i|\mathbf{b} \) are conditionally independent given \( \mathbf{b}_i \).
Historical Genesis

• Shared Parameter Models for Nonignorable Missing (NIM) Data:
  > Follmann and Wu (*Biometrics*, 1995)
  > Ten Have et al (*Biometrics*, 1998)

• Assuming the $y_{ij}$ are conditionally independent given $b_i$,
  the Shared Parameter Model factors as

$$
\pi(y_i^o, T_i) = \int_{y_m} \int_b \pi(y_i^o, y_i^m, T_i | b) \pi(b) \, db \, dy_m
$$

$$
= \int_b \pi(y_i^o | b) \pi(T_i | b) \pi(b) \, db \left\{ \int_{y_m} \frac{\pi(y_i^m | b)}{\pi(b)} \, dy_i^m \right\}
$$

$$
= \int_b \pi(y_i^o | b) \pi(T_i | b) \pi(b) \, db
$$

• This model belongs to the class of random-effects dependent
  selection models as described by Little (*JASA*, 1995)
Shared Parameter Models - Key Features

- The shared parameter model induces marginal correlation between \( y \) and \( T \) through their joint dependence on \( b \).

- Inference under a shared parameter model does not require missing or unobserved data to be MCAR or MAR.

\[
\pi(T_i | y_i^o, y_i^m) = \frac{\int_b \pi(T_i, y_i^o, y_i^m | b_i) \pi(b_i) db_i}{\int_b \pi(y_i^o, y_i^m | b_i) \pi(b_i) db_i} \\
= \frac{\int_b \pi(T_i | b_i) \pi(y_i^o, y_i^m | b_i) \pi(b_i) db_i}{\int_b \pi(y_i^o, y_i^m | b_i) \pi(b_i) db_i} \\
= \int_b \pi(T_i | b) \pi(b | y_i^o, y_i^m) db
\]

so that, in general, the conditional distribution of \( T_i | y_i \) depends on \( y_i^m \) through posterior distribution of \( b_i \).
Shared Parameter Model Specifications:

- A shared parameter model is obtained by specifying
  
  > A conditional model for the longitudinal data: \( \pi(y_i|b_i) \)
    - The \( b_i \) serve as subject-specific random effects
  > A conditional model for the event time data: \( \pi(T_i|b_i) \)
    - The \( b_i \) serve as subject-specific covariates
3.1 Conditional Models for Longitudinal Data

A Generalized Nonlinear Mixed-Effects (GNLME) Model:

- For broad applications entailing continuous or discrete data, the conditional pdf of $y_i|b_i$ is assumed to come from the quadratic exponential family (Prentice and Zhao, 1991, *Biometrics*):

$$
\pi(y_i|b_i) = \Delta_i^{-1} \exp\{y_i'\xi_i + w_i'\zeta_i + c_i(y_i)\} 
$$

where $w_i = Vech(y_i'y_i')$, $c_i(\cdot)$ is a “shape” function, $\Delta_i$ is a normalization constant, $\xi_i$ and $\zeta_i$ are canonical parameter vectors expressed in terms of the conditional mean and variance, $\mu_i$ and $\sigma_i$, via $\xi'_i = \xi'_i(\mu_i, \sigma_i)$, and $\zeta'_i = \zeta'_i(\mu_i, \sigma_i)$

- Examples: Multivariate Normal, Gamma, Inverse Gaussian, Poisson, Binomial, Correlated Binary Data
Conditional Models for Longitudinal Data

- The GNLME model may be written in terms of the vector of conditional means, $\mu_i$, and vector of conditional variances, $\sigma_i$, as:

$$E(y_i|b_i) = \mu_i(\beta, b_i, x_i, z_i) = \mu_i(\beta, b_i); \quad i=1,\ldots,n$$
$$Var(y_i|b_i) = \Lambda_i(\beta, \alpha, b_i, x_i, z_i) = \Lambda_i(\beta, \alpha, b_i)$$

where

- $y_i = [y_1 \ldots y_{p_i}]'$ is a $p_i \times 1$ vector of repeated measures
- $x_i$ and $z_i$ are vectors of covariates associated with the fixed- and random-effects parameters, respectively
- $\beta$ is a $s \times 1$ vector of fixed-effects parameters
- $\sigma_i = Vech[\Lambda_i(\beta, \alpha, b_i)]$ is the vector of conditional variances
- $\alpha$ is a $u \times 1$ vector of conditional covariance parameters
- $b_i$ is a $v \times 1$ vector of random-effects $\sim iid \mathcal{N}(0, \Psi)$
3.2 Conditional Models for Event Time Data

