

**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...”

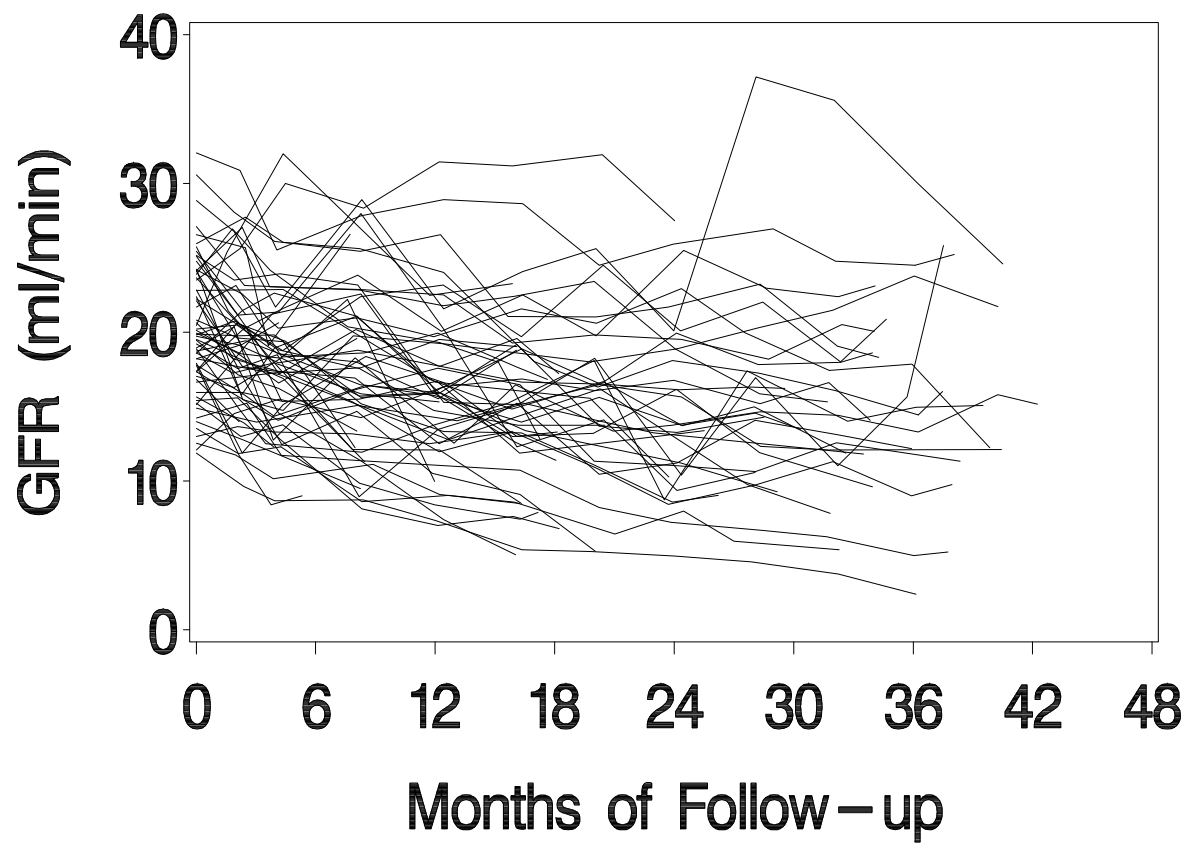
2 Motivating Examples

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

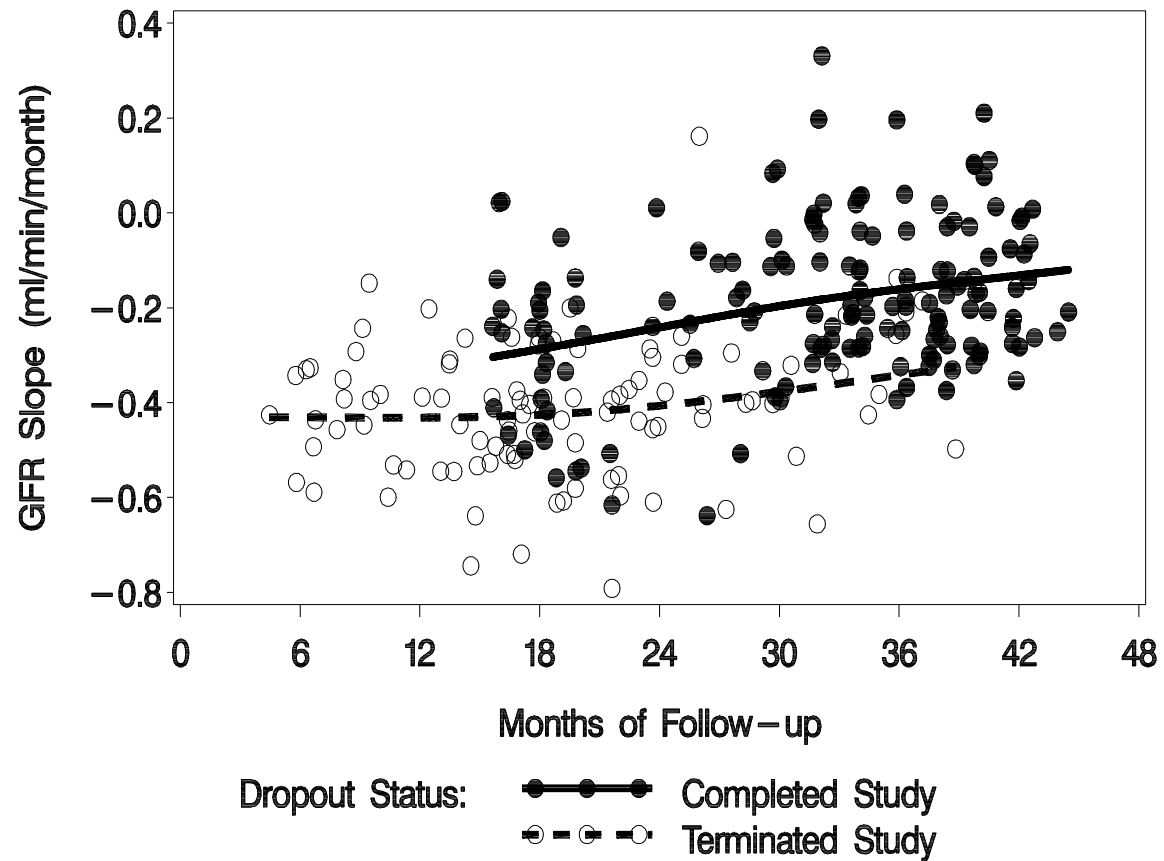


MDRD Problems/Issues:

- 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 1½ 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

