Joint Modelling of Event Counts and Survival Times: Example Using Data from the MESS Trial

J. K. Rogers
J. L. Hutton  K. Hemming

Young Researchers Meeting
13th January 2009
Outline

Background and MESS
  Epilepsy
  MESS

Survival Analysis
  Survival Data
  Censoring

Time to First Tonic-Clonic Seizure
  Assessing the Risk of Subsequent Tonic-Clonic Seizures

Exploratory Analysis
  Summary Statistics
  Modelling

Joint Model
  Joint Modelling of Event Counts and Survival Times
  Extension to Joint Modelling of Event Counts and Multiple Survival Times

Research Plan
  Cure-Rate Model
  Modification of Parameters
  Further Plans
<table>
<thead>
<tr>
<th>Outline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Background and MESS</strong></td>
</tr>
<tr>
<td>Epilepsy</td>
</tr>
<tr>
<td>MESS</td>
</tr>
<tr>
<td><strong>Survival Analysis</strong></td>
</tr>
<tr>
<td>Survival Data</td>
</tr>
<tr>
<td>Censoring</td>
</tr>
<tr>
<td><strong>Time to First Tonic-Clonic Seizure</strong></td>
</tr>
<tr>
<td>Assessing the Risk of Subsequent Tonic-Clonic Seizures</td>
</tr>
<tr>
<td><strong>Exploratory Analysis</strong></td>
</tr>
<tr>
<td>Summary Statistics</td>
</tr>
<tr>
<td>Modelling</td>
</tr>
<tr>
<td><strong>Joint Model</strong></td>
</tr>
<tr>
<td>Joint Modelling of Event Counts and Survival Times</td>
</tr>
<tr>
<td>Extension to Joint Modelling of Event Counts and Multiple Survival Times</td>
</tr>
<tr>
<td><strong>Research Plan</strong></td>
</tr>
<tr>
<td>Cure-Rate Model</td>
</tr>
<tr>
<td>Modification of Parameters</td>
</tr>
<tr>
<td>Further Plans</td>
</tr>
</tbody>
</table>
Epilepsy

- Epilepsy is defined as the occurrence of recurrent and unprovoked seizures.
- ILAE classification scheme divides seizures into partial, generalised or unclassified seizures.

Partial seizures - involve only part of the brain; simple or complex. Include motor, sensory, occipital, frontal lobe and temporal lobe seizures. Partial epilepsy can sometimes occur with secondary generalisation

Generalised seizures - involve all of the brain and are categorised as tonic-clonic (grand mal), absence (petit mal), myoclonic and atonic
Early Epilepsy and Single Seizures

- On average 50% of people do not experience a recurrence following a first seizure
- Around 20 – 30% of people will never achieve long-term remission
- Risk of future seizures increases with the number of previous seizures
- Antiepileptic drugs come with unpleasant side effects
Aim of Trial

MRC Multicentre Trial for Early Epilepsy and Single Seizures

- When should treatment with antiepileptic drugs commence
Aim of Trial

MRC Multicentre Trial for Early Epilepsy and Single Seizures

- When should treatment with antiepileptic drugs commence
- Comparison of policies: immediate versus deferred treatment in those patients where uncertainty about starting treatment remained
- Interest lay in effects on short-term remission and long-term prognosis

[Marson et al., 2005]
Outline

Background and MESS
  Epilepsy
  MESS

Survival Analysis
  Survival Data
  Censoring

Time to First Tonic-Clonic Seizure
  Assessing the Risk of Subsequent Tonic-Clonic Seizures

Exploratory Analysis
  Summary Statistics
  Modelling

Joint Model
  Joint Modelling of Event Counts and Survival Times
  Extension to Joint Modelling of Event Counts and Multiple Survival Times

Research Plan
  Cure-Rate Model
  Modification of Parameters
  Further Plans
Modelling Survival Data

[Collett, 2003]

- Analysis of data which is in the form of times from a well defined start point, up to a particular event of interest.
- Time to 1st, 2nd, 5th seizure, 1st tonic-clonic, 1 and 2 yr remission
Modelling Survival Data

[Collett, 2003]