Strategies for modeling $T$

- Let $T_i$ be time to event (e.g., death, dropout, etc.)
- We observe $T_i^* = \min(T_i, C_i)$ and $\delta_i$ where
  
  $C_i =$ observed censoring times (e.g., staggered entries)

  $$\delta_i = \begin{cases} 
  1 & \text{if } T_i^* = T_i \\
  0 & \text{if } T_i = C_i. 
  \end{cases}$$

- Nonparametric models for $T_i$ include
  - Kaplan-Meier survival (Hogan and Laird, 1997)
  - Cox proportional hazards model (Tsiatis et al, 1995)

- Parametric (Accelerated Failure-Time) Models
  - Log-normal (Schluchter, 1992; DeGruttola & Tu, 1994)
  - Weibull, Piecewise exponential
Conditional Models for Event Time Data

Accelerated Failure-Time (AFT) Models:

- For continuous-time survival data, the conditional pdf of $T_i | b_i$ is assumed to be from the class of accelerated failure-time models:

$$T_i | b_i = \exp \{ f_i (\beta, b_i)' \eta \} T_{0i}$$  \hspace{1cm} (2)

where

- $f_i (\beta, b_i)' \eta$ is a conditional linear function in $\eta$ of some vector function of covariates $f_i (\beta, b_i) = f_i (\beta, b_i, x_i, z_i)$ that may depend on $\beta$ and $b_i$, as well as covariates $x_i$ and $z_i$.
- $T_{0i} = \exp (\eta_0) T_{\epsilon_i}^\phi$ are iid event times from some baseline distribution with $\eta_0$ and $\phi$ representing location and scale parameters for $W_{0i} = \log (T_{0i}) = \eta_0 + \phi \log (T_{\epsilon_i})$.

- Examples: Exponential, Weibull, Extreme Value, Log-Normal
Conditional Models for Event Time Data

Accelerated Failure-Time (AFT) Models:

- Weibull model with shape parameter $\gamma$
  
  - Parametric but fairly flexible
  
  - The Weibull model is a proportional hazards model
    
    > Baseline Hazard rate: $\lambda_0(t) = \gamma \exp(\eta_0) t^{\gamma-1}$
    
    > Hazard rate: $\lambda(t) = \lambda_0(t) \exp\{ f_i(\beta, b_i)' \eta \}$

  - Fairly robust to misspecification of the baseline hazard
Conditional Models for Event Time Data

Accelerated Failure-Time (AFT) Models:

- Piecewise Exponential Model (Interval Poisson Model)
  - Semi-parametric and flexible:
    > Partition $T$ into $k$ disjoint intervals $(t_0, t_1], \ldots, (t_{k-1}, t_k]$
    > Baseline Hazard rate: $\lambda_0(t) = \sum_{h=1}^{k} \lambda_{0h} I(t \in (t_{h-1}, t_h])$
    > Hazard rate: $\lambda(t) = \lambda_0(t) \exp\{f_{ih}(\beta, b_i, t_{h-1})'\eta\}$
  - Easily handles time-dependent covariates
  - Proportional and Non-proportional hazards
Example - The MDRD Study

- Let $T_i =$ time to dropout, and $y_{ij} =$ GFR at time $t_{ij}$.
- For patients in the Diet K, Low BP Group, assume

**GNLME Model:** (Linear mixed-effects regression)

$$y_{ij}|b_i = \beta_{1i} + \beta_{2i}t_{ij} + \epsilon_{ij}$$

$$\beta_{1i} = \beta_1 + b_{1i}$$

$$\beta_{2i} = \beta_2 + b_{2i}$$

**AFT Model:**

$$T_i|b_i = \exp\{f_i(\beta, b_i)'\eta\}T_{0i}$$

$$f_i(\beta, b_i)'\eta = \eta_1\beta_{1i} + \eta_2\beta_{2i}$$

$T_{0i} \sim \text{Weibull}(\lambda_0(t), \gamma)$ with baseline hazard $\lambda_0(t) = \gamma \exp(\eta_0)t^{\gamma-1}$

- The hazard rate is $\lambda(t) = \lambda_0(t)\exp\{\eta_1\beta_{1i} + \eta_2\beta_{2i}\}$
- The joint model for $(y_{ij}, T_i)$ is nonlinear in $b_i$. 
Example - The ADEMEX Study

- Let $T_i =$time to dropout, and $y_{ij} =$GFR at time $t_{ij}$.
- For patients in Control or Treated Groups, assume