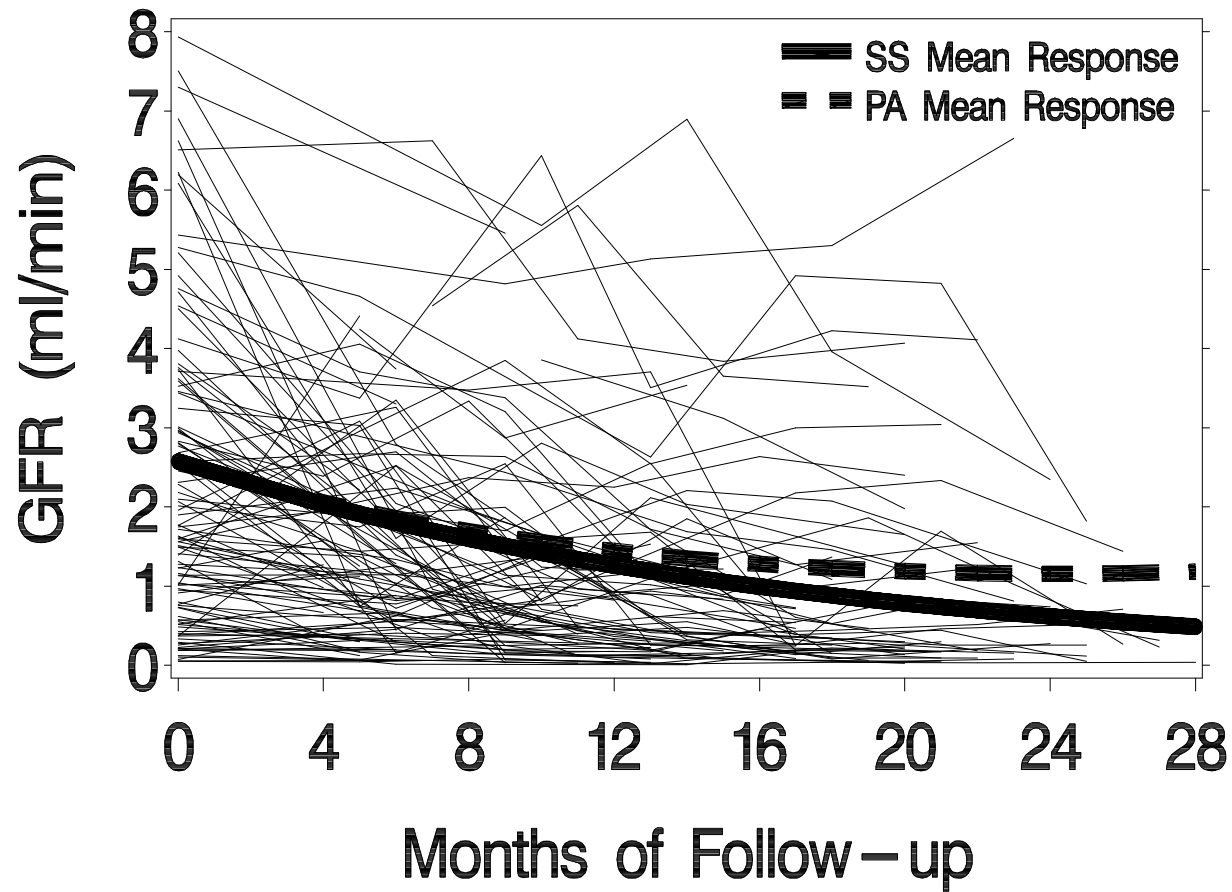


ADEMEX Study (Paniagua et al., JASN, 2002)

- 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

Control Patients



ADEMEX Problems/Issues

ADEMEX Patient status at end of trial

Patient status	Control		Treated		p
	N	%	N	%	
-Lost to followup	48	10%	42	9%	NS
-Completed	191	39%	201	42%	
-Died	112	23%	103	21%	
-Drop to HD	48	10%	63	13%	
-Drop to PD	47	10%	45	9%	
-Return of Kidney	1	0.2%	1	0.2%	
-Transplant	37	8%	26	5%	
Total	484	100%	481	100%	

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

Evaluation Point	N	%
-Pretreatment (tx)	255	100%
- 1 months post-tx	186	73%
- 3 months post-tx	148	58%
- 6 months post-tx	126	49%
-12 months post-tx	90	35%

3 Shared Parameter (SP) Models

- Let $\mathbf{y}_i = (\mathbf{y}_i^o, \mathbf{y}_i^m)' = (y_{i1}, \dots, y_{ip})'$ be the “complete-data” vector for the i^{th} subject ($i = 1, \dots, n$) with $\mathbf{y}_i^o = (y_{i1}, \dots, y_{ip_i})'$ 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 $\mathbf{y}_i^o = \{y_i(t) : t = 1, \dots, p_i\}$
- 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)

$$\begin{aligned}\pi(\mathbf{y}_i, T_i) &= \int_{\mathbf{b}} \pi(\mathbf{y}_i, T_i | \mathbf{b}_i) \pi(\mathbf{b}_i) d\mathbf{b}_i \\ &= \int_{\mathbf{b}} \pi(\mathbf{y}_i | \mathbf{b}_i) \pi(T_i | \mathbf{b}_i) \pi(\mathbf{b}_i) d\mathbf{b}_i\end{aligned}$$

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 \mathbf{b}_i ,

the Shared Parameter Model factors as

$$\begin{aligned}\pi(\mathbf{y}_i^o, T_i) &= \int_{\mathbf{y}^m} \int_{\mathbf{b}} \pi(\mathbf{y}_i^o, \mathbf{y}_i^m, T_i | \mathbf{b}) \pi(\mathbf{b}) d\mathbf{b} d\mathbf{y}^m \\ &= \int_{\mathbf{b}} \pi(\mathbf{y}_i^o | \mathbf{b}) \pi(T_i | \mathbf{b}) \pi(\mathbf{b}) d\mathbf{b} \left\{ \int_{\mathbf{y}^m} \frac{\pi(\mathbf{y}_i^m, \mathbf{b})}{\pi(\mathbf{b})} d\mathbf{y}_i^m \right\} \\ &= \int_{\mathbf{b}} \pi(\mathbf{y}_i^o | \mathbf{b}) \pi(T_i | \mathbf{b}) \pi(\mathbf{b}) d\mathbf{b}\end{aligned}$$

- 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 \mathbf{y} and T through their joint dependence on \mathbf{b} .
- Inference under a shared parameter model does not require missing or unobserved data to be MCAR or MAR.

$$\begin{aligned}\pi(T_i|\mathbf{y}_i^o, \mathbf{y}_i^m) &= \frac{\int_{\mathbf{b}} \pi(T_i, \mathbf{y}_i^o, \mathbf{y}_i^m|\mathbf{b}_i)\pi(\mathbf{b}_i)d\mathbf{b}_i}{\int_{\mathbf{b}} \pi(\mathbf{y}_i^o, \mathbf{y}_i^m|\mathbf{b}_i)\pi(\mathbf{b}_i)d\mathbf{b}_i} \\ &= \frac{\int_{\mathbf{b}} \pi(T_i|\mathbf{b}_i)\pi(\mathbf{y}_i^o, \mathbf{y}_i^m|\mathbf{b}_i)\pi(\mathbf{b}_i)d\mathbf{b}_i}{\int_{\mathbf{b}} \pi(\mathbf{y}_i^o, \mathbf{y}_i^m|\mathbf{b}_i)\pi(\mathbf{b}_i)d\mathbf{b}_i} \\ &= \int_{\mathbf{b}} \pi(T_i|\mathbf{b})\pi(\mathbf{b}|\mathbf{y}_i^o, \mathbf{y}_i^m)d\mathbf{b}\end{aligned}$$

so that, in general, the conditional distribution of $T_i|\mathbf{y}_i$ depends on \mathbf{y}_i^m through posterior distribution of \mathbf{b}_i .