- Analysis of data which is in the form of times from a well defined start point, up to a particular event of interest.
- Time to 1st, 2nd, 5th seizure, 1st tonic-clonic, 1 and 2 yr remission
- Two functions are of central interest:
  - Survivor function - $S(t) = \mathbb{P}(T \geq t)$
  - Hazard function - $h(t) = \lim_{\delta t \to 0} \left\{ \frac{\mathbb{P}(t \leq T \leq t + \delta t | T \geq t)}{\delta t} \right\}$
Censoring

- Individual’s actual survival time cannot be measured
- We have instead some measurable censored time associated with them
- Three types of censoring:
  1. right censoring, occurring when the censored survival time is less than the actual, unknown survival time
  2. left censoring, occurring when the censored survival time is greater than the actual, unknown survival time, and
  3. interval censoring, which is evident if the actual survival time is only known up to some interval.
Outline

Background and MESS
  Epilepsy
  MESS

Survival Analysis
  Survival Data
  Censoring

**Time to First Tonic-Clonic Seizure**
  Assessing the Risk of Subsequent Tonic-Clonic Seizures

Exploratory Analysis
  Summary Statistics
  Modelling

Joint Model
  Joint Modelling of Event Counts and Survival Times
  Extension to Joint Modelling of Event Counts and Multiple Survival Times

Research Plan
  Cure-Rate Model
  Modification of Parameters
  Further Plans
Time to First Tonic-Clonic Seizure

Time to first tonic–clonic seizure

- Tonic–Clonic only
- 2 Degree Tonic–Clonic
- Tonic–Clonic with other Gen
- Minor

Time in years
Survival
0.0 0.2 0.4 0.6 0.8 1.0
0 2 4 6 8

Tonic–Clonic only
2 Degree Tonic–Clonic
Tonic–Clonic with other Gen
Minor
Time to First Tonic-Clonic Seizure

Time to first tonic–clonic seizure

Survival

Time in years

Tonic–Clonic

Minor
Time to First Tonic-Clonic Seizure

Survival Analysis

- Time to First Tonic-Clonic Seizure
- Exploratory Analysis
- Joint Model
- Research Plan

Time to first tonic–clonic seizure

Survival

Immediate
Deferred

Tonic–Clonic
Minor

Time in years

0 2 4 6 8

0.0 0.2 0.4 0.6 0.8 1.0
Outline

Background and MESS
  Epilepsy
  MESS
Survival Analysis
  Survival Data
  Censoring
Time to First Tonic-Clonic Seizure
  Assessing the Risk of Subsequent Tonic-Clonic Seizures
Exploratory Analysis
  Summary Statistics
  Modelling
Joint Model
  Joint Modelling of Event Counts and Survival Times
  Extension to Joint Modelling of Event Counts and Multiple Survival Times
Research Plan
  Cure-Rate Model
  Modification of Parameters
  Further Plans
# Seizure Type Pre-Randomisation

<table>
<thead>
<tr>
<th>Seizure Type</th>
<th>Immediate</th>
<th>Deferred</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple Partial</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>Complex Partial</td>
<td>36</td>
<td>32</td>
</tr>
<tr>
<td>Partial with 2° Tonic-Clonic</td>
<td>239</td>
<td>215</td>
</tr>
<tr>
<td>Myoclonic</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Absence</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Tonic-Clonic</td>
<td>375</td>
<td>406</td>
</tr>
<tr>
<td>Other</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>Combination of Generalised</td>
<td>21</td>
<td>19</td>
</tr>
</tbody>
</table>
Number of Seizures Pre-Randomisation

- **All Seizures**
- **Partial Seizures**
- **Partial Seizures with Secondary Tonic–Clonic**
- **Tonic–Clonic Seizures**
- **Other Seizures**
- **Combination of Generalised Seizures**
Period of Time From First Seizure to Randomisation

![Bar chart showing the frequency of periods from first seizure to randomisation across different time intervals.](chart.png)
Pre-Randomisation Seizure Rates
Kaplan-Meier Curves by Treatment Type

— immediate, - - - deferred
Kaplan-Meier Curves by Seizure Type

- Time to first seizure
- Time to second seizure
- Time to fifth seizure

Survival curves for different seizure types:
- All Seizures
- Partial
- Partial with 2 Degree T–C
- Tonic–Clonic
- Other
- Combination
Accelerated Failure Time Assumption

\[ z = 0 \text{ - allocated to immediate treatment group} \]
\[ z = 1 \text{ - allocated to deferred treatment group} \]