**GNLME Model: (Nonlinear mixed-effects regression)**

$$y_{ij} | b_i = \beta_{1i} \exp\{-\beta_{2i} t_{ij}\} + \epsilon_{ij}$$

$$\beta_{1i} = \beta_{11} + \beta_{12} G_i + b_{1i}$$

$$\beta_{2i} = \beta_{21} + \beta_{22} G_i + b_{2i}$$

**AFT Model:**

$$T_i | b_i = \exp\{f_i(\beta, b_i)' \eta\} T_{0i}$$

$$f_i(\beta, b_i)' \eta = \eta_1 \beta_{1i} + \eta_2 \beta_{2i}$$

where $G_i$ is a treatment group indicator (control=0, treated=1), $T_{0i} \sim Weibull(\lambda_0(t), \gamma)$ with baseline hazard $\lambda_0(t) = \gamma \exp(\eta_0) t^{\gamma-1}$

- The hazard rate is $\lambda(t) = \lambda_0(t) \exp\{\eta_1 \beta_{1i} + \eta_2 \beta_{2i}\}$
Example - The Head & Neck Cancer Study

- Let $T_i$ be a discrete survival time variable defined by $T_i = j$ ($j = 1, 2, 3, 4$) when dropout occurs in $(t_{j-1}, t_j]$; $t_j = 1, 3, 6, 12$ months. Let $y_{ij} = 1$ when % oral intake < 50% and 0 otherwise.

**GNLME Model: (Logistic mixed-effects regression)**

$$P(y_{ij} = 1|b_i) = \frac{exp\{\beta_i\}}{1 + exp\{\beta_i\}} = E(y_{ij}|b_i)$$

$$\beta_i = (\beta_0 + b_i) + \beta_1 Sex_i + \beta_2 Age_i + \sum_{j=1}^{4} \beta_{3j} I(t_j)$$

**Discrete Time Failure Model:**

$$P(T_i = j|b_i) = \lambda_{ij} \prod_{k=1}^{j-1} (1 - \lambda_{ik})$$

$$\lambda_{ik}|b_i = 1 - exp(-exp\{\eta_{ik} + \eta_1 Sex_i + \eta_2 Age_i + \eta_3 b_i\})$$

where $I(t_j)$ are 0-1 indicators, $\lambda_{ij}$ is a discrete-time hazard rate, and $\eta_{ik}$ defines baseline conditional survival probability in $(t_{k-1}, t_k]$. 
4 Estimation

- Maximum likelihood (ML) estimation requires maximizing the integrated log-likelihood

\[ L(\beta, \alpha, \theta, \eta, \phi; y, T) = \]

\[ \sum_{i=1}^{n} \log \int_b \pi(y_i | b_i) \pi(T_i^* | b_i)^\delta_i S(T_i^* | b_i)^{1-\delta_i} \pi(b_i) db_i \]

where \( \theta = Vech(\Psi) \) and \( S(T_i^* | b_i) \) is the conditional survivor function.

- Methods include
  - Numerical integration (Gauss-Hermite quadrature)
  - Conditional Generalized Estimating Equations (CGEE)
    > First-order approximations (NLME, PQL)
    > Second-order approximations (Laplace MLE or LMLE)
Overview of Estimation Procedures

- Numerical integration techniques (NLMIXED)
- Markov Chain Monte Carlo techniques (WinBUGS)
- First-order approximations (S-Plus, R, %NLINMIX)
  Breslow & Clayton (1993, *JASA*) - PQL
  Vonesh, Wang, Nie and Mujamdar (2002, *JASA*) - CGEE2
- Second-order Laplace approximation (NLMIXED, qpoints=1)
  Vonesh (1996, *Biometrika*) - Laplace MLE (LMLE)
  Raudenbush, Yang & Yosef (2000, *J Comp and Graph Stat*)
5 Examples:

MDRD Study:
• Preliminary survival analysis comparing various life-time distributions
  - Kaplan-Meier survival
  - Log-normal survival (Schluchter et al, 2001)
  - Exponential survival (i.e., Weibull(1))
  - Weibull survival
Figure 1: Survival curves for Diet K, Low BP Group
Examples:

MDRD Study:

• The longitudinal model assuming ignorable dropout (MAR) is:

\[ GFR_{ij} = \beta_{1i} + \beta_{2i}t_{ij} + \epsilon_{ij} \]

\[ \beta_{1i} = \beta_{11}G_{1i} + \beta_{12}G_{2i} + \beta_{13}G_{3i} + \beta_{14}G_{4i} + b_{1i} \]

\[ \beta_{2i} = \beta_{21}G_{1i} + \beta_{22}G_{2i} + \beta_{23}G_{3i} + \beta_{24}G_{4i} + b_{2i} \]

where \( G_{1i}, G_{2i}, G_{3i} \) and \( G_{4i} \) are group indicator variables that define which diet and blood pressure group the \( i^{th} \) patient belongs.