Shared Parameter Model Specifications:

- A shared parameter model is obtained by specifying
 - > A conditional model for the longitudinal data: $\pi(\mathbf{y}_i | \mathbf{b}_i)$
 - The \mathbf{b}_i serve as subject-specific random effects
 - > A conditional model for the event time data: $\pi(T_i | \mathbf{b}_i)$
 - The \mathbf{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 $\mathbf{y}_i|\mathbf{b}_i$ is assumed to come from the quadratic exponential family (Prentice and Zhao, 1991, *Biometrics*):

$$\pi(\mathbf{y}_i|\mathbf{b}_i) = \Delta_i^{-1} \exp\{\mathbf{y}_i' \boldsymbol{\xi}_i + \mathbf{w}_i' \boldsymbol{\zeta}_i + c_i(\mathbf{y}_i)\} \quad (1)$$

where $\mathbf{w}_i = \text{Vech}(\mathbf{y}_i \mathbf{y}_i')$, $c_i(\cdot)$ is a “shape” function, Δ_i is a normalization constant, $\boldsymbol{\xi}_i$ and $\boldsymbol{\zeta}_i$ are canonical parameter vectors expressed in terms of the conditional mean and variance, $\boldsymbol{\mu}_i$ and $\boldsymbol{\sigma}_i$, via $\boldsymbol{\xi}_i' = \boldsymbol{\xi}_i'(\boldsymbol{\mu}_i, \boldsymbol{\sigma}_i)$, and $\boldsymbol{\zeta}_i' = \boldsymbol{\zeta}_i'(\boldsymbol{\mu}_i, \boldsymbol{\sigma}_i)$

- 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, $\boldsymbol{\mu}_i$, and vector of conditional variances, $\boldsymbol{\sigma}_i$, as:

$$E(\mathbf{y}_i|\mathbf{b}_i) = \boldsymbol{\mu}_i(\boldsymbol{\beta}, \mathbf{b}_i, \mathbf{x}_i, \mathbf{z}_i) = \boldsymbol{\mu}_i(\boldsymbol{\beta}, \mathbf{b}_i); \quad i=1, \dots, n$$

$$Var(\mathbf{y}_i|\mathbf{b}_i) = \boldsymbol{\Lambda}_i(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{b}_i, \mathbf{x}_i, \mathbf{z}_i) = \boldsymbol{\Lambda}_i(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{b}_i)$$

where

- $\mathbf{y}_i = [y_1 \dots y_{p_i}]'$ is a $p_i \times 1$ vector of repeated measures
- \mathbf{x}_i and \mathbf{z}_i are vectors of covariates associated with the fixed- and random-effects parameters, respectively
- $\boldsymbol{\beta}$ is a $s \times 1$ vector of fixed-effects parameters
- $\boldsymbol{\sigma}_i = Vech[\boldsymbol{\Lambda}_i(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{b}_i)]$ is the vector of conditional variances
- $\boldsymbol{\alpha}$ is a $u \times 1$ vector of conditional covariance parameters
- \mathbf{b}_i is a $v \times 1$ vector of random-effects $\sim iid N(\mathbf{0}, \boldsymbol{\Psi})$

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 δ_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 $T_i|\mathbf{b}_i$ is assumed to be from the class of accelerated failure-time models:

$$T_i|\mathbf{b}_i = \exp\{\mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i)'\boldsymbol{\eta}\}T_{0i} \quad (2)$$

where

- $\mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i)'\boldsymbol{\eta}$ is a conditional linear function in $\boldsymbol{\eta}$ of some vector function of covariates $\mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i) = \mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i, \mathbf{x}_i, \mathbf{z}_i)$ that may depend on $\boldsymbol{\beta}$ and \mathbf{b}_i , as well as covariates \mathbf{x}_i and \mathbf{z}_i .
 - $T_{0i} = \exp(\eta_0)T_{\epsilon_i}^\phi$ are iid event times from some baseline distribution with η_0 and ϕ 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 γ
 - 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: $\lambda(t) = \lambda_0(t) \exp\{\mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i)' \boldsymbol{\eta}\}$
 - 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 $(t_0, t_1], \dots, (t_{k-1}, t_k]$
 - > Baseline Hazard rate: $\lambda_0(t) = \sum_{h=1}^k \lambda_{0h} I(t \in (t_{h-1}, t_h])$
 - > Hazard rate: $\lambda(t) = \lambda_0(t) \exp\{\mathbf{f}_{ih}(\boldsymbol{\beta}, \mathbf{b}_i, t_{h-1})' \boldsymbol{\eta}\}$
 - 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)

$$y_{ij}|\mathbf{b}_i = \beta_{1i} + \beta_{2i}t_{ij} + \epsilon_{ij}$$

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

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

AFT Model:

$$T_i|\mathbf{b}_i = \exp\{f_i(\boldsymbol{\beta}, \mathbf{b}_i)'\boldsymbol{\eta}\}T_{0i}$$

$$f_i(\boldsymbol{\beta}, \mathbf{b}_i)'\boldsymbol{\eta} = \eta_1\beta_{1i} + \eta_2\beta_{2i}$$

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

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

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

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

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

AFT Model:

$$T_i|\mathbf{b}_i = \exp\{f_i(\boldsymbol{\beta}, \mathbf{b}_i)'\boldsymbol{\eta}\}T_{0i}$$

$$f_i(\boldsymbol{\beta}, \mathbf{b}_i)'\boldsymbol{\eta} = \eta_1\beta_{1i} + \eta_2\beta_{2i}$$

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)

$$P(y_{ij} = 1 | \mathbf{b}_i) = \exp\{\beta_i\} / (1 + \exp\{\beta_i\}) = E(y_{ij} | \mathbf{b}_i)$$

$$\beta_i = (\beta_0 + b_i) + \beta_1 \text{Sex}_i + \beta_2 \text{Age}_i + \sum_{j=1}^4 \beta_{3j} I(t_j)$$

Discrete Time Failure Model:

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

$$\lambda_{ik} | \mathbf{b}_i = 1 - \exp(-\exp\{\eta_{ik} + \eta_1 \text{Sex}_i + \eta_2 \text{Age}_i + \eta_3 b_i\})$$

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

4 Estimation

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

$$L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\eta}, \phi; \mathbf{y}, T) =$$

$$\sum_{i=1}^n \log \int_{\mathbf{b}} \pi(\mathbf{y}_i | \mathbf{b}_i) \pi(T_i^* | \mathbf{b}_i)^{\delta_i} S(T_i^* | \mathbf{b}_i)^{1-\delta_i} \pi(\mathbf{b}_i) d\mathbf{b}_i$$

where $\boldsymbol{\theta} = \text{Vech}(\boldsymbol{\Psi})$ and $S(T_i^* | \mathbf{b}_i)$ 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)

Pinheiro & Bates (1995, *J Comp. and Graphical Stat*)

- Markov Chain Monte Carlo techniques (WinBUGS)

Guo and Carlin (2004, *American Statistician*) - see website.

