An introduction to frailty models

Methodology considerations for building prognostic models for esophageal cancer using SEER-Medicare database

Xi Kathy Zhou, Ph.D.
Division of statistics and epidemiology
Department of public health
Weill Cornell Medical College
Outline

- Background about Esophageal Cancer
  - Need prognostic model
  - Need SEER-Medicare data
- Motivation through SEER-Medicare data
  - Need to deal with Clustered Survival Data
- A brief review of survival data analysis
- Methods for Clustered Survival Data with a focus on frailty models
- Statistical Simulation to illustrate the usefulness of the frailty models
Motivation and Background
Esophageal Cancer

- 8th most common cancer and 6th leading cause of cancer death worldwide.
- Increasing incidence of esophageal adenocarcinoma at a 2% annual rate in western countries.
- Poor prognosis in general: ~10% 5-year survival.
- Prognostic factors:
  - Stage, age, gender, number of positive lymph nodes
Controversies in treating esophageal cancer

- Surgery
  - only curative approach
  - high risk
  - great variation (esp. extensiveness)

- Regionalization of treatment
  - based on strong volume and short-term outcome relationship
Issues of interest

- Long term implication of the surgical treatment – the prognostic role of the number of dissected lymph nodes
- Long term implication of the volume outcome relationship
- The extent of heterogeneity in long term outcome may be attributed to providers

Require multi-center clinical trial data or large observational databases.
SEER-MEDICARE Database

- Have all information available in SEER
  - demographics, diagnosis, primary treatment, survival
- Added information from Medicare
  - comorbidity, treatments sequence, provider information
- Has been used to build a short-term surgical mortality score for patients with esophageal cancer.
Methodological challenge

- Primary outcome: censored survival times.
- Data structure: naturally clustered.
- Covariates: individual level & cluster level

Consequences of failing to take the clustered data structure into account:

- Regression parameters may be biased (Hougaard, 2000)
- Wrong effects assessment (Wintrebert, et al. 2004)

A model for clustered survival data allows for evaluation of cluster level covariates is needed
Survival Data Analysis Review
Cornell Esophageal Cancer Data

- Consecutive patients between 1987-2006
- N=264 treated with surgery only
- Median follow up: 4.1 years
- Overall 5-year survival: 37.9% yrs (95% CI = (31.8%, 45.1%))

- Main objective:
  Comparison of surgical procedures
Notation

- \( i=1, \ldots, N \) subjects
- \( T_i \) : survival time for subject \( i \)
- \( C_i \) : time to censoring for subject \( i \)
- \( X_i = \min(T_i, C_i) \) observed survival time for \( i \)
- \( Z_i \) : vector of risk factors for subject \( i \)
  - stage, age, gender, tumor size, tumor differentiation, surgical procedure, number of positive lymph nodes
Describe Univariate Survival Time

- Cumulative distribution function: \( F(t) = P(T \leq t) \)
- Density function: \( f(t) = F'(t) \)
- Survivor function: \( S(t) = 1 - F(t) = \Pr(T > t) \)
- Hazard function: \( \lambda(t) = \frac{f(t)}{S(t)} \)

Equivalent

Due to censoring, commonly working with the hazard function \( \lambda(t) \)
Modeling covariate effect

Proportional hazard model

\[ \lambda(t \mid Z) = \lambda_0(t) \exp(\beta^T Z) \]

\( \lambda_0(t) \): baseline hazard

\( \lambda_0(t)=1 \): exponential PH model

\( \lambda_0(t)=\alpha t^{\alpha-1} \): weibull PH model

\( \lambda_0(t)=\alpha_j I(a_{j-1} < t \leq a_j) \): piecewise exponential PH model

\( \lambda_0(t) \) unspecified: Cox PH model
Parameter Estimation

- Partial likelihood:

$\prod_{i: \text{Cases}} \frac{\exp(X_i \beta)}{\sum_{j \in R_i} \exp(X_j \beta)}$

Widely available in statistical packages, SAS, R/Splus, Stata, etc.

