Joint modelling of pre-randomisation event counts and multiple post-randomisation survival times with cure rates: application to data for early epilepsy and single seizures

J.K. Rogers[a] and J.L. Hutton[b]

[a] *Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK*;

[b] *Department of Statistics, University of Warwick, Coventry, UK*

This is an Accepted Manuscript of an article published by Taylor and Francis in Journal of Applied Statistics, available online: http://wwww.tandfonline.com/10.1080/02664763.2012.748720.

Abstract

We consider the analysis of recurrent event data that examines the differences between two treatments. The outcomes that are considered in the analyses are the pre-randomisation event count and post-randomisation times to first and second event with associated cure fractions.

We develop methods that allow pre-randomisation counts and two post-randomisation survival times to be jointly modelled under a Poisson process framework, assuming that outcomes are predicted by (unobserved) event rates. We apply these methods to data that examines the difference between immediate and deferred treatment policies in patients presenting with single seizures or early epilepsy.

We find evidence to suggest that post-randomisation seizure rates change at randomisation and following a first seizure post-randomisation. We also find that there are cure rates associated with the post-randomisation times to first and second seizure. The increase in power over standard survival techniques, offered by the joint models we propose, resulted in more precise estimates of the treatment effect and the ability to detect interactions with covariate effects.

Epilepsy; Poisson process; Proportion cured; Recurrent events; Survival analysis

1 Introduction

Our work is motivated by the analysis of recurrent event data. Examples of recurrent event data include asthma attacks [10], gout flares [1] and epileptic seizures, which we consider here. The analysis

of recurrent event data usually considers either the event counts or the gap times between successive events. In studies of epileptic seizures, there is frequently information about individuals' seizure patterns over a period of time prior to randomisation that is not fully utilised in analysis. Epilepsy is characterised by multiple seizures, not a single, isolated event, yet in many treatment studies it is only time to first event that is analysed. We develop methodology that allows the pre-randomisation seizure counts and multiple post-randomisation survival times to be jointly modelled, assuming that outcomes are predicted by (unobserved) seizure rates, and consider the possibility of there being cure rates present post-randomisation. We suppose that each patient has an underlying constant seizure rate, which depends on baseline attributes, and that post-randomisation seizure rates are modified relative to associated baseline seizure rates and treatment policy. A greater reduction results in a longer time to seizure post-randomisation, indicating a better therapy.

A model typically used for the analysis of count data is the Poisson Generalised Linear Model [21]. When presented with count data that are overdispersed, random effect mixture distributions are often used. The most convenient choice of random effects distribution is the Gamma, which consequently yields the Negative Binomial distribution [12]. Alternatives to the Gamma distribution as the mixing distribution are considered by Hougaard *et al.* [13], with special consideration given to the analysis of the frequency of epileptic seizures, but the inclusion of covariates is not considered.

Cook and Lawless [5] discuss the use of models that are appropriate in the specification and testing of treatment effects on recurrent events. They suggest that methods based on rates and mean functions offer the most straightforward specification of treatment effects for recurrent events, but recognise that there may be situations where the equivalent analyses based on gap times are more natural. For example, if datasets exhibit cure rates, models that focus on rates and mean functions would assess post-randomisation event rates that are zero, but to consider this, focus would need to be on a suitable length of observation time, or gap time, with no events recorded.

An additional consideration in the analysis of recurrent event data is the use of baseline count data. In this scenario, Cook and Lawless [5] consider the use of mixed Poisson processes for the analysis of recurrent event data that incorporates a period of observation in which subjects are monitored prior to randomisation to treatment. Interest in this instance would be focussed on the change in event rate pre and post-randomisation.