- First-order approximations (S-Plus, R, %NLINMIX)

Lindstrom & Bates (1990, *Biometrics*) - NLME

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)

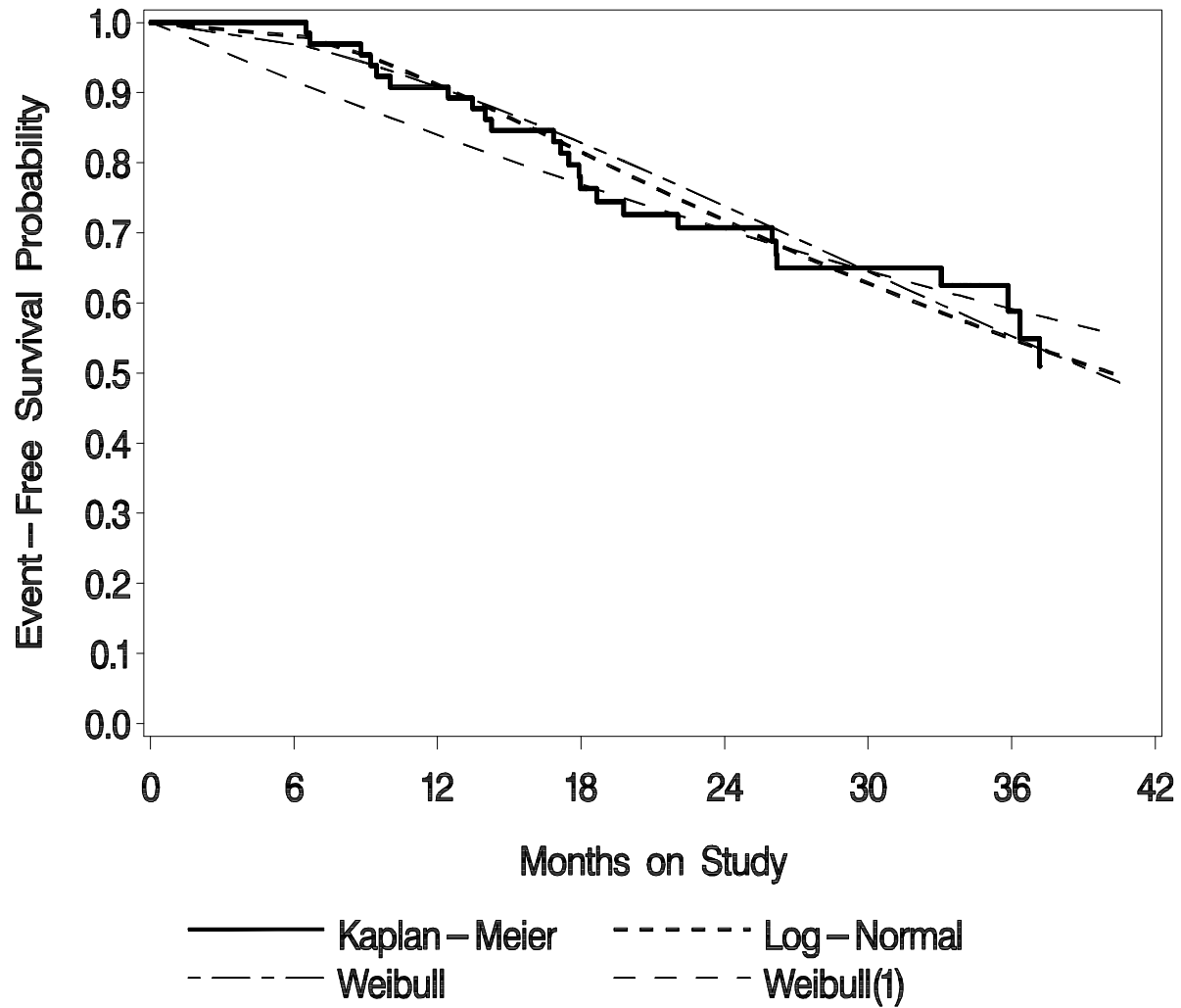
Raudenbush, Yang & Yosef (2000, *J Comp and Graph Stat*)

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)
 - $T \sim$ Weibull with hazard rate $\lambda_0(t)\exp(\sum_{k=1}^4 \eta_k G_{ki})$ - MCAR
 - $T \sim$ Weibull with hazard rate $\lambda_0(t)\exp(\eta_1\beta_{1i} + \eta_2\beta_{2i})$ - NIM
 - $T \sim$ 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

	Standard MLE	Weibull SP	Piecewise Exp SP
Parameter	$\pi(\mathbf{y}, T) = \pi(\mathbf{y})\pi(T)$	(Shape=4.87)	(6-months)
Ψ	$\begin{pmatrix} 19.9 & .061 \\ .061 & .050 \end{pmatrix}$	$\begin{pmatrix} 19.4 & .158 \\ .158 & .067 \end{pmatrix}$	$\begin{pmatrix} 19.4 & .144 \\ .144 & .064 \end{pmatrix}$
σ^2	5.23	5.14	5.13
Slopes: β_{21}	-0.251	-0.327 (30%)	-0.315 (25%)
β_{22}	-0.252	-0.296 (17%)	-0.288 (14%)
β_{23}	-0.292	-0.319 (9%)	-0.313 (7%)
β_{24}	-0.326	-0.388 (19%)	-0.382 (17%)
$L(\mathbf{y} \hat{\mathbf{b}})$	-2428.16	-2422.90	-2416.58
$-2L(\mathbf{y}, T)$	7617.1	7387.6	7406.0