- Maximum likelihood approach:
Generalized linear regression for poisson data
Estimating Model Parameters: Bayesian Approach

- Utilize full likelihood
- Need prior specification for model parameters
- Semiparametric models require specifying a prior random process to generate a distribution function for survival times
- MCMC based methods for generating samples from joint posterior distribution
- BUGS can be used
Cornell Data: Predictors

- **Stage (invasiveness):**
  - 1 (28.4%); 2 (18.2%); 3 (50.8%); 4 (2.6%)
- **Age:**
  - median: 64 range: 29-88
- **Gender:** Male (78.4%)
- **Procedure:**
  - non-invasive (34.8%); 2-field (30.7%); 3-field (34.5%)
- **Tumor size (cm):**
  - median 3.5 (range 0-12.5)
- **No. of positive lymph nodes:**
  - 0 (40.9%), 1-5 (36.0%), 6-34 (23.1%)
- **Tumor differentiation:**
  - Well (10.2%), moderate (42.4%), poor (44.3%)
<table>
<thead>
<tr>
<th>Variables</th>
<th>Hazard Ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tumor stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>0.002</td>
</tr>
<tr>
<td>II</td>
<td>2.65 (1.42, 4.96)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>III</td>
<td>3.49 (1.97, 6.17)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>IV</td>
<td>5.22 (1.89, 14.43)</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Tumor differentiation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>1</td>
<td>0.84</td>
</tr>
<tr>
<td>Moderate</td>
<td>1.04 (0.73, 1.46)</td>
<td>0.07</td>
</tr>
<tr>
<td>Good</td>
<td>0.53 (0.27, 1.05)</td>
<td></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>1</td>
<td>0.02</td>
</tr>
<tr>
<td>M</td>
<td>1.64 (1.07, 2.51)</td>
<td></td>
</tr>
<tr>
<td><strong>Surgical procedure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-radical</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Radical 2-field</td>
<td>0.53 (0.34, 0.81)</td>
<td>0.004</td>
</tr>
<tr>
<td>Radical 3-field</td>
<td>0.44 (0.29, 0.67)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Age at surgery</strong></td>
<td>1.02 (1.01, 1.04)</td>
<td>0.01</td>
</tr>
<tr>
<td><strong>Tumor size</strong></td>
<td>1.04 (0.96, 1.13)</td>
<td>0.35</td>
</tr>
<tr>
<td><strong>No. of positive LN</strong></td>
<td>1.04 (1.01, 1.08)</td>
<td>0.005</td>
</tr>
</tbody>
</table>
Summary

- Extensiveness of surgery has great implication for survival
  - More extensive surgery, better survival

- Difficult to generalize the results (single institution study, 4 surgeons)
Methods for Clustered Survival Data
Types of Clustered Survival Data

- Multiple event time data
  - Each subject experiences multiple failures, recurrent events in bladder cancer
  - Dependence structure is induced serially

- Individuals under the study are naturally clustered
  - paired data, family data
  - dependence structure is induced in a parallel manner
SEER-Medicare Data on Esophageal Cancer

- Right censored time to event data
- Differences in patient selection, treatment, post treatment care may induce heterogeneity in outcome across providers
- Cluster size varies greatly across providers
- Include both individual covariates and cluster level covariates
- Observational database – nonrandom treatment assignment, treatment effect may be different for different providers
Notation

- $i=1,\ldots, n_k$
- $k=1, \ldots, K$
- $T_{ik}$: survival time for subject $i$ in cluster $k$
- $Z_{ik}$: individual level covariates
- $V_k$: cluster level covariates

Key feature in the data:
- $(T_{1k}, \ldots, T_{nk})$ dependent
Problems to be dealt with