Survivor function at \( z = 0 \) is \( S_0(t) \)

Under AFT assumption survivor function at \( z = 1 \), \( S_1(t) \) is given by

\[ S_1(t) = S_0(\omega t) \quad (1) \]
Testing the AFT Assumption

We define $t_0^{(a)}$, $t_1^{(a)}$, for $0 < a < 1$, by:

$$a = S_0(t_0^{(a)}) \quad t_0^{(a)} = S_0^{-1}(a)$$

$$a = S_1(t_1^{(a)}) \quad t_1^{(a)} = S_1^{-1}(a)$$

Then by (1) $t_1^{(a)} = t_0^{(a)}/\omega$, that is the resulting Q-Q plots should produce a straight line that passes through the origin.
Q-Q Plots Justifying the AFT Distributions

Distributions that are AFT include the Weibull, Exponential, Log-logistic, Lognormal, Gamma, and Inverse Gaussian distributions.
Log-Logistic Distribution

Survivor function given by:

\[ S(t) = \frac{1}{1 + \left(\frac{t}{b}\right)^a} \]

Consider the following transformation:

\[ \ln \left\{ \frac{S(t)}{1 - S(t)} \right\} = -a \ln(t) + a \ln(b) \]

which is linear in \( \ln(t) \)

Hence a plot of \( \ln(t) \) against \( \ln \left\{ \frac{S(t)}{1 - S(t)} \right\} \) should give a straight line with gradient \(-a\), passing through \( a \ln(b) \)
Justifying the Log-Logistic Distribution

Time to first seizure

\[ \ln(t) \]

\[ \ln\left( \frac{S(t)}{1-S(t)} \right) \]

Time to second seizure

\[ \ln(t) \]

\[ \ln\left( \frac{S(t)}{1-S(t)} \right) \]

Time to fifth seizure

\[ \ln(t) \]

\[ \ln\left( \frac{S(t)}{1-S(t)} \right) \]
## Outline

**Background and MESS**
- Epilepsy
- MESS

**Survival Analysis**
- Survival Data
- Censoring

**Time to First Tonic-Clonic Seizure**
- Assessing the Risk of Subsequent Tonic-Clonic Seizures

**Exploratory Analysis**
- Summary Statistics
- Modelling

**Joint Model**
- Joint Modelling of Event Counts and Survival Times
- Extension to Joint Modelling of Event Counts and Multiple Survival Times

**Research Plan**
- Cure-Rate Model
- Modification of Parameters
- Further Plans
Response Variables

For individual $i$ we take:

1. $X_i$ - a pre-randomisation seizure count over a period $u_i$
2. $Y_i$ - time to first seizure post-randomisation

[Cowling et al., 2006]
Joint Model

\[ X_i \mid \nu_i \sim \text{Poisson}(\lambda_i u_i \nu_i) \]
\[ Y_i \mid \nu_i \sim \text{Exponential}(\lambda_i \psi_i \nu_i) \]
\[ \nu_i \sim \text{Gamma}(\alpha, \alpha) \]

\[ \lambda_i = \exp(\beta_1' z_{1i}) \]
\[ \psi_i = \exp(\beta_2' z_{2i}) \]

Unconditional distribution of \( Y_i \) turns out to be Lomax distribution with scale \( \alpha/\lambda_i \psi_i \) and shape \( \alpha \)
Joint Model Continued

Fig. 1. Graphical model of the underlying event process

\( \beta_1 \) \( Z_{1i} \) \( \alpha \) \( \beta_2 \) \( Z_{2i} \)

\( \lambda_i \) \( v_i \) \( \psi_i \)

\( u_i \) \( X_i \) \( Y_i \delta_i \)

\( i = 1, ..., n \)
Log-Logistic and Lomax distributions

- Log-logistic

\[ F_Y(y_i) = 1 - \{1 + (y_i/b)^a\}^{-1} \]

- Lomax

\[ F_Y(y_i) = 1 - \{1 + (y_i/b)^a\}^{-a} \]

Log-logistic distribution with shape parameter set to 1 is equivalent to a Lomax distribution with its shape parameter also equal to 1.
Joint Model Continued