• To assess whether dropout is ignorable, we fit models assuming \( GFR \) and \( T \) are independent (MCAR) versus dependent (NIM)

  - \( T \sim \text{Weibull with hazard rate } \lambda_0(t)exp(\sum_{k=1}^{4} \eta_kG_{ki}) \) - MCAR
  - \( T \sim \text{Weibull with hazard rate } \lambda_0(t)exp(\eta_1\beta_{1i} + \eta_2\beta_{2i}) \) - NIM
  - \( T \sim \text{Piecewise Exponential with } \lambda_0(t)exp(\eta_1\beta_{1i} + \eta_2\beta_{2i}) \) - NIM
Table 1: Standard MLE (MCAR) versus SP (NIM) Models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Standard MLE</th>
<th>Weibull SP (Shape=4.87)</th>
<th>Piecewise Exp SP (6-months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi$</td>
<td>$\pi(y, T) = \pi(y) \pi(T)$</td>
<td>(19.4, 0.158)</td>
<td>(19.4, 0.144)</td>
</tr>
<tr>
<td></td>
<td>(19.9, 0.061)</td>
<td>(0.158, 0.067)</td>
<td>(0.144, 0.064)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>5.23</td>
<td>5.14</td>
<td>5.13</td>
</tr>
<tr>
<td>Slopes: $\beta_{21}$</td>
<td>$-0.251$</td>
<td>$-0.327 (30%)$</td>
<td>$-0.315 (25%)$</td>
</tr>
<tr>
<td>$\beta_{22}$</td>
<td>$-0.252$</td>
<td>$-0.296 (17%)$</td>
<td>$-0.288 (14%)$</td>
</tr>
<tr>
<td>$\beta_{23}$</td>
<td>$-0.292$</td>
<td>$-0.319 (9%)$</td>
<td>$-0.313 (7%)$</td>
</tr>
<tr>
<td>$\beta_{24}$</td>
<td>$-0.326$</td>
<td>$-0.388 (19%)$</td>
<td>$-0.382 (17%)$</td>
</tr>
<tr>
<td>$L(y</td>
<td>b)$</td>
<td>$-2428.16$</td>
<td>$-2422.90$</td>
</tr>
<tr>
<td>$-2 L(y, T)$</td>
<td>7617.1</td>
<td>7387.6</td>
<td>7406.0</td>
</tr>
</tbody>
</table>
Association Between Rate of Decline in GFR and Dropout

Piecewise Exponential (Interval Poisson) Shared Parameter Model

Empirical Bayes Estimates of GFR Slopes Assuming Nonignorable Dropout

GFR Slope (ml/min/month)

Months of Follow-up

Dropout Status:  
Completed Study  
Terminated Study
Table 2: Laplace vs Numerical Quadrature for Weibull SP Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>QPOINTS=1</th>
<th>QPOINTS=10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull Shape, $\gamma$</td>
<td>4.873 (0.677)</td>
<td>5.355 (0.806)</td>
</tr>
<tr>
<td>$\eta_0^*$</td>
<td>-3.421 (0.152)</td>
<td>-3.444 (0.147)</td>
</tr>
<tr>
<td>$\eta_1^*$</td>
<td>-0.061 (0.009)</td>
<td>-0.061 (0.009)</td>
</tr>
<tr>
<td>$\eta_2^*$</td>
<td>-2.661 (0.202)</td>
<td>-2.699 (0.201)</td>
</tr>
<tr>
<td>$\Psi, \sigma^2$</td>
<td>$\begin{pmatrix} 19.4 &amp; 0.158 \ 0.158 &amp; 0.067 \end{pmatrix}, 5.14$</td>
<td>$\begin{pmatrix} 19.4 &amp; 0.158 \ 0.158 &amp; 0.068 \end{pmatrix}, 5.15$</td>
</tr>
<tr>
<td>Intercept: Diet K, Low BP</td>
<td>19.62 (0.571)</td>
<td>19.61 (0.570)</td>
</tr>
<tr>
<td>Slope: Diet K, Low BP</td>
<td>-0.327 (0.035)</td>
<td>-0.331 (0.035)</td>
</tr>
<tr>
<td>$-2\log L(y, T)$</td>
<td>7387.6</td>
<td>7386.2</td>
</tr>
<tr>
<td>CPU (hr:min:sec)</td>
<td>1:27:12.79</td>
<td>12:39:42.25</td>
</tr>
</tbody>
</table>
Examples:

**MDRD Study:**

- We compared different SP models for Diet K, Low BP Group
- Purpose: Evaluate the performance of different SP models with respect to inference on both dropout and rate of change in GFR
  - Normal SP model of Schluchter et al (2001)
    > $g(T_i) = T_i$ (Linear transformation $\Rightarrow$ Normality)
    > $g(T_i) = \log(T_i)$ (Logarithmic transformation $\Rightarrow$ Normality)
  - Weibull SP model (Laplace-based MLE)
  - Piecewise Exponential SP model (Laplace-based MLE)
- Laplace approximation implemented via NLMIXED (qpoints=1).
  - See also website in Guo and Carlin (2004, *Amer Stat*) for WinBUGS (MCMC) and SAS (NLMIXED) code
SAS Data Structure for Weibull Model:

Diet K, Low BP Group

(Dropout=1, Dropout Time=23.66 months)

<table>
<thead>
<tr>
<th>ptid</th>
<th>ind</th>
<th>months</th>
<th>GFR</th>
<th>Drop</th>
<th>T</th>
<th>response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0.00</td>
<td>16.36</td>
<td>1</td>
<td>23.66</td>
<td>16.36</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>4.14</td>
<td>14.93</td>
<td>1</td>
<td>23.66</td>
<td>14.93</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>8.41</td>
<td>12.41</td>
<td>1</td>
<td>23.66</td>
<td>12.41</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>12.35</td>
<td>10.95</td>
<td>1</td>
<td>23.66</td>
<td>10.95</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>16.30</td>
<td>8.56</td>
<td>1</td>
<td>23.66</td>
<td>8.56</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>20.47</td>
<td>6.12</td>
<td>1</td>
<td>23.66</td>
<td>6.12</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.00</td>
<td>16.36</td>
<td>1</td>
<td>23.66</td>
<td>23.66</td>
</tr>
</tbody>
</table>
SAS Code for Weibull Model:

```
proc nlmixed data=CGEE2 start qpoints=1;
  parms b11=20 b21=0 eta0=-3 eta11=0 eta21=0
       Gamma=2 s11=1 s12=0 s22=1 Sigma_Sq=1;
  b1i=b11 + u1;    b2i=b21 + u2;
  Mu = (b1i + b2i*Months);  SD = sqrt(Sigma_Sq);
  Li = exp(eta0 + eta11*b1i + eta21*b2i);
  Hi = Gamma*(Li**Gamma)*(T**(Gamma-1));
  ll_Y = (1-ind)*(-0.5*((RESPONSE-Mu)**2)/Sigma_Sq
       - 0.5*log(Sigm_Sq)) ;
  ll_T = ind*(RESPONSE*log(Hi) - (Li*T)**Gamma );
model response ~ general(ll_Y + ll_T);
random u1 u2 ~ normal([0,0],[s11,s12,s22]) sub=ptid;
```
### Table 3a. Results under different SP models for Diet K, Low BP

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lognormal†</th>
<th>Weibull‡‡</th>
<th>Piecewise Exp‡‡††</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g(T) = \log(T)$</td>
<td>Shape($\gamma$) = 2.965</td>
<td>(6-months)</td>
<td></td>
</tr>
<tr>
<td>$\Psi$</td>
<td>(16.4, 0.064)</td>
<td>(16.3, 0.073)</td>
<td>(16.3, 0.072)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>5.45</td>
<td>5.45</td>
<td>5.45</td>
</tr>
<tr>
<td>$\beta_1$ (ml/min)</td>
<td>19.50 (0.51)</td>
<td>19.51 (0.53)</td>
<td>19.49 (0.53)</td>
</tr>
<tr>
<td>$\beta_2$ (ml/min/mo)</td>
<td>-0.294 (0.044)</td>
<td>-0.298 (0.034)</td>
<td>-0.292 (0.034)</td>
</tr>
<tr>
<td>$\rho(T, b_0)$ or $RR$</td>
<td>$\rho = 0.16$</td>
<td>$RR(b_0) = 0.91^*$</td>
<td>$RR(b_0) = 0.90^*$</td>
</tr>
<tr>
<td>$\rho(T, b_1)$ or $RR$</td>
<td>$\rho = 0.74$</td>
<td>$RR(b_1) = 0.44^*$</td>
<td>$RR(b_1) = 0.48^*$</td>
</tr>
</tbody>
</table>

† Schluchter et. al. (2001, *Stat in Med*) - *EM algorithm*

‡‡ Laplace approximation
Examples:

MDRD Study:

- We compared the effect of different covariate specifications
- Purpose: Evaluate different dropout mechanisms under a piecewise exponential model using different covariate functions.