- How to modeling association structure
  - For accurate covariate estimation
  - To assess the amount of heterogeneity
- How to make inference
  - Possible complicated semiparametric estimation
- How to assess model fit
  - Model selection and diagnosis
Modeling the within-cluster dependence
Under proportional hazard framework
Within-cluster dependence

- Not directly observed
- Shared by subjects within the cluster
  - Could be described by a latent cluster level covariate $\xi_k$
- Different across clusters
  - Hazard function should be different for different clusters
Fixed Effects Approach

$$\lambda_{ik}(t \mid Z_{ik}, \xi_k) = \lambda_0(t) \exp(\beta^T Z_{ik} + \xi_k)$$

- Subject Specific regression
- The latent covariate $\xi$ is treated as fixed effects
- Affect the hazard proportionally
- Easy calculation conceptually, available methods

- require large cluster size for good estimation
- Large number of parameters, may encounter computational problem
Stratified PH Model

\[ \lambda_{ik}(t \mid Z_{ik}) = \lambda_{0k}(t) \exp(\beta^T Z_{ik}) \]

- Subject specific regression
- Allows baseline hazard different for each cluster (no proportionality assumed)
- No distribution assumed for the difference in baseline hazard
- Easy calculation
- Parameter estimation is based on intra-cluster comparisons

- Can not estimate effects due to cluster level covariate (no variation within strata)
- Can not assess the amount of heterogeneity across clusters
Frailty PH Model

\[ \lambda_{ik}(t | Z_{ik}, \xi_k) = \xi_k \lambda_0(t) \exp(\beta^T Z_{ik}) \]

- Subject specific regression
- Latent covariate \( \xi_k \) is treated as random effect
- Could be generated from various distributions
- Affect the hazard proportionally
- Much smaller number of parameters
- Allows for modeling cluster level covariates
- Allows for estimating between cluster heterogeneity
- Can be very flexible

- Computationally can be challenging
- Still under active development
Conditional Models

- Fixed effects approach: using indicator variables for each cluster
- Stratified approach where each center is treated as a strata
- frailty models (Random effects models for survival outcome)
Marginal Model

(Wei, Lin and Weissfeld 1988, Cai 1992)

- Population average model
- Correlation structure completely unspecified
- within-cluster correlation is adjusted through the robust sandwich type estimator
- No estimate for heterogeneity
- Covariate estimate interpretation: average difference between any two individuals across cluster.
- useful for assess the significance of a particular covariate
- not very useful for prediction or for assessing cluster effects
- Statistical programs available in R/Splus, SAS, Stata
Frailty PH Models
Early developments of frailty models

- 1978: Clayton
  - Bivariate data without covariates
  - Gamma frailty model

- 1979: Vaupel et al first used the term “frailty”
  - Gamma frailty
  - Account for the unobserved population heterogeneity

- Hougaard (1986)
  - Positive stable frailty model

- Oakes (1989)
  - Other choices of frailty distributions
Understanding the frailties

\[ \lambda_{ik}(t \mid Z_{ik}, \xi_k) = \xi_k \lambda_0(t) \exp(\beta^T Z_{ik}) \]

Frailties ($\xi_k$)
- Unobserved, latent
- independent of the measured covariates
- affect the average baseline risk multiplicatively
- specifies the heterogeneity of the population

- independent across clusters
- shared by members within a cluster

- distributional choices were mainly based on computational convenience
- frailties are scaled to have mean 1, for identifiability
Heterogeneity & Dependence

- Frailty variation characterizes the heterogeneity across clusters
- Frailty induces dependence within a cluster
- Methods to quantifying the unconditional dependence in survival times
  - Kendall’s tau (a global measure of dependence)
- Dependence is different for different frailty distribution specification
Commonly used frailty models
Gamma frailty model

\[ \xi_k \sim G(\theta^{-1}, \theta^{-1}) \]

- Mean(\(\xi_k\))=1, Var(\(\xi_k\))=\(\theta\)
- Intraclass correlation: \(\tau=\theta/(2+\theta)\)
- Cross ratio: \(1+\theta\)
- Larger \(\theta\), stronger dependence