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

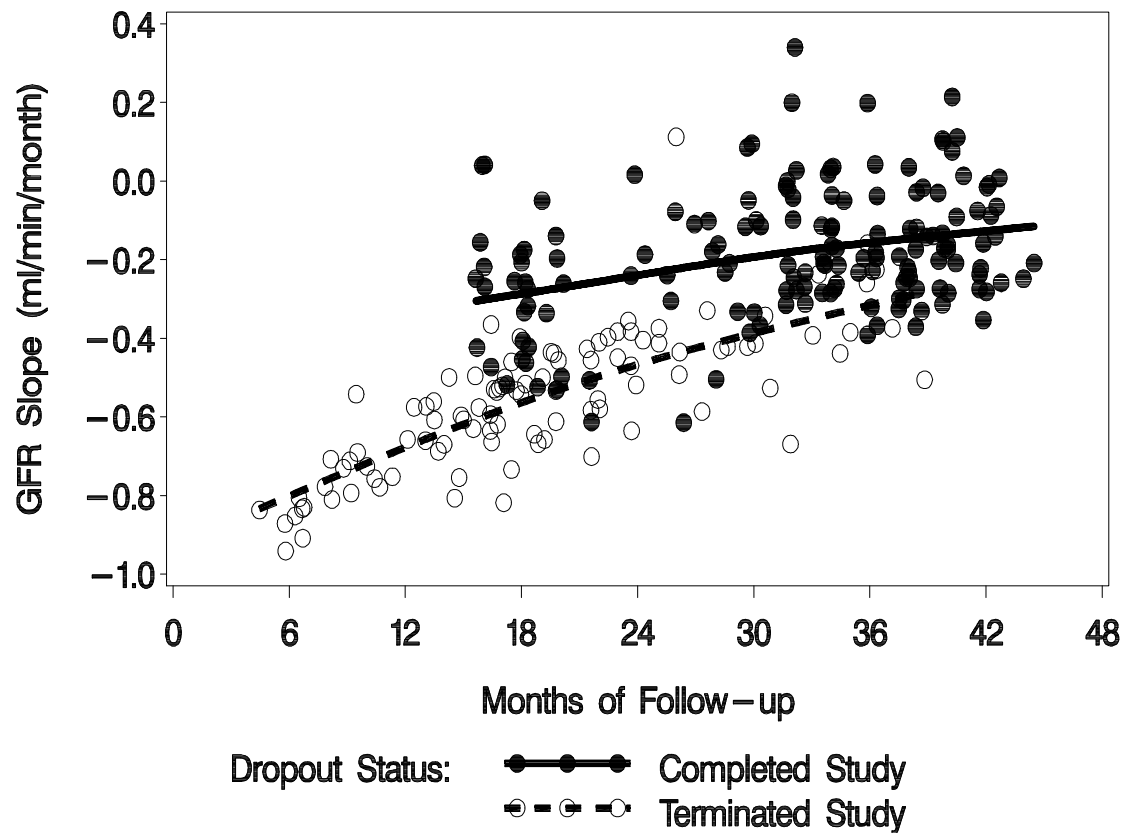


Table 2: Laplace vs Numerical Quadrature for Weibull SP Model

Parameter	QPOINTS=1	QPOINTS=10
Weibull Shape, γ	4.873 (0.677)	5.355 (0.806)
η_0^*	-3.421 (0.152)	-3.444 (0.147)
η_1^*	-0.061 (0.009)	-0.061 (0.009)
η_2^*	-2.661 (0.202)	-2.699 (0.201)
Ψ, σ^2	$\begin{pmatrix} 19.4 & .158 \\ .158 & .067 \end{pmatrix}, 5.14$	$\begin{pmatrix} 19.4 & .158 \\ .158 & .068 \end{pmatrix}, 5.15$
Intercept: Diet K, Low BP	19.62 (0.571)	19.61 (0.570)
Slope: Diet K, Low BP	-0.327 (0.035)	-0.331 (0.035)
$-2\text{Log } L(\mathbf{y}, T)$	7387.6	7386.2
CPU (hr:min:sec)	1:27:12.79	12:39:42.25

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)

ptid	ind	months	GFR	Drop	T	response
1	0	0.00	16.36	1	23.66	16.36
1	0	4.14	14.93	1	23.66	14.93
1	0	8.41	12.41	1	23.66	12.41
1	0	12.35	10.95	1	23.66	10.95
1	0	16.30	8.56	1	23.66	8.56
1	0	20.47	6.12	1	23.66	6.12
1	1	0.00	16.36	1	23.66	23.66

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			
	Lognormal [†]	Weibull ^{††}	Piecewise Exp ^{††}
Parameter	$g(T) = \log(T)$	Shape(γ)=2.965	(6-months)
Ψ	$\begin{pmatrix} 16.4 & .064 \\ .064 & .049 \end{pmatrix}$	$\begin{pmatrix} 16.3 & .073 \\ .073 & .050 \end{pmatrix}$	$\begin{pmatrix} 16.3 & .072 \\ .072 & .049 \end{pmatrix}$
σ^2	5.45	5.45	5.45
β_1 (ml/min)	19.50 (0.51)	19.51 (0.53)	19.49 (0.53)
β_2 (ml/min/mo)	-0.294 (0.044)	-0.298 (0.034)	-0.292 (0.034)
$\rho(T, b_0)$ or RR	$\rho = 0.16$	$RR(b_0) = 0.91^*$	$RR(b_0) = 0.90^*$
$\rho(T, b_1)$ or RR	$\rho = 0.74$	$RR(b_1) = 0.44^*$	$RR(b_1) = 0.48^*$

† 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.

Model	Covariates: $\mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i)'$	$\lambda(t) = \lambda_{0h} \exp\{\mathbf{f}_i(\boldsymbol{\beta}, \mathbf{b}_i)' \boldsymbol{\eta}\}$
1	(β_{1i}, β_{2i})	$\lambda_{0h} \exp\{\eta_1 \beta_{1i} + \eta_2 \beta_{2i}\}$
2	$\mu_{ih}(t_{h-1}) = \beta_{1i} + \beta_{2i} t_{h-1}$	$\lambda_{0h} \exp\{\eta_1 [\beta_{1i} + \beta_{2i} t_{h-1}]\}$
3	$y_{ih}(t_{i(h-1)})^\dagger$	$\lambda_{0h} \exp\{\eta_1 y_{ih}\}$
4	$(\beta_{1i}, \beta_{2i}, y_{ih})$	$\lambda_{0h} \exp\{\eta_1 \beta_{1i} + \eta_2 \beta_{2i} + \eta_3 y_{ih}\}$

† $y_{ih}(t_{i(h-1)})$ = last observed GFR prior to $(t_{h-1}, t_h]$. This model tests for an ignorable threshold effect and yields standard MLE's.

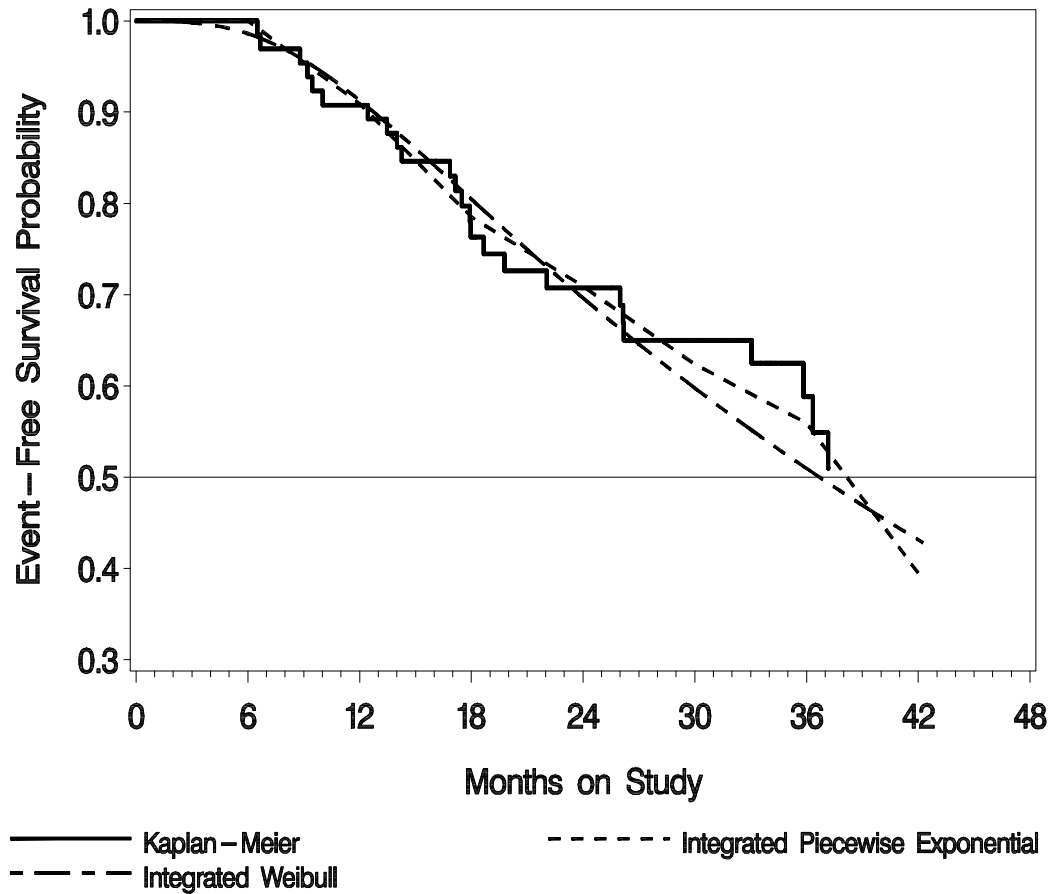
Table 3d. Piecewise Exponential SP Model for Different Covariates