<table>
<thead>
<tr>
<th>Model</th>
<th>Covariates: $f_i(\beta, b_i)'$</th>
<th>$\lambda(t) = \lambda_0h \exp{f_i(\beta, b_i)'\eta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(\beta_1i, \beta_2i)$</td>
<td>$\lambda_0h \exp{\eta_1\beta_1i + \eta_2\beta_2i}$</td>
</tr>
<tr>
<td>2</td>
<td>$\mu_{ih}(t_{h-1}) = \beta_1i + \beta_2it_{h-1}$</td>
<td>$\lambda_0h \exp{\eta_1[\beta_1i + \beta_2it_{h-1}]}$</td>
</tr>
<tr>
<td>3</td>
<td>$y_{ih}(t_{i(h-1)})^\dagger$</td>
<td>$\lambda_0h \exp{\eta_1y_{ih}}$</td>
</tr>
<tr>
<td>4</td>
<td>$(\beta_1i, \beta_2i, y_{ih})$</td>
<td>$\lambda_0h \exp{\eta_1\beta_1i + \eta_2\beta_2i + \eta_3y_{ih}}$</td>
</tr>
</tbody>
</table>

$\dagger y_{ih}(t_{i(h-1)}) = \text{last observed GFR prior to } (t_{h-1}, t_h]$. This model tests for an ignorable threshold effect and yields standard MLE’s.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta_{1i}, \beta_{2i}$</th>
<th>$\mu_{ih}(t_{h-1})$</th>
<th>$y_{ih}(t_{i(h-1)})$</th>
<th>$\beta_{1i}, \beta_{2i}, y_{ih}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi$</td>
<td>(16.3, 0.072)</td>
<td>(16.5, 0.026)</td>
<td>(16.6, 0.020)</td>
<td>(16.3, 0.073)</td>
</tr>
<tr>
<td></td>
<td>(0.072, 0.049)</td>
<td>(0.026, 0.042)</td>
<td>(0.020, 0.038)</td>
<td>(0.073, 0.049)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>5.45 (0.394)</td>
<td>5.45 (0.395)</td>
<td>5.50 (0.401)</td>
<td>5.45 (0.394)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>19.49 (0.53)</td>
<td>19.41 (0.53)</td>
<td>19.35 (0.54)</td>
<td>19.49 (0.53)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.292 (0.034)</td>
<td>-0.264 (0.032)</td>
<td>-0.244 (0.030)</td>
<td>-0.292 (0.034)</td>
</tr>
<tr>
<td>$RR(\beta_{1i})$</td>
<td>1.11 (0.074)†</td>
<td>–</td>
<td>–</td>
<td>1.11 (0.139)†</td>
</tr>
<tr>
<td>$RR(\beta_{2i})$</td>
<td>2.08 (0.413)†</td>
<td>–</td>
<td>–</td>
<td>2.10 (0.494)†</td>
</tr>
<tr>
<td>$RR(y_t)$</td>
<td>–</td>
<td>1.21 (0.065)†</td>
<td>1.14 (0.050)†</td>
<td>0.99 (0.097)</td>
</tr>
<tr>
<td>$lnL(y</td>
<td>\hat{b})$</td>
<td>-632.7</td>
<td>-632.1</td>
<td>-634.3</td>
</tr>
<tr>
<td>$lnL(T</td>
<td>\hat{b})$</td>
<td>-94.8</td>
<td>-112.9</td>
<td>-117.2</td>
</tr>
<tr>
<td>$AIC(y,T)$</td>
<td>1925.2</td>
<td>1933.4</td>
<td>1937.4</td>
<td>1927.2</td>
</tr>
</tbody>
</table>
Table 4: GOF comparing Weibull to Piecewise Exp. SP model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Weibull $\beta_{1i}, \beta_{2i}$</th>
<th>Piecewise Exponential $\beta_{1i}, \beta_{2i}$</th>
<th>$y_{ih}(t_{i(h-1)})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Psi$</td>
<td>(16.3, 0.073)</td>
<td>(16.3, 0.072)</td>
<td>(16.6, 0.020)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>5.45 (0.394)</td>
<td>5.45 (0.394)</td>
<td>5.50 (0.401)</td>
</tr>
<tr>
<td>$\beta_1$ (ml/min)</td>
<td>19.51 (0.53)</td>
<td>19.49 (0.53)</td>
<td>19.35 (0.54)</td>
</tr>
<tr>
<td>$\beta_2$ (ml/min/month)</td>
<td>-0.298 (0.034)</td>
<td>-0.292 (0.034)</td>
<td>-0.244 (0.030)</td>
</tr>
<tr>
<td>$RR(\beta_{1i})$</td>
<td>1.10 (0.074)$^\dagger$</td>
<td>1.11 (0.074)$^\dagger$</td>
<td>$-$</td>
</tr>
<tr>
<td>$RR(\beta_{2i})$</td>
<td>2.29 (0.513)$^\dagger$</td>
<td>2.08 (0.413)$^\dagger$</td>
<td>$-$</td>
</tr>
<tr>
<td>$-2LogL(y,T)$</td>
<td>1900.6</td>
<td>1895.2</td>
<td>1909.4</td>
</tr>
<tr>
<td>$AIC(y,T)$</td>
<td>1920.6</td>
<td>1925.2</td>
<td>1937.4</td>
</tr>
</tbody>
</table>
Unadjusted and Adjusted Survival Curves - Diet K, Low BP Group