- Marginal hazard ratio due to covariate becomes time-dependent
Positive Stable Frailty
(Hougaard, 1986)

\[ \xi_k \sim S_\alpha (\sigma = 1, \beta = 1, \delta = 0) \]

Stable distribution:
- \( \alpha \): stability parameter \((0,2]\), degree of peakedness and heaviness of tail
- \( \sigma \): scale parameter
- \( \beta \): skewness parameter \([-1,1]\)
- \( \delta \): location parameter

Positive stable:
- \( \alpha \) in \((0,1)\), \( \beta = 1, \delta = 0 \)
- \( \sigma = 1 \) for identifiability

- No closed form density function
- Characterized by its Laplace transformation
Mean($\xi_k$)=1, Var($\xi_k$)=inf

Kendall’s $\tau=1-\alpha$

Cross ratio:
- complex
- time varying
- converge to 1 as time increases

The only frailty model maintained the marginal proportionality of the covariates
Log Normal Frailty Model

\[ \varepsilon_k = \log(\xi_k) \sim N(0, \sigma^2) \]

Mean(\(\varepsilon_k\))=0, Var(\(\varepsilon_k\))=\(\sigma^2\)

No simple explicit form for dependence measures
Extensions of simple frailty models

Multivariate frailty models:
- Could arise if there is subject-cluster interaction for some covariates

Time-dependent frailty models
- Paik: piecewise gamma frailty model
- Dunson: Bayesian semiparametric dynamic frailty model
Estimation & Inference under frailty models

- More complicated likelihood
- Partial likelihood approach no longer workable
  - Need to deal with baseline hazard
- Computational difficulty is different for different frailty models
EM Algorithm (Nielsen)

- Conceptually simple, view the frailties as missing data
- E-step: estimate the expected frailty for each cluster (simple expression for gamma frailty models)
- M-step: obtain the maximum likelihood estimate for the covariates and baseline hazard by treating frailties as an offset

- Slow convergence, 30-50 steps
- SAS Macros for gamma & positive frailty models
Pennalized Partial Likelihood
(Therneau & Grambsch, 2003)

- Rewrote the conditional likelihood
- Maximizes
  - log likelihood - \( \eta \ p(\text{parameters}) \)
  - \( P() \) is a penalty function, \( \eta \) is a smoothing parameter
- fast calculation
- yield MLE for gamma frailties
- Available in R/Splus for gamma & log-normal frailty models
Bayesian Methods

- Working with joint posterior density of parameters given data
- Need to specify a prior process for baseline in semiparametric setting (gamma process, beta process, dirichlet process)
- Need to specify prior distributions for parameters or hyperparameters
- MCMC/Gibbs sampling approach used to update the sample
- Can be readily implemented for gamma, log-normal, inverse-gaussian frailty models in BUGS
- Allow easy extension for some more complicated frailty models
Software availability

- **R/ Splus:**
  - gamma & log normal frailty models (PPL)

- **SAS:**
  - gamma & positive stable frailty models (EM algorithm, slower)
  - Score tests of independence of failures

- **Stata:**
  - gamma, inverse gaussian frailty models

- **BUGS:**
  - gamma, log normal, inverse-gaussian, some time-dependent, multivariate, additive frailty models
Illustration / Comparison of methods using simulated data

- Data generated using parameters estimated from WCMC database
- Number of subjects N=2000
- Exponential PH model
- Censoring rate: 30%
- Gamma frailty:
  - strong dependence: Gamma(0.5, 0.5), Kendall’s $\tau=0.5$
  - weak dependence: Gamma(20, 20), Kendall’s $\tau =0.02$
- number of clusters K=200
- Cluster size:
  - fixed: 10/cluster
  - varying: small volume providers 130
Simulated data: covariates

Individual level covariates:

- Tumor stage
  - 1(25%), 2(20%), 3(45%), 4(10%)
  - HR: 1:2.5:3.5:5

- Gender:
  - F (25%), M (75%)
  - HR: 1:1.2

- Number of positive lymph node:
  - rounded gamma variable
  - HR: 1.04

Cluster level covariate:

- Fixed cluster size: Quality (Low vs High: 1.10)
- Varying cluster size: Volume (0.98)

Methods used:
No adjustment, Marginal, stratified, gamma frailty, log normal frailty
Simulation result: Individual level covariate estimates
(Strong dependence, fixed cluster size)

<table>
<thead>
<tr>
<th></th>
<th>Noadjustment</th>
<th>Marginal</th>
<th>stratified</th>
<th>Gamma Frailty</th>
<th>LogNorm Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias</strong></td>
<td>-0.440</td>
<td>-0.440</td>
<td>0.045</td>
<td>0.019</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>-0.859</td>
<td>-0.859</td>
<td>0.050</td>
<td>-0.001</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>-1.514</td>
<td>-1.514</td>
<td>0.108</td>
<td>0.027</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>-0.048</td>
<td>-0.048</td>
<td>0.001</td>
<td>-0.001</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>-0.010</td>
<td>-0.010</td>
<td>0.000</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td><strong>RMSE</strong></td>
<td>0.472</td>
<td>0.472</td>
<td>0.251</td>
<td>0.223</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>0.884</td>
<td>0.884</td>
<td>0.324</td>
<td>0.287</td>
<td>0.287</td>
</tr>
<tr>
<td></td>
<td>1.564</td>
<td>1.564</td>
<td>0.633</td>
<td>0.585</td>
<td>0.580</td>
</tr>
<tr>
<td></td>
<td>0.089</td>
<td>0.089</td>
<td>0.086</td>
<td>0.078</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td>0.014</td>
<td>0.014</td>
<td>0.011</td>
<td>0.011</td>
<td>0.010</td>
</tr>
<tr>
<td><strong>95% CI Coverage</strong></td>
<td>0.35</td>
<td>0.34</td>
<td>0.95</td>
<td>0.95</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td>0.05</td>
<td>0.94</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.07</td>
<td>0.08</td>
<td>0.96</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.86</td>
<td>0.86</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>0.80</td>
<td>0.81</td>
<td>0.95</td>
<td>0.95</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Simulation result: Individual level covariate estimates
(Weak dependence, fixed cluster size)

<table>
<thead>
<tr>
<th></th>
<th>Noadjustment</th>
<th>Marginal</th>
<th>Stratified</th>
<th>Gamma Frailty</th>
<th>LogNorm Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-0.050</td>
<td>-0.050</td>
<td>0.022</td>
<td>-0.001</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>-0.117</td>
<td>-0.117</td>
<td>0.024</td>
<td>-0.018</td>
<td>-0.002</td>
<td></td>
</tr>
<tr>
<td>-0.219</td>
<td>-0.219</td>
<td>0.024</td>
<td>-0.040</td>
<td>-0.010</td>
<td></td>
</tr>
<tr>
<td>0.002</td>
<td>0.002</td>
<td>0.010</td>
<td>0.007</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>-0.001</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td><strong>RMSE</strong></td>
<td>0.223</td>
<td>0.253</td>
<td>0.227</td>
<td>0.228</td>
<td></td>
</tr>
<tr>
<td>0.282</td>
<td>0.282</td>
<td>0.311</td>
<td>0.279</td>
<td>0.277</td>
<td></td>
</tr>
<tr>
<td>0.545</td>
<td>0.545</td>
<td>0.609</td>
<td>0.534</td>
<td>0.534</td>
<td></td>
</tr>
<tr>
<td>0.077</td>
<td>0.077</td>
<td>0.086</td>
<td>0.079</td>
<td>0.079</td>
<td></td>
</tr>
<tr>
<td>0.010</td>
<td>0.010</td>
<td>0.011</td>
<td>0.010</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td><strong>95% CI</strong></td>
<td><strong>Coverage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.96</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>0.93</td>
<td>0.93</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>0.92</td>
<td>0.91</td>
<td>0.95</td>
<td>0.94</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>0.94</td>
<td>0.93</td>
<td>0.96</td>
<td>0.94</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>0.95</td>
<td>0.93</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>
Simulation result: Individual level covariate estimates (Strong dependence, varying cluster size)