Parameter	β_{1i}, β_{2i}	$\mu_{ih}(t_{h-1})$	$y_{ih}(t_{i(h-1)})$	$\beta_{1i}, \beta_{2i}, y_{ih}$
Ψ	$\begin{pmatrix} 16.3 & .072 \\ .072 & .049 \end{pmatrix}$	$\begin{pmatrix} 16.5 & .026 \\ .026 & .042 \end{pmatrix}$	$\begin{pmatrix} 16.6 & .020 \\ .020 & .038 \end{pmatrix}$	$\begin{pmatrix} 16.3 & .073 \\ .073 & .049 \end{pmatrix}$
σ^2	5.45 (0.394)	5.45 (0.395)	5.50 (0.401)	5.45 (0.394)
β_1	19.49 (0.53)	19.41 (0.53)	19.35 (0.54)	19.49 (0.53)
β_2	-0.292 (0.034)	-0.264 (0.032)	-0.244 (0.030)	-0.292 (0.034)
$RR(\beta_{1i})$	1.11 (0.074) [†]	—	—	1.11 (0.139)
$RR(\beta_{2i})$	2.08 (0.413) [†]	—	—	2.10 (0.494) [†]
$RR(y_t)$	—	1.21 (0.065) [†]	1.14 (0.050) [†]	0.99 (0.097)
$\ln L(\mathbf{y} \hat{\mathbf{b}})$	-632.7	-632.1	-634.3	-632.7
$\ln L(T \hat{\mathbf{b}})$	-94.8	-112.9	-117.2	-94.7
$AIC(\mathbf{y}, T)$	1925.2	1933.4	1937.4	1927.2

Table 4: GOF comparing Weibull to Piecewise Exp. SP model

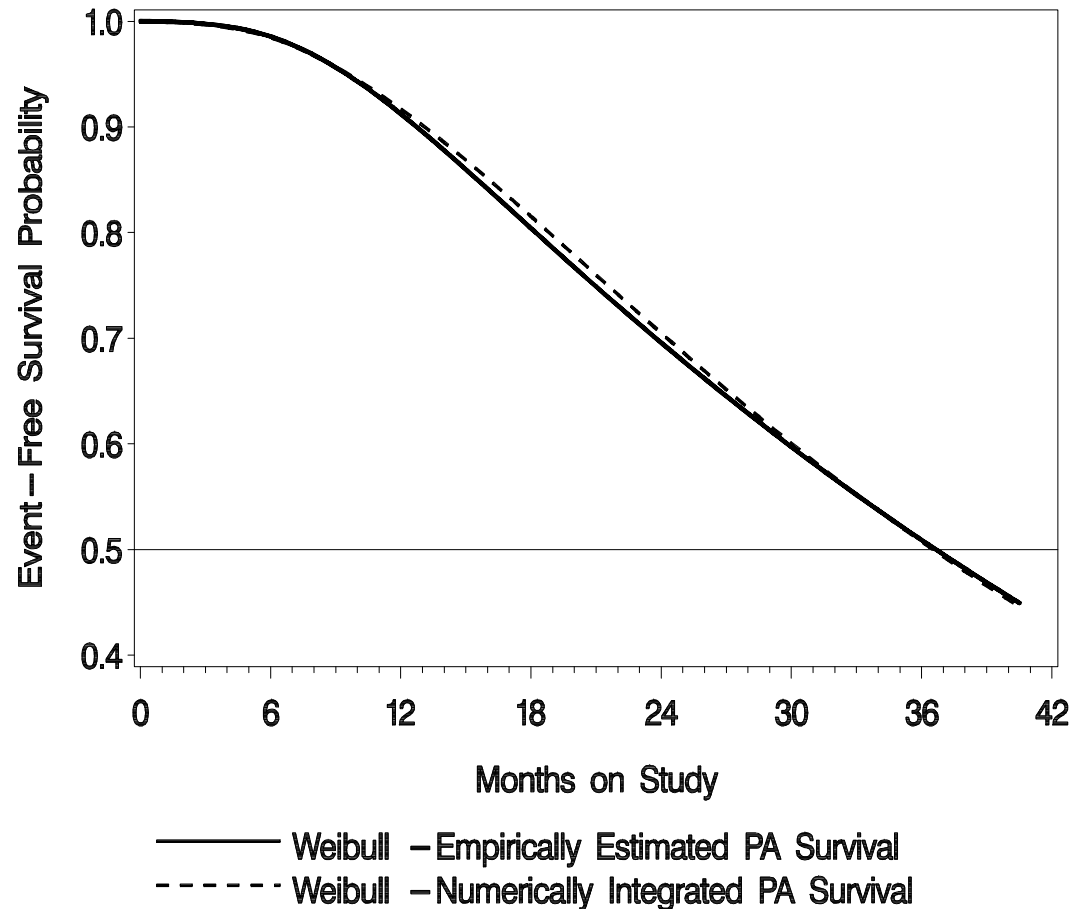
	Weibull	Piecewise Exponential	
Parameter	β_{1i}, β_{2i}	β_{1i}, β_{2i}	$y_{ih}(t_{i(h-1)})$
Ψ	$\begin{pmatrix} 16.3 & .073 \\ .073 & .050 \end{pmatrix}$	$\begin{pmatrix} 16.3 & .072 \\ .072 & .049 \end{pmatrix}$	$\begin{pmatrix} 16.6 & .020 \\ .020 & .038 \end{pmatrix}$
σ^2	5.45 (0.394)	5.45 (0.394)	5.50 (0.401)
$\beta_1(\text{ml/min})$	19.51 (0.53)	19.49 (0.53)	19.35 (0.54)
$\beta_2(\text{ml/min/month})$	-0.298 (0.034)	-0.292 (0.034)	-0.244 (0.030)
$RR(\beta_{1i})$	1.10 (0.074) [†]	1.11 (0.074) [†]	—
$RR(\beta_{2i})$	2.29 (0.513) [†]	2.08 (0.413) [†]	—
$-2\text{Log } L(\mathbf{y}, T)$	1900.6	1895.2	1909.4
$AIC(\mathbf{y}, T)$	1920.6	1925.2	1937.4