Diggle *et al.* [9] present statistical methodology for the joint modelling of repeated measurements and a single survival time with common random effects (note that in survival analysis, random effects

are often termed frailties). In this scenario, the repeated measurements may be of primary interest and the survival time is a secondary outcome, possibly a potentially informative dropout time, or the survival time may be of primary interest and the repeated measurements represent a time-varying explanatory variable. We propose methodology that differs from this as it considers the joint modelling of a pre-randomisation event count and multiple post-randomisation survival times. In this scenario, most standard survival analysis would treat the pre-randomisation event count information as a covariate [27], possibly a covariate with measurement error. As an alternative, Cook and Wei [6] consider a semi-parametric approach for data that are a combination of event counts and survival times. Cowling *et al.* [8] considers a fully parametric technique that jointly models an individual's pre-randomisation event count, and a single post-randomisation failure time under a Poisson process framework. There has been no discussion, however, of the analysis of recurrent event data which is recorded as a pre-randomisation baseline count and multiple post-randomisation gap times with cure rates.

We first introduce the dataset that has motivated our methodology, with some exploratory analysis. We then present the joint model that we have developed and discuss the results obtained following its application to the dataset.

2 The MRC Multicentre Trial for Early Epilepsy and Single Seizures (MESS)

The question of whether to start patients on a course of anticonvulsants after a single epileptic seizure remains an area of uncertainty. Several studies have shown that intervention after a single seizure reduces the risk of short-term recurrence, but does not affect the long-term remission rates in individuals with single or infrequent seizures [20, 3]. Warrell *et al.* [29] state that seizure recurrence after a single untreated seizure is around 80%, Berg and Shinnar [2] put seizure recurrence at 50%. It is also thought that the risk of future seizures increases with the number of previous seizures, with around 30% of epilepsy sufferers never achieving long-term remission [4]. Antiepileptic drugs (AEDs) are often accompanied with adverse effects, which include weight loss or weight gain, altered mood, drowsiness, hair loss, polycystic ovarian disease, visual field defects and teratogenicity. For most epilepsy sufferers, the benefits of AEDs will far outweigh the associated risks. For those individuals, however, who have had only a single seizure, or have infrequent and mild epileptic seizures, the question of whether to

withhold treatment until absolutely necessary becomes clinically important.

The MRC Multicentre Trial for Early Epilepsy and Single Seizures (MESS) [20] compared two treatment schemes: immediate versus delayed treatment in those patients presenting with only a single, or few epileptic seizures. MESS was a multicentre, unmasked randomised controlled trial that randomised 1443 individuals in the early stages of epilepsy to immediate or deferred treatment. Interest lay in the effects of these treatment policies on both short-term recurrence, and long-term prognosis. Those allocated to the deferred treatment group had treatment withheld until both clinician and patient agreed that treatment was necessary. MESS was a pragmatic trial, meaning that all subsequent choices of antiepileptic drug, dose and duration were in line with the clinicians' usual practice.

Baseline covariates collected at randomisation include age, sex, and information on patients' pre-randomisation seizure history, such as seizure type and seizure frequency. The pre-randomisation seizure types are categorised as follows: 'Tonic-Clonic' contains those presenting with tonic-clonic seizures only, '2° Tonic-Clonic' are those individuals presenting with partial seizures accompanied by secondary tonic-clonic seizures and the 'Generalised' group contains those presenting with any types of generalised seizures (this group could include those having a combination of tonic-clonic and other generalised seizures). The 'Partial' group contains those individuals who presented with partial seizures only pre-randomisation (either simple or complex), whilst 'Other' contains those presenting with seizures that do not fit into any of the above categories. An electroencephalogram (EEG) was also requested for each individual. EEG outcome is simply defined here as being abnormal or normal.

Detailed methods and primary analyses can be found in Marson *et al.* [20] and Kim *et al.* [16]. Previous analyses of the MESS data have adopted the Cox Proportional-Hazards model, despite the data violating the proportional-hazards assumption [24]; exploratory analysis has suggested that the data should be modelled by an accelerated failure time distribution. Statistical analyses are by intention to treat; interest lay in the treatment policy to which an individual was assigned rather than whether an individual was on treatment at the time of future seizures.