<table>
<thead>
<tr>
<th></th>
<th>Noadjustment</th>
<th>Marginal</th>
<th>stratified</th>
<th>Gamma Frailty</th>
<th>LogNorm Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias</strong></td>
<td>-0.910</td>
<td>-0.910</td>
<td>0.036</td>
<td>0.017</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>-1.651</td>
<td>-1.651</td>
<td>0.019</td>
<td>-0.015</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>-2.807</td>
<td>-2.870</td>
<td>0.087</td>
<td>0.011</td>
<td>-0.033</td>
</tr>
<tr>
<td></td>
<td>-0.107</td>
<td>-0.107</td>
<td>-0.001</td>
<td>-0.003</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>-0.021</td>
<td>-0.021</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>RMSE</strong></td>
<td>0.921</td>
<td>0.921</td>
<td>0.294</td>
<td>0.263</td>
<td>0.267</td>
</tr>
<tr>
<td></td>
<td>1.656</td>
<td>1.656</td>
<td>0.339</td>
<td>0.307</td>
<td>0.323</td>
</tr>
<tr>
<td></td>
<td>2.819</td>
<td>2.819</td>
<td>0.716</td>
<td>0.642</td>
<td>0.657</td>
</tr>
<tr>
<td></td>
<td>0.130</td>
<td>0.130</td>
<td>0.087</td>
<td>0.079</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>0.023</td>
<td>0.023</td>
<td>0.012</td>
<td>0.012</td>
<td>0.012</td>
</tr>
<tr>
<td><strong>95% CI Coverage</strong></td>
<td>0.004</td>
<td>0.004</td>
<td>0.940</td>
<td>0.928</td>
<td>0.930</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.948</td>
<td>0.932</td>
<td>0.928</td>
</tr>
<tr>
<td></td>
<td>0.000</td>
<td>0.000</td>
<td>0.926</td>
<td>0.906</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td>0.682</td>
<td>0.680</td>
<td>0.956</td>
<td>0.962</td>
<td>0.962</td>
</tr>
<tr>
<td></td>
<td>0.468</td>
<td>0.482</td>
<td>0.936</td>
<td>0.940</td>
<td>0.934</td>
</tr>
</tbody>
</table>
Simulation result: Individual level covariate estimates
(Weak dependence, varying cluster size)

<table>
<thead>
<tr>
<th></th>
<th>Noadjustment</th>
<th>Marginal</th>
<th>stratified</th>
<th>Gamma Frailty</th>
<th>LogNorm Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias</strong></td>
<td>-0.393</td>
<td>-0.393</td>
<td>0.022</td>
<td>-0.001</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>-0.752</td>
<td>-0.752</td>
<td>0.032</td>
<td>-0.008</td>
<td>-0.022</td>
</tr>
<tr>
<td></td>
<td>-1.373</td>
<td>-1.373</td>
<td>0.078</td>
<td>-0.009</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>-0.046</td>
<td>-0.046</td>
<td>-0.001</td>
<td>-0.002</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>-0.008</td>
<td>-0.008</td>
<td>0.001</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>RMSE</strong></td>
<td>0.442</td>
<td>0.442</td>
<td>0.267</td>
<td>0.246</td>
<td>0.246</td>
</tr>
<tr>
<td></td>
<td>0.785</td>
<td>0.785</td>
<td>0.325</td>
<td>0.292</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>1.434</td>
<td>1.434</td>
<td>0.669</td>
<td>0.594</td>
<td>0.591</td>
</tr>
<tr>
<td></td>
<td>0.092</td>
<td>0.092</td>
<td>0.092</td>
<td>0.086</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>0.013</td>
<td>0.013</td>
<td>0.012</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td><strong>95% CI</strong></td>
<td>0.562</td>
<td>0.560</td>
<td>0.966</td>
<td>0.950</td>
<td>0.948</td>
</tr>
<tr>
<td>Coverage</td>
<td>0.168</td>
<td>0.160</td>
<td>0.960</td>
<td>0.952</td>
<td>0.950</td>
</tr>
<tr>
<td></td>
<td>0.140</td>
<td>0.144</td>
<td>0.940</td>
<td>0.938</td>
<td>0.940</td>
</tr>
<tr>
<td></td>
<td>0.890</td>
<td>0.892</td>
<td>0.956</td>
<td>0.942</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>0.880</td>
<td>0.874</td>
<td>0.934</td>
<td>0.934</td>
<td>0.934</td>
</tr>
</tbody>
</table>
Simulation result: cluster level covariate estimates