\[ \hat{S}_T(t) = \frac{1}{n} \sum_{i=1}^{n} \hat{S}_{T|b}(t|\hat{\omega}, \hat{b}_i) \]
\[ \hat{S}_T(t) = \frac{1}{n} \sum_{i=1}^{n} \hat{S}_{T|b}(t|\hat{\omega}, \hat{\beta_i}); \quad \tilde{S}_T(t) = \int_b S_{T|b}(t|\hat{\omega}, b) \pi(b) db \]
Estimated Weibull Survival Curves for Diet K, Low BP Group

\[
R(t | b) = R(b) \exp(b) 
\]

Adjusted PA: \( \int_b S_{T|b}(t | \omega, b) \pi(b) db \), Adjusted SS: \( \hat{S}_{T|b}(t | \omega, b = 0) \).
PA Survival and SS Survival Curves for Diet K, Low BP Group

PA Curve (Marginal), --- SS Curve (at $b_i = E(b_i) = 0$)
Examples:

Head and Neck Cancer Study:

- We compared two different SP models versus traditional GEE
- Purpose: Jointly evaluate the performance of different shared parameter models with respect to estimating both dropout and percent of patients with < 50% oral intake over time
  - GEE model based on exchangeable correlation (MCAR)
  - Weibull model (proportional hazards)
    > $T =$ midpoint between last evaluation and last contact
    > Adaptive Gaussian quadrature
  - Discrete time survival model (proportional hazards)
    > Binary regression with complementary log-log link
    > Adaptive Gaussian quadrature
Table 2: Percent of patients with oral intake < 50%. Comparison of observed (unadjusted) versus adjusted\(^\dagger\) percentages obtained via logistic regression (GEE) and shared parameter model (SPM).

<table>
<thead>
<tr>
<th>Evaluation Point</th>
<th>N</th>
<th>Observed</th>
<th>MCAR</th>
<th>SP Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretreatment</td>
<td>253</td>
<td>7%</td>
<td>7%</td>
<td>4%</td>
</tr>
<tr>
<td>1 month</td>
<td>186</td>
<td>39%</td>
<td>39%</td>
<td>40%</td>
</tr>
<tr>
<td>3 months</td>
<td>148</td>
<td>26%</td>
<td>28%</td>
<td>30%</td>
</tr>
<tr>
<td>6 months</td>
<td>127</td>
<td>21%</td>
<td>22%</td>
<td>21%</td>
</tr>
<tr>
<td>12 months</td>
<td>91</td>
<td>10%</td>
<td>13%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Results exclude two patients with suspect site information

† Percentages adjusted for sex, age, race, site of disease
7 Conclusions:

- We consider a class of generalized linear and nonlinear mixed-effects models from a quadratic exponential family which includes multivariate models for continuous and discrete data.

- We propose using Piecewise Exponential (Interval Poisson) or Weibull regression models for modeling the “survival” times.

  1. The Piecewise Exponential model is semi-parametric and allows for proportional or non-proportional hazards as well as time-dependent covariates.

  2. The Weibull model is a proportional hazards model that offers great flexibility in modeling the baseline hazard function.