Typically, the questions asked by individuals presenting with a single seizure are: 'Will I have another seizure, and if so, when?' or 'If I have a second seizure am I likely to have more, and with increased frequency?'. When answering questions of this type it is more intuitive and clinically relevant to analyse recurrent event data in terms of the survival times. Furthermore, time to first seizure is an internationally agreed outcome in epilepsy trials [15], so analysis of gap times jointly with

pre-randomisation event counts is natural. We consider models for the effect of treatment policies on times to first and second seizures, and exploit information on pre-randomisation seizure history, including seizure type, to enhance conclusions.

3 Exploratory analysis

The small number of patients with generalised or other seizures were excluded, as were 23 individuals with missing data, resulting in a final sample size of 1334.

The pre-randomisation period at risk (the time from first seizure to randomisation) is essential in the estimation of pre-randomisation seizure rates. Of the 1334 individuals analysed, 782 (59%) presented a single seizure pre-randomisation. The times from these single seizures to randomisation ranged from the same day to 464 days, with the median number of days being 26. These times may be inaccurately small, as the time at risk but seizure free before the first seizure, is not known. This results in high, but imprecise estimates of their associated underlying seizure rates and an ensuing overestimation of the seizure rate reductions. Following discussions with clinicians, we subsequently made adjustments to the the number of days at risk before randomisation, denoted u_i: if the time from first seizure to randomisation was greater than 182 days, u_i was this time, otherwise $u_i = 182$. As a sensitivity analysis the data were re-analysed with minimum times of 91 and 365 days. The resulting parameter estimates can be found in Rogers *et al.* [25]. The magnitudes of differences observed in seizure rates between the groups were maintained through each adjustment. The log-likelihoods associated with each model would suggest that having a minimum pre-randomisation period of 365 days is optimal, but the likelihood function is very flat and the relative differences between groups were maintained, hence the clinicians' suggestion of a minimum pre-randomisation period of 182 days was used for all future analyses of the MESS data. The subsequent absolute estimates of the expected number of days until the next seizure post-randomisation were not sensitive to the choice of minimum value for u_i.

Allocation to treatment policy was almost equal: 663 patients were allocated to immediate treatment and 671 to deferred treatment. Additionally, 641 (48%) of those analysed experienced at least one seizure following randomisation (290 immediate, 351 deferred), with a subsequent 441 (69%) of these experiencing a second (203 immediate, 238 deferred).

There were essentially two possible time scales of interest: the total time, measured from randomi-

sation, to the occurrence of the first and second post-randomisation seizures, or the gap times between post-randomisation seizures. Our analysis of the MESS data focusses on the analysis of the times from randomisation to first seizure, and the times from first to second seizure, with the overall follow-up time subject to right censoring. In this censoring scenario, the times from first to second seizure may be subject to dependent censoring, as the duration of the time to first seizure may have an effect on the potential censoring value of the second duration [17]. A long time to first seizure post-randomisation implies a short observation period for the time from first to second seizure post-randomisation, and vice versa.

The severity of the dependent censoring in our data was assessed by comparing non-parametric estimates of the survivor function of the times from first to second seizure, conditional on the times to first seizure [28], with Kaplan-Meier estimates of the marginal survivor function for the times from first to second seizure. In general, Kaplan-Meier estimates were lower than the conditional estimates of the survivor function [24]. The differences are due to longer times to first seizure resulting in shorter subsequent observation periods for second seizure, and higher rates of censoring. One method for handling dependent censoring, which we use here, is to fit models with individual-specific independent and identically distributed random effects to induce associations among gap times [5]. Such models assume that given the random effect, the gap times for an individual are independent. Alternative approaches include the specification of a multivariate model for a specified set of gap times or conditional models [5].

The Kaplan-Meier estimates (Figure 1) show that treatment policy appears to be influential in determining an individual's time to first seizure post-randomisation, but not their time from first to second seizure. A plausible explanation for this is that those individuals randomised to deferred treatment, who subsequently experience a seizure post-randomisation, would most likely start treatment with AEDs, bringing them in line with those allocated to immediate treatment. Also note that the times from first to second seizure are typically shorter than times to first seizure.