Table 1. Maximum likelihood parameter estimates for Poisson and negative binomial GLMs and full joint models†

<table>
<thead>
<tr>
<th>Term</th>
<th>Regression coefficient</th>
<th>Estimates (standard errors) for the following models:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Poisson GLM</td>
<td>Negative binomial GLM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \lambda_i )</td>
<td>( \alpha )</td>
<td>( \infty )</td>
<td>1.221 (0.055)</td>
</tr>
<tr>
<td>( \beta_0 )</td>
<td></td>
<td>-3.093 (0.033)</td>
<td>-3.059 (0.092)</td>
</tr>
<tr>
<td>( \beta_{\text{type}} )</td>
<td></td>
<td>0.541 (0.013)</td>
<td>0.557 (0.037)</td>
</tr>
<tr>
<td>( \beta_{\text{age}} )</td>
<td></td>
<td>0.035 (0.009)</td>
<td>0.025 (0.022)</td>
</tr>
<tr>
<td>( \psi_i )</td>
<td>( \beta_0 )</td>
<td></td>
<td>-2.492 (0.042)</td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{trt}} )</td>
<td></td>
<td>0.050 (0.041)</td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{type}} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{trt} \times \text{type}} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{age}} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \beta_{\text{trt} \times \text{age}} )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( -\log\text{-likelihood} )</td>
<td></td>
<td>7489 (1137)</td>
<td>3311 (1136)</td>
</tr>
</tbody>
</table>

†The term \( \lambda_i \) contains parameters from equation (3) corresponding to the effect of covariates on the underlying event rate. The term \( \psi_i \) contains parameters from equation (4) corresponding to the effect of covariates in modifying the post-randomization event rate. Type is –1 or 1 for generalized or partial onset epilepsy respectively; age is the original age minus 30 years, in decades; trt is –1 or 1 for CBZ or VPS respectively.
Joint Model Continued

Table 2. Estimates for typical survival models fitted to the times to first seizure†

<table>
<thead>
<tr>
<th>Regression coefficient</th>
<th>Estimates (standard errors) for the following models:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Exponential</td>
<td>Weibull</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>$-6.984 (0.074)$</td>
<td>$-3.345 (0.151)$</td>
</tr>
<tr>
<td>$\beta_{\log\text{(count)}}$</td>
<td>$0.406 (0.022)$</td>
<td>$0.302 (0.036)$</td>
</tr>
<tr>
<td>$\beta_{\text{type}}$</td>
<td>$0.184 (0.032)$</td>
<td>$0.138 (0.047)$</td>
</tr>
<tr>
<td>$\beta_{\text{age}}$</td>
<td>$-0.144 (0.018)$</td>
<td>$-0.098 (0.027)$</td>
</tr>
<tr>
<td>$\beta_{\text{tt}}$</td>
<td>$-0.004 (0.026)$</td>
<td>$0.008 (0.038)$</td>
</tr>
<tr>
<td>$\beta_{\text{tt}\times\text{type}}$</td>
<td>$0.182 (0.026)$</td>
<td>$0.115 (0.038)$</td>
</tr>
<tr>
<td>Scale</td>
<td>$1(0)$</td>
<td>$0.482 (0.014)$</td>
</tr>
<tr>
<td>$-\text{Log-likelihood}$ (degrees of freedom)</td>
<td>$5736 (1134)$</td>
<td>$5269 (1133)$</td>
</tr>
</tbody>
</table>

†Type is $-1$ or $1$ for generalized or partial onset epilepsy respectively; age is the original age minus 30 years, in decades; tt is $-1$ or $1$ for CBZ or VPS respectively.
Times to First and Second Seizures Only

\[ X_i \mid \nu_i \sim \text{Poisson}(\lambda_i u_i \nu_i) \]
\[ Y_{ji} \mid \nu_i \sim \text{Exponential}(\lambda_i \psi_i \nu_i), \quad j = 1, 2, \]

\[
\lambda_i = \exp(\beta'_1 z_{1i}) \\
\psi_i = \exp(\beta'_2 z_{2i})
\]

The unconditional joint distribution of the \( Y_{ji}, j = 1, 2 \) is the bivariate Lomax distribution with scale \( \alpha/\lambda_i \psi_i \) and shape \( \alpha \)
Bivariate Lomax Distribution