Unadjusted and Adjusted Survival Curves - Diet K, Low BP Group



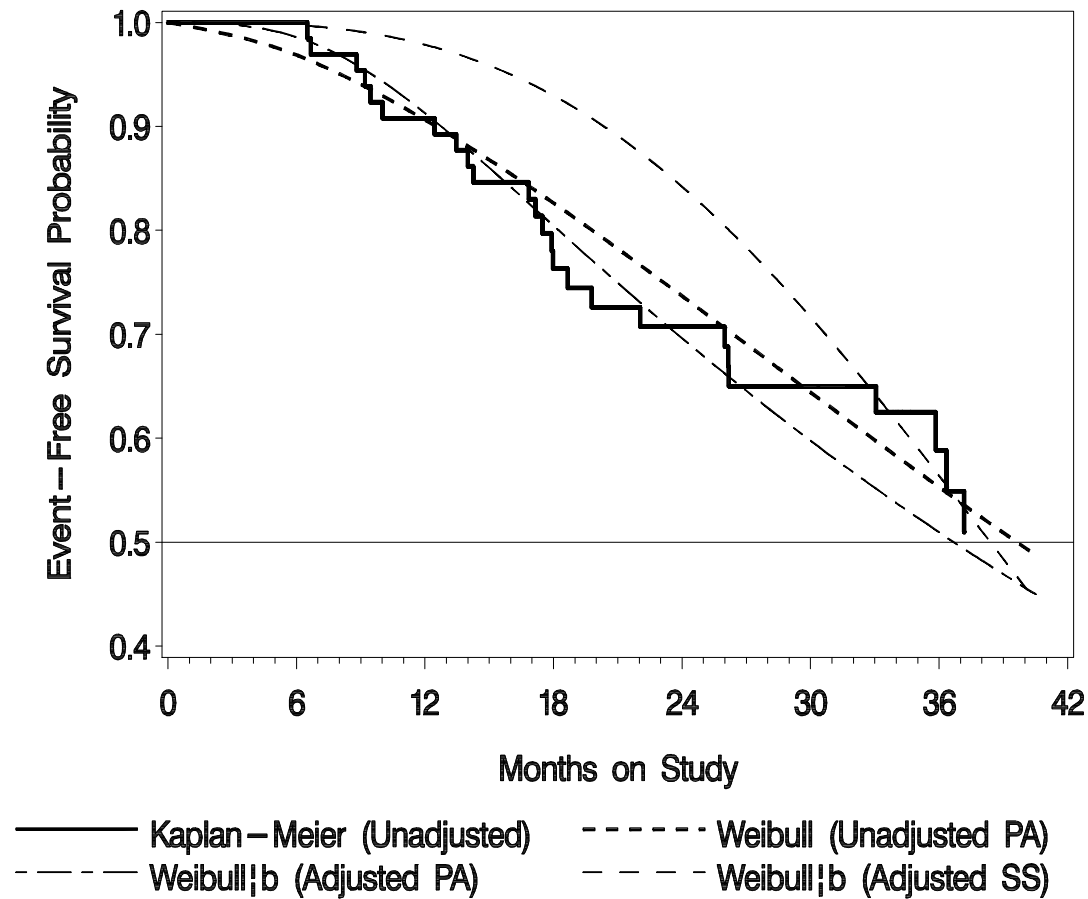
$$\hat{S}_T(t) = \frac{1}{n} \sum_{i=1}^n \hat{S}_{T|b}(t|\hat{\omega}, \hat{b}_i)$$

Empirical vs Numerically Integrated PA Survival - Diet K, Low BP



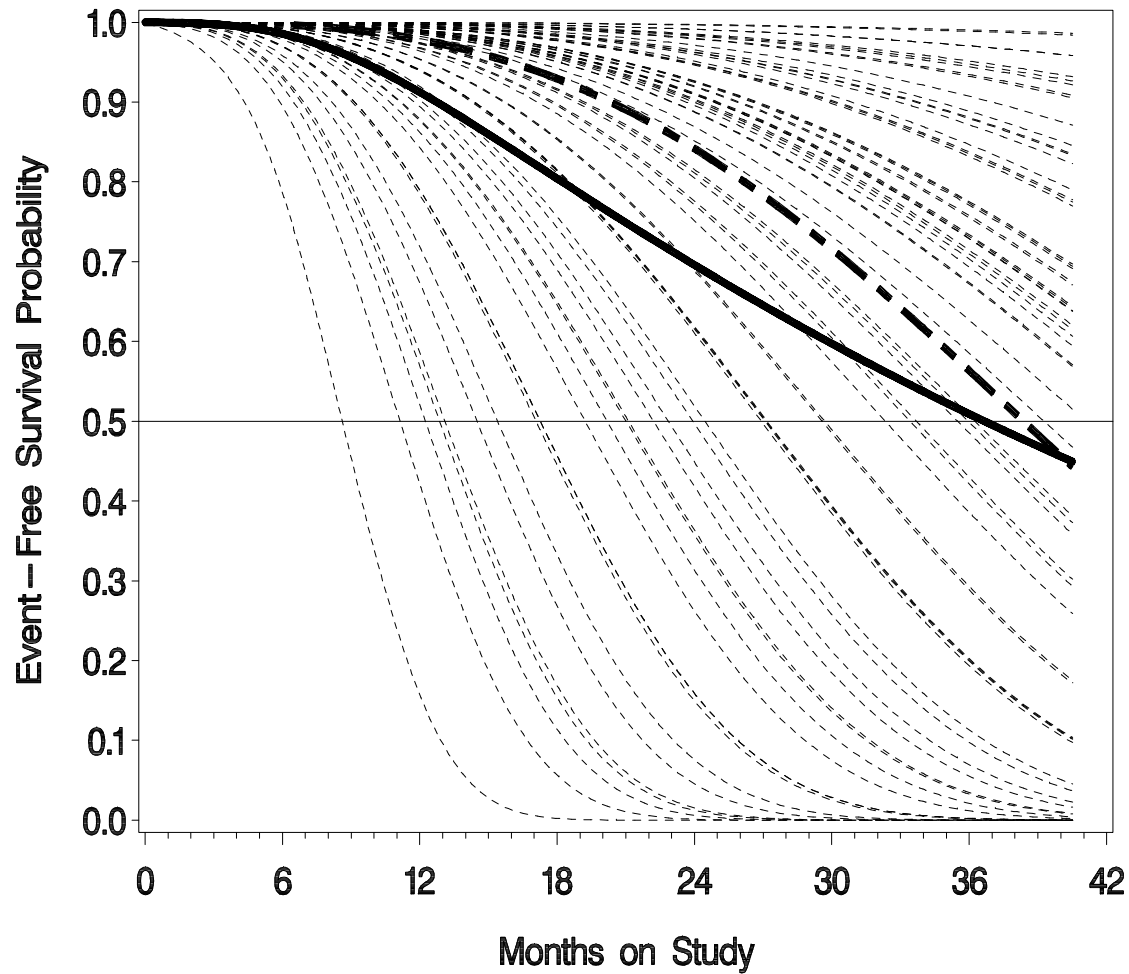
$$\hat{S}_T(t) = \frac{1}{n} \sum_{i=1}^n \hat{S}_{T|b}(t|\hat{\omega}, \hat{\mathbf{b}}_i); \tilde{S}_T(t) = \int_{\mathbf{b}} S_{T|b}(t|\hat{\omega}, \mathbf{b})\pi(\mathbf{b})d\mathbf{b}$$

Estimated Weibull Survival Curves for Diet K, Low BP Group



Adjusted PA: $\int_{\mathbf{b}} S_{T|\mathbf{b}}(t|\hat{\omega}, \mathbf{b})\pi(\mathbf{b})d\mathbf{b}$, Adjusted SS: $\hat{S}_{T|\mathbf{b}}(t|\hat{\omega}, \mathbf{b} = \mathbf{0})$.