Figure 2 shows the empirical cumulative distribution function for time to first seizure as a proportion of the total time to second seizure, for only those individuals presenting with at least two seizures post-randomisation, as dependent censoring is present. The median is 0.623, confirming that those experiencing at least two seizures post-randomisation typically have shorter gap times between first and second seizures than randomisation and first seizure. This suggests clustering of seizures. Those experiencing partial seizures pre-randomisation typically have a shorter time to first seizure

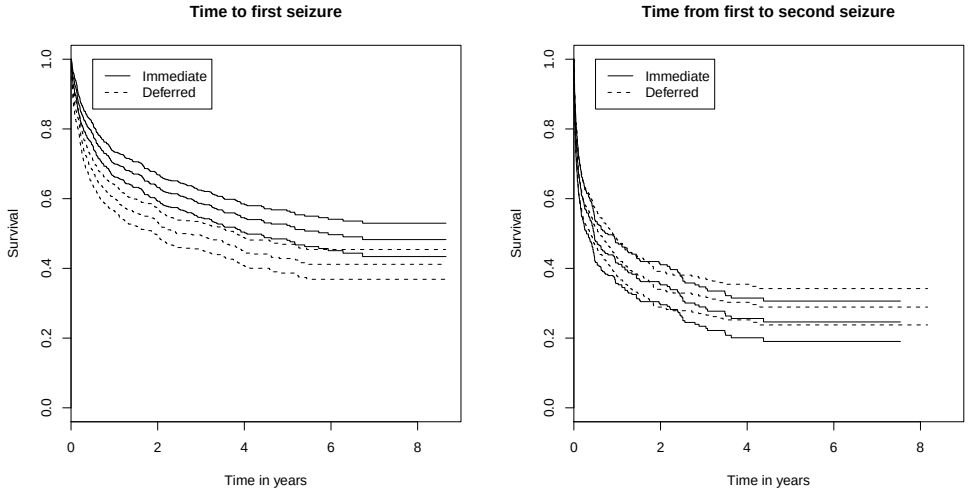

Figure 1: Kaplan-Meier curves for times to first and from first to second seizure (with 95% CI), stratified by treatment policy.

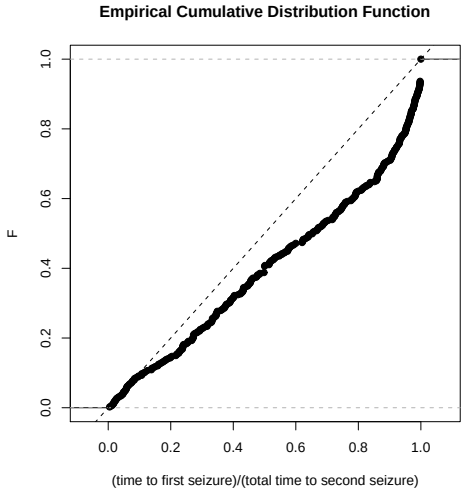

Figure 2: Empirical cumulative distribution function for (time to first seizure)/(total time to second seizure).

post-randomisation than the tonic-clonic seizure groups, but for times from first to second seizure the difference between the Kaplan-Meier curves is less pronounced (Figure 3).

4 Joint modelling of event counts and survival times

We describe how pre-randomisation event counts and post-randomisation times to first seizure and from first to second seizure can be jointly modelled under a censored Poisson process with cure rates framework.

Recall that seizure recurrence after a single, untreated seizure is around $50\% - 80\%$ and over half of the trial participants presented only a single seizure pre-randomisation. It is therefore reasonable