<table>
<thead>
<tr>
<th></th>
<th>Noadjustment</th>
<th>Marginal</th>
<th>stratified</th>
<th>Gamma Frailty</th>
<th>LogNorm Frailty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Strong.fixed</td>
<td>-0.047</td>
<td>-0.047</td>
<td>NA</td>
<td>0.013</td>
<td>0.026</td>
</tr>
<tr>
<td>Weak.fixed</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>Strong.vary</td>
<td>0.010</td>
<td>0.010</td>
<td></td>
<td>0.000</td>
<td>-0.007</td>
</tr>
<tr>
<td>Weak.vary</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>RMSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.153</td>
<td>0.153</td>
<td>NA</td>
<td>0.228</td>
<td>0.287</td>
</tr>
<tr>
<td></td>
<td>0.070</td>
<td>0.070</td>
<td></td>
<td>0.072</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>0.014</td>
<td>0.014</td>
<td></td>
<td>0.011</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>0.003</td>
<td>0.003</td>
<td></td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td><strong>95% CI Coverage</strong></td>
<td>0.52</td>
<td>0.93</td>
<td>NA</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.94</td>
<td></td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>0.18</td>
<td>0.70</td>
<td></td>
<td>0.92</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>0.90</td>
<td>0.93</td>
<td></td>
<td>0.96</td>
<td>0.97</td>
</tr>
</tbody>
</table>
Simulation summary: individual level covariates

- Marginal model provide almost the same estimates as the model without any adjustment
- Estimated effects attenuated in marginal models (frailty selection)
- Attenuation in effects weaker when dependence is weaker

- Stratified approach has good performance when the cluster size is large

- Frailty models offer the best estimates
  - miss specifying frailty distribution have small effect on estimates

- Overall, balanced cluster size retains more information and better estimates

- Frailty models have good performance in general, even when there is no intra-class correlation, or frailty distribution is miss-specified
Simulation summary: cluster level covariates

- Models without adjustment to clustered data have poorer 95% CI coverage (prone to spurious finding).
- When there is strong dependence, marginal models perform slightly better than frailty models in terms of bias and RMSE, but comparable coverage.
- Frailty models have better performance when the cluster size varies.
- When there is weak dependence, frailty models have comparable performance as the marginal model.
Future research

- Application to SEER-Medicare data
  - data could be more complicated
  - standard PH frailty models, additive frailty models, time dependent frailty model, multivariate frailty model
  - model diagnosis, model selection, model validation
Conclusions

- Frailty models would be a useful tool for modeling survival data from SEER-Medicare Database
- Gamma frailty model is well developed, tested and widely available
- Developments is ongoing for other frailty models, their extensions, model diagnosis and model selection
- Involving complex computation
- BUGS program offers greatest flexibility without the need for writing own algorithm
References
Thank you!