The unconditional joint survivor function of the \( Y_{ji}, j = 1, 2 \) is

\[
S_{Y_1,Y_2}(y_{1i}, y_{2i}; \lambda_i, \psi_i, \alpha) = \left\{ 1 + \frac{\lambda_i \psi_i (y_{1i} + y_{2i})}{\alpha} \right\}^\alpha
\]

which has expectation and variance

\[
\mathbb{E}(Y_j) = \frac{\alpha}{\lambda_i \psi_i (\alpha - 1)} \quad \text{Var}(Y_j) = \frac{\alpha^3}{(\lambda_i \psi_i)^2 (\alpha - 1)^2 (\alpha - 2)}
\]

The multivariate Lomax distribution has a decreasing joint hazard function, which is particularly useful for epilepsy data.
Three different ways censoring can arise:

(i) $Y_{1i}$ and $Y_{2i}$ both observed,

(ii) $Y_{1i}$ is observed, but $Y_{2i}$ is censored, and

(iii) $Y_{1i}$ is censored, so we have no information about $Y_{2i}$.

Straightforward to maximise log-likelihood, allowing inference on the parameters $\alpha$, $\beta_1$ and $\beta_2$, using a numerical method such as Newton-Raphson.
Outline

Background and MESS
  Epilepsy
  MESS

Survival Analysis
  Survival Data
  Censoring

Time to First Tonic-Clonic Seizure
  Assessing the Risk of Subsequent Tonic-Clonic Seizures

Exploratory Analysis
  Summary Statistics
  Modelling

Joint Model
  Joint Modelling of Event Counts and Survival Times
  Extension to Joint Modelling of Event Counts and Multiple Survival Times

Research Plan
  Cure-Rate Model
  Modification of Parameters
  Further Plans
Allowing for ‘Immunes’

- ‘Proper’ survival distribution has $S(t) \in [0, 1]$.
- An improper survival distribution allows some of the ‘true’ lifetimes to be infinite, representing the immunes.
- Let $p$ be the susceptible proportion, then

\[
p = F(\infty) = \lim_{t \to \infty} F(t), \quad 0 < p \leq 1
\]

- Assume that $F_0(t)$ is the distribution function for the susceptible individuals, then

\[
F(t) = pF_0(t)
\]

[Maller, 1996]
Allowing Treatment Parameter to Change Through Time

Unrealistic to assume $\psi_i$ constant through time

- Modify treatment parameter
  - $\psi_i = (s\gamma + 1) \exp(\beta_2'z_{2i})$
  - $\psi_i = \exp(\beta_2'z_{2i} + s\gamma)$
  - $s$: previous number of seizures

- Also consider $t$: time since last seizure

- May want to include binary covariate 1/0 indicating whether an individual is currently taking AEDs or not respectively
Long-Term Remission

- Model remission as seizure rates
  One-year remission $\equiv$ Seizure rate less than 1 in 365 days
- Investigate dropout
Randomisation Issues

- Two randomisation forms used during the trial

- Second randomisation strategy allows comparisons between specific drugs
Randomisation Issues

- Two randomisation forms used during the trial
  1. Randomisation $\rightarrow$ Drug (614 participants)

- Second randomisation strategy allows comparisons between specific drugs
Randomisation Issues

- Two randomisation forms used during the trial
  1. Randomisation → Drug (614 participants)
  2. Drug → Randomisation (811 participants)
- Second randomisation strategy allows comparisons between specific drugs
Handling Randomisation Issues

• Antiepileptic drug strongly dependent on a number of baseline covariates, such as:
  • age
  • type of epilepsy
  • nature of the seizures

• Regress missing items on those influential baseline covariates we have observed

• Multiple imputation
For Further Reading I

A. Marson, A. Jacoby, A. Johnson, L. Kim, C. Gamble, D. Chadwick, on behalf of the Medical Research Council MESS Study Group.
Immediate versus deferred antiepileptic drug treatment for early epilepsy and single seizures: a randomised controlled trial.

Joint modelling of event counts and survival times.
For Further Reading II

D. Collett.  
*Modelling Survival Data in Medical Research, 2nd Edition.*  

*Survival Analysis with Long Term Survivors.*  
John Wiley and Sons, 1996.