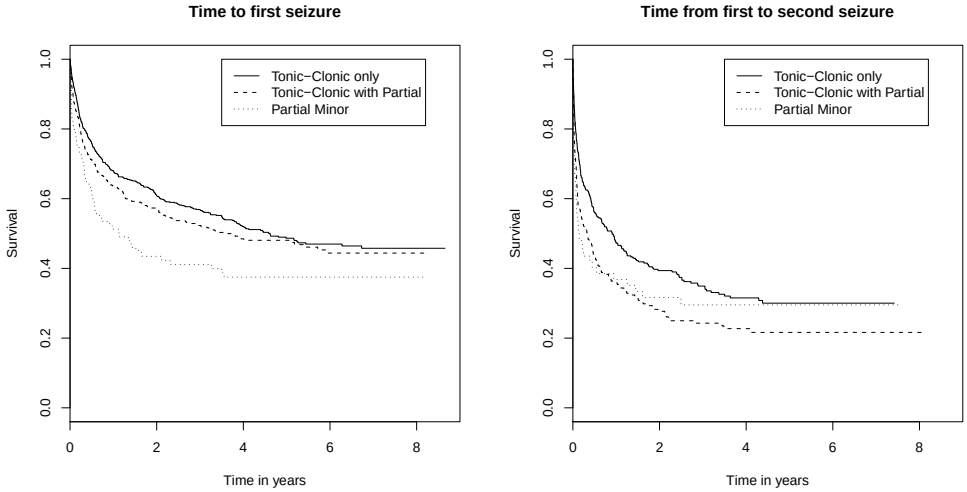

Figure 3: Kaplan-Meier curves for times to first and from first to second seizure, stratified by seizure type pre-randomisation.

to suspect that a substantial proportion of patients would never have a seizure post-randomisation, regardless of the length of time for which they were followed. We consider these individuals as cured, or immune to seizure recurrence. If survival data has a proportion that are immune to the event of interest, a model that ignores this could result in underestimates of the post-randomisation seizure rates. Kaplan-Meier curves for times to first and second seizure both level off well above zero, with narrow confidence intervals despite substantial censoring, suggesting that there may be an immune component for both times to first seizure and from first to second seizure post-randomisation.

A survival distribution that is improper allows, formally, infinite survival times. Cure rate models allow the quantity $p = F(\infty) = \lim_{t\to\infty} F(t)$ (where $F(t)$ is the cumulative distribution function of the survival times) to be strictly less than 1, corresponding to the presence of immunes in the population [19]. Furthermore, for all $t \geq 0$, the cumulative distribution function of the survival times is $F(t) = pG(t)$, where $G(t)$ is the cumulative distribution function of the survival times for those that are susceptible to the event of interest. To ensure that p remains within the interval $[0, 1]$ a logistic reparameterisation is often considered.

We assume that individual i, $i = 1 \ldots n$, experiences events according to a Poisson process with rate $\lambda_i \nu_i$. The parameter λ_i is a function of the baseline covariates with additional heterogeneity in the population being modelled through the frailty term ν_i, assumed to follow a Gamma distribution with expectation 1 and variance $1/\alpha$. Smaller values of α are indicative of higher levels of heterogeneity. Consequently, the pre-randomisation event count, X_i, over a period u_i, for individual i, follows a Poisson distribution with mean $\lambda_i u_i \nu_i$ and the gap times are Exponential with the same rate. After

randomisation, treatment policy will modify the event rate. We assume that the treatment acts multiplicatively on the rate, so the post-randomisation event rate for individual i will be $\lambda_i\psi_i\nu_i$, where ψ_i models the treatment effects. Let T_{1i} and T_{2i} be the times from randomisation to first and second seizures respectively, for individual i. We set $Y_{1i} = T_{1i}$ and $Y_{2i} = T_{2i} - T_{1i}$, so that Y_{1i} is the time to first seizure and Y_{2i} is the time from first seizure to the second. Both Y_{1i} and Y_{2i} for the susceptible proportion are Exponentially distributed with rate $\lambda_i\psi_i\nu_i$ and the two survival times are independent given ν_i. In summary, the joint model is specified by the following equations:

$$\begin{aligned} f_{X|\nu}(x_i \mid \nu_i; \lambda_i, u_i) &= \frac{(\lambda_i u_i \nu_i)^{x_i} \exp(-\lambda_i u_i \nu_i)}{x_i!}, \\ f_{Y_1,Y_2|\nu}(y_{1i}, y_{2i} \mid \nu_i; \lambda_i, \psi_i, p_{1i}, p_{2i}) &= p_{1i}p_{2i}(\lambda_i\psi_i\nu_i)^2 \exp(-\lambda_i\psi_i\nu_i(y_{1i} + y_{2i})), \\ f_\nu(\nu_i; \alpha) &= \frac{\alpha^\alpha \nu_i^{\alpha-1} \exp(-\alpha\nu_i)}{\Gamma(\alpha)}, \end{aligned}$$