- Alternatively, one can apply these techniques to discrete time survival models for interval-censored survival data.
Conclusions:

- One may use CGEE or Laplace MLE (LMLE) to estimate parameters from a Joint Shared Parameter model. These techniques are easily implemented using existing software
  
  CGEE1: SAS macro NLINMIX
  
  CGEE2: SAS macro CGEE2 (a modified version of NLINMIX)
  
  LMLE: SAS procedure NLMIXED (QPOINTS=1)

- LMLE generally provides a better approximation and lower mean bias compared to CGEE (PQL, CGEE2) methods

- For discrete, highly sparse data, numerical integration is recommended (adaptive Gaussian quadrature)
Conclusions:

- The paradox of jointly modeling event time data and serial data.
  1. The modeling of survival data requires large numbers of events to achieve reasonable estimates of those effects associated with event times.
  2. Subject-specific inference in the longitudinal setting requires a moderately large number of observations per subject.
  3. Jointly modeling these two outcome variables may be at odds with one another
- Inference in the presence of NIM data is model-dependent. Sensitivity analyses are needed to ensure the results are not overly sensitive to the model specifications.
CGEE (PQL, CGEE2) versus LMLE:

- CGEE/PQL may be viewed as an approximate Laplace-based ML estimator in that, for fixed $\Psi$, the set of CGEE for $\tau = (\beta, \alpha)$ are related to the Laplacian-based estimating equations via:

$$U_{\text{LMLE}}(\tau, \hat{b}) = U_{\text{CGEE}}(\tau, \hat{b}) - \frac{1}{2} \sum_{i=1}^{n} \left\{ \frac{\partial}{\partial \tau} \log \left( | -L''_i(\tau, \hat{b}_i(\tau)) | \right) \right\}$$

$$= U_{\text{CGEE}}(\tau, \hat{b}) - \frac{1}{2} \sum_{i=1}^{n} \left\{ \frac{\partial}{\partial \tau} \log \left( | -l''_i(\tau, \hat{b}_i(\tau)) + \Psi^{-1} | \right) \right\}$$

where

$$U_{\text{CGEE}}(\tau, \hat{b}) = \frac{\partial}{\partial \tau} \sum_{i=1}^{n} \left\{ l_i(\tau, b_i) - \frac{1}{2} b'_i \Psi^{-1} b_i \right\} \bigg|_{b_i = \hat{b}_i(\tau)}$$

$$l_i(\tau, b_i) = l_i(\beta, \alpha, b_i; y_i) = \log[\pi(y_i|b_i)]$$

$$L''_i(\tau, \hat{b}_i(\tau)) = \frac{\partial^2}{\partial b_i \partial b_i'} \left\{ l_i(\tau, b_i) - \frac{1}{2} b'_i \Psi^{-1} b_i \right\} \bigg|_{b_i = \hat{b}_i(\tau)}$$
Table 1: Select estimates from a SP model for head & neck cancer data using a discrete time survival model for time to dropout.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Parameter</th>
<th>Laplace Approx.</th>
<th>Adaptive Gaussian</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_{ij}$</td>
<td>$\beta_0$ (intercept)</td>
<td>$11.802 (3.455)^\dagger$</td>
<td>$9.057 (2.334)^\dagger$</td>
</tr>
<tr>
<td></td>
<td>$\beta_1$ (sex)</td>
<td>$1.528 (0.989)$</td>
<td>$1.159 (0.657)$</td>
</tr>
<tr>
<td></td>
<td>$\beta_2$ (age)</td>
<td>$-0.127 (0.045)^\dagger$</td>
<td>$-0.101 (0.031)^\dagger$</td>
</tr>
<tr>
<td></td>
<td>$\psi$</td>
<td>$24.236 (12.861)$</td>
<td>$9.195 (2.637)$</td>
</tr>
<tr>
<td>$T_i$</td>
<td>$\lambda_1$ (sex)</td>
<td>$0.088 (0.205)$</td>
<td>$0.086 (0.206)$</td>
</tr>
<tr>
<td></td>
<td>$\lambda_2$ (age)</td>
<td>$0.004 (0.009)$</td>
<td>$0.004 (0.009)$</td>
</tr>
<tr>
<td></td>
<td>$\lambda_3$ ($b_i$)</td>
<td>$-0.066 (0.030)^\dagger$</td>
<td>$-0.106 (0.046)^\dagger$</td>
</tr>
<tr>
<td></td>
<td>$-2Log L(y, T)$</td>
<td>1189.6</td>
<td>1196.4</td>
</tr>
<tr>
<td></td>
<td>CPU (hr:min:sec)</td>
<td>0:9:19.00</td>
<td>0:20:54.31</td>
</tr>
</tbody>
</table>