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[†] percentages obtained via logistic regression (GEE) and shared parameter model (SPM).

			MCAR	SP Models	
Evaluation Point	N	Observed	GEE	Weibull	Discrete
Pretreatment	253	7%	7%	4%	5%
1 month	186	39%	39%	40%	40%
3 months	148	26%	28%	30%	29%
6 months	127	21%	22%	21%	20%
12 months	91	10%	13%	13%	12%

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 Ψ , the set of CGEE for $\tau = (\beta, \alpha)$ are related to the Laplacian-based estimating equations via:

$$\begin{aligned} U_{\text{LMLE}}(\tau, \hat{\mathbf{b}}) &= U_{\text{CGEE}}(\tau, \hat{\mathbf{b}}) - \frac{1}{2} \sum_{i=1}^n \left\{ \frac{\partial}{\partial \tau} \log \left(\left| -L_i''(\tau, \hat{\mathbf{b}}_i(\tau)) \right| \right) \right\} \\ &= U_{\text{CGEE}}(\tau, \hat{\mathbf{b}}) - \frac{1}{2} \sum_{i=1}^n \left\{ \frac{\partial}{\partial \tau} \log \left(\left| -l_i''(\tau, \hat{\mathbf{b}}_i(\tau)) + \Psi^{-1} \right| \right) \right\} \end{aligned}$$

where

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

$$l_i(\tau, \mathbf{b}_i) = l_i(\beta, \alpha, \mathbf{b}_i; \mathbf{y}_i) = \log[\pi(\mathbf{y}_i | \mathbf{b}_i)]$$

$$L_i''(\tau, \hat{\mathbf{b}}_i(\tau)) = \frac{\partial^2}{\partial \mathbf{b}_i \partial \mathbf{b}_i'} \left\{ l_i(\tau, \mathbf{b}_i) - \frac{1}{2} \mathbf{b}_i' \Psi^{-1} \mathbf{b}_i \right\} \Big|_{\mathbf{b}_i = \hat{\mathbf{b}}_i(\tau)}$$

Pearson Residuals for GFR versus Time-to-Dropout
Model 1

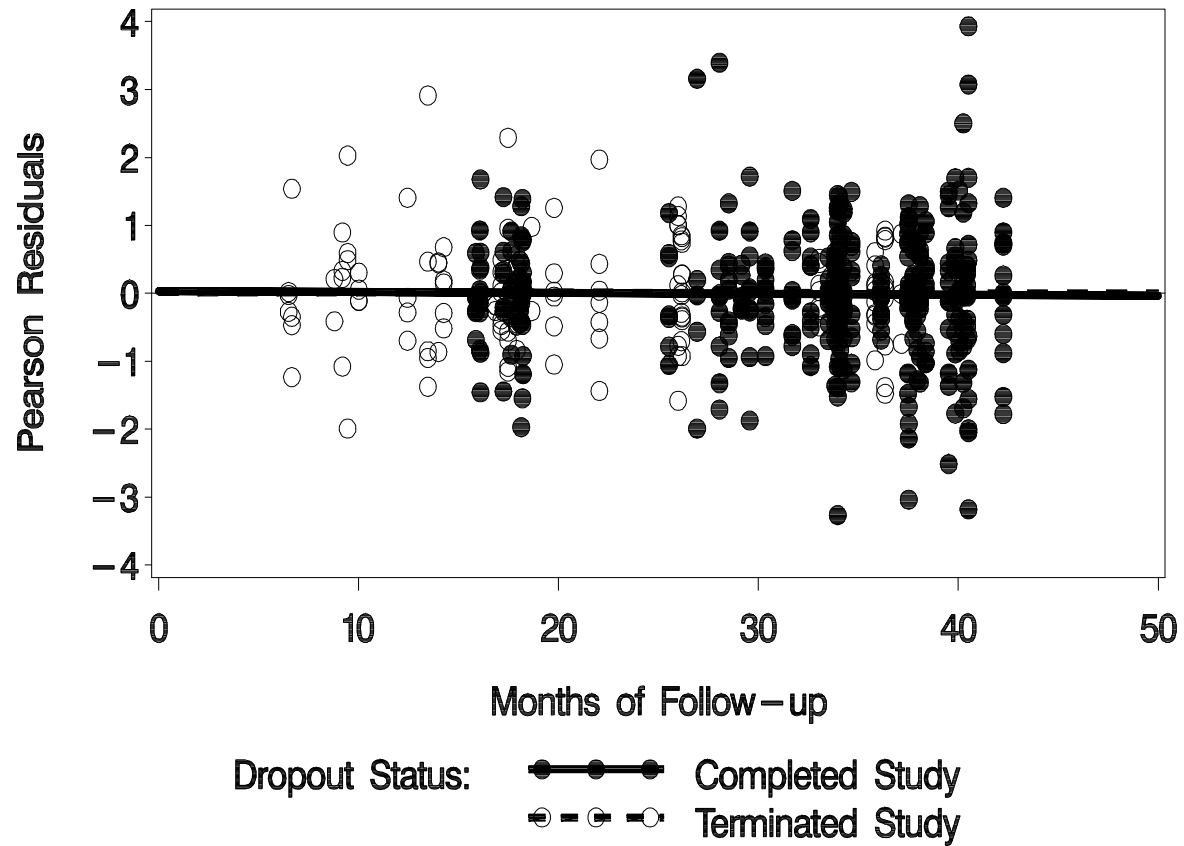


Table 1: Select estimates from a SP model for head & neck cancer data using a discrete time survival model for time to dropout.

Outcome	Parameter	Laplace Approx.	Adaptive Gaussian
y_{ij}	$\beta_0(\textit{intercept})$	11.802 (3.455) [†]	9.057 (2.334) [†]
	$\beta_1(\textit{sex})$	1.528 (0.989)	1.159 (0.657)
	$\beta_2(\textit{age})$	-0.127 (0.045) [†]	-0.101 (0.031) [†]
	ψ	24.236 (12.861)	9.195 (2.637)
T_i	$\lambda_1(\textit{sex})$	0.088 (0.205)	0.086 (0.206)
	$\lambda_2(\textit{age})$	0.004 (0.009)	0.004 (0.009)
	$\lambda_3(b_i)$	-0.066 (0.030) [†]	-0.106 (0.046) [†]
	$-2\textit{Log } L(\mathbf{y}, T)$	1189.6	1196.4
	CPU (hr:min:sec)	0:9:19.00	0:20:54.31