where $\lambda_i = \exp(\boldsymbol{\beta}_1'\boldsymbol{z}_{1i})$, $\psi_i = \exp(\boldsymbol{\beta}_2'\boldsymbol{z}_{2i})$,

$$\begin{aligned} p_{1i} &= \frac{\exp(\boldsymbol{\kappa}_1'\boldsymbol{w}_{1i})}{1 + \exp(\boldsymbol{\kappa}_1'\boldsymbol{w}_{1i})}, \\ p_{2i} &= \frac{\exp(\boldsymbol{\kappa}_2'\boldsymbol{w}_{2i})}{1 + \exp(\boldsymbol{\kappa}_2'\boldsymbol{w}_{2i})}. \end{aligned}$$

and $\boldsymbol{z}_{1i}$, $\boldsymbol{z}_{2i}$, $\boldsymbol{w}_{1i}$, $\boldsymbol{w}_{2i}$ are vectors of covariates, not necessarily distinct.

The unconditional density of X_i, $f_X(x_i; \lambda_i, u_i, \alpha)$, is the Negative Binomial. The unconditional joint distribution of the Y_{ji}, $j = 1, 2$, for the susceptible proportion, is the bivariate Lomax distribution [22], with each of the Y_{ji} having univariate Lomax marginal distributions with shape and scale parameters α and $\alpha/\lambda_i\psi_i$ respectively (see Appendix A). This model formulation gives an accelerated failure time model with Lomax baseline distribution.

When formulating the likelihood, the different ways that censoring can occur need to be considered. There are four different ways censoring can arise in this setting, namely: (i) Y_{1i} and Y_{2i} are both observed, (ii) Y_{1i} is observed, but Y_{2i} is censored, (iii) Y_{1i} is censored, but exists, so Y_{2i} is taken to be censored at zero, and (iv) Y_{1i} is censored, and cured, so Y_{2i} doesn't exist. We introduce, for each individual, an allocation variable, q_i, which is an indicator function taking the value 1 if the individual is susceptible to post-randomisation seizure recurrence, and zero if the individual is immune. Let δ_{ji} be the indicator function for the jth survival time, taking the value 1 if the seizure is observed, and zero if the survival time is censored. These censoring scenarios must be considered separately

and combined with the censoring indicators and allocation variable, q_i, for the formulation of the log-likelihood. The log-likelihood for the observed data $\mathcal{D}$, for all the n individuals, is subsequently given by

$$\ell(\alpha, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2 \mid \mathcal{D}) = \sum_{i=1}^{n} \left\{ \left[\sum_{k=0}^{x_i-1} \ln(\alpha + k) \right] + x_i \ln(u_i) - \ln(x_i!) + \alpha \ln(\alpha) \right.$$

$$\begin{aligned}
&+(x_i + \delta_{1i}(1+\delta_{2i}))\ln(\lambda_i) + \delta_{1i}(1+\delta_{2i})\ln(\psi_i) + \delta_{1i}\ln(p_{1i}) \\
&+\delta_{1i}\delta_{2i}\ln(p_{2i}) + \delta_{1i}\ln(x_i+\alpha) + \delta_{1i}\delta_{2i}\ln(x_i+\alpha+1) \\
&-\delta_{1i}\delta_{2i}(x_i+\alpha+2)\ln(\lambda_i u_i + \lambda_i\psi_i(y_{1i}+y_{2i}) + \alpha) \\
&+(1-\delta_{1i})q_i \ln\left(\frac{p_{1i}}{(\lambda_i u_i + \lambda_i\psi_i y_{1i} + \alpha)^{x_i+\alpha}}\right) \\
&+(1-\delta_{1i})(1-q_i)\ln\left(\frac{1-p_{1i}}{(\lambda_i u_i + \alpha)^{x_i+\alpha}}\right) \\
&+\delta_{1i}(1-\delta_{2i})\ln\left(\frac{1-p_{2i}}{(\lambda_i u_i + \lambda_i\psi_i y_{1i} + \alpha)^{x_i+\alpha+1}}\right. \\
&\left.\left.+\frac{p_{2i}}{(\lambda_i u_i + \lambda_i\psi_i(y_{1i}+y_{2i}) + \alpha)^{x_i+\alpha+1}}\right)\right\}, \qquad (1)
\end{aligned}$$

It is straightforward to obtain the first and second derivatives of the log-likelihood, allowing inference on the parameters α, $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, using numerical methods. As we do not observe the allocation variable, q_i, for all i, we use an EM algorithm to maximise the log-likelihood (full details can be found in Appendix B).

The exploratory analysis suggested that clustering of seizures may be evident. There is also evidence to suggest that covariate effects may be different for times to first seizure and times from first to second seizure. To allow for this, we can consider a joint model that includes seizure rate modifiers, both at randomisation and following a first post-randomisation seizure. In this scenario, the joint probability distribution function for the survival times, given the random effects, is

$$f_{Y_1,Y_2|\nu}(y_{1i}, y_{2i} \mid \nu_i; \lambda_i, \psi_i, p_{1i}, p_{2i}) = p_{1i}p_{2i}(\lambda_i\psi_{1i}\nu_i)^2\psi_{2i}\exp(-\lambda_i\psi_{1i}\nu_i(y_{1i}+\psi_{2i}y_{2i})),$$

where now $\psi_{1i} = \exp(\boldsymbol{\beta}_2'\boldsymbol{z}_{2i})$, $\psi_{2i} = \exp(\boldsymbol{\beta}_3'\boldsymbol{z}_{3i})$ and $\boldsymbol{z}_{2i}$, $\boldsymbol{z}_{3i}$ are vectors of covariates, not necessarily distinct. Now, the log-likelihood for the observed data $\mathcal{D}$, for all the n individuals, is given by

$$\ell(\alpha, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2 \mid \mathcal{D}) = \sum_{i=1}^{n} \left\{ \left[\sum_{k=0}^{x_i-1} \ln(\alpha + k) \right] + x_i \ln(u_i) - \ln(x_i!) + \alpha \ln(\alpha) \right.$$

$$+\{x_i + \delta_{1i}(1 + \delta_{2i})\} \ln(\lambda_i) + \delta_{1i}(1 + \delta_{2i}) \ln(\psi_{1i}) + \delta_{1i}\delta_{2i} \ln(\psi_{2i})$$

$$+\delta_{1i} \ln(p_{1i}) + \delta_{1i}\delta_{2i} \ln(p_{2i}) + \delta_{1i} \ln(x_i + \alpha) + \delta_{1i}\delta_{2i} \ln(x_i + \alpha + 1)$$

$$-\delta_{1i}\delta_{2i}(x_i + \alpha + 2) \ln(\lambda_i u_i + \lambda_i \psi_{1i} y_{1i} + \lambda_i \psi_{1i} \psi_{2i} y_{2i} + \alpha)$$

$$+(1 - \delta_{1i}) q_i \ln \left(\frac{p_{1i}}{(\lambda_i u_i + \lambda_i \psi_{1i} y_{1i} + \alpha)^{x_i+\alpha}} \right)$$

$$+(1 - \delta_{1i})(1 - q_i) \ln \left(\frac{1 - p_{1i}}{(\lambda_i u_i + \alpha)^{x_i+\alpha}} \right)$$

$$+\delta_{1i}(1 - \delta_{2i}) \ln \left(\frac{1 - p_{2i}}{(\lambda_i u_i + \lambda_i \psi_{1i} y_{1i} + \alpha)^{x_i+\alpha+1}} \right.$$

$$\left. \left. + \frac{p_{2i}}{(\lambda_i u_i + \lambda_i \psi_{1i} y_{1i} + \lambda_i \psi_{1i} \psi_{2i} y_{2i} + \alpha)^{x_i+\alpha+1}} \right) \right\}. \tag{2}$$

5 Results

The joint model that allowed the seizure rate to be modified both at randomisation and following a first post-randomisation seizure was used to analyse the MESS data (Equation (2)). Further results, obtained when a marginal Negative Binomial Generalised Linear Model was fitted to the count data and standard survival models were applied to the two survival times, have been considered, but are presented elsewhere [24].

The estimated regression coefficients for the joint model are given in Table 1. The term λ_i contains parameter estimates corresponding to the effect of covariates on the underlying event rate. The terms ψ_{1i} and ψ_{2i} contain parameter estimates corresponding to the effect of covariates on the post-randomisation reduction in event rates at randomisation and following a first seizure post-randomisation respectively. A positive (negative) regression coefficient would indicate an increased (decreased) seizure rate relative to the seizure rate in the reference group. A positive (negative) κ regression coefficient would indicate an increase (decrease) in the susceptible proportion relative to the susceptible proportion in the reference group. The reference group contains individuals presenting with partial seizures pre-randomisation, a normal EEG and randomised to deferred treatment.

Interpretations of the regression coefficients is most straightforward once they have been used to calculate estimated pre-randomisation seizure rates, $\widehat{\lambda}_i$ cure rates, $\widehat{p}_{1i}$, $\widehat{p}_{2i}$ and seizure rate modifiers,

$\widehat{\psi}_{1i}$, $\widehat{\psi}_{2i}$ (Tables 2, 3 and 4 respectively).

Table 1: Estimated regression coefficients in λ_i, ψ_{1i} and ψ_{2i} for the full joint model.

	Regression Coefficient	Estimates (standard errors)
	α	2.023 (0.107)
λ_i	$\beta_{1,0}$	-4.145 (0.086)
	$\beta_{1,t-c}$	-1.076 (0.096)
	$\beta_{1,2^\circ t-c}$	-0.701 (0.098)
	$\beta_{1,partial}$	reference
ψ_i	$\beta_{2,0}$	-0.958 (0.320)
	$\beta_{2,trt}$	0.307 (0.347)
	$\beta_{2,t-c}$	0.577 (0.334)
	$\beta_{2,2^\circ t-c}$	0.483 (0.344)
	$\beta_{2,partial}$	reference
	$\beta_{2,eeg}$	0.595 (0.350)
	$\beta_{2,t-c\times trt}$	-0.468 (0.352)
	$\beta_{2,2^\circ t-c\times trt}$	-0.594 (0.362)
	$\beta_{2,partial\times trt}$	reference
	$\beta_{2,eeg\times trt}$	-0.593 (0.202)
	$\beta_{2,t-c\times eeg}$	-0.518 (0.363)
	$\beta_{2,2^\circ t-c\times eeg}$	-0.361 (0.374)
	$\beta_{2,partial\times eeg}$	reference
ψ_{2i}	$\beta_{3,0}$	1.537 (0.524)
	$\beta_{3,trt}$	-0.393 (0.537)
	$\beta_{3,t-c}$	-1.219 (0.544)
	$\beta_{3,2^\circ t-c}$	-0.590 (0.551)
	$\beta_{3,partial}$	reference
	$\beta_{3,eeg}$	-0.820 (0.545)
	$\beta_{3,t-c\times trt}$	0.599 (0.549)
	$\beta_{3,2^\circ t-c\times trt}$	0.450 (0.555)
	$\beta_{3,partial\times trt}$	reference
	$\beta_{3,eeg\times trt}$	0.690 (0.315)
	$\beta_{3,t-c\times eeg}$	0.944 (0.564)
	$\beta_{3,2^\circ t-c\times eeg}$	-0.266 (0.573)
	$\beta_{3,partial\times eeg}$	reference
p_{1i}	$\kappa_{1,0}$	0.706 (0.359)
	$\kappa_{1,trt}$	-0.067 (0.146)
	$\kappa_{1,t-c}$	-0.750 (0.365)
	$\kappa_{1,2^\circ t-c}$	-0.979 (0.374)
	$\kappa_{1,partial}$	reference
	$\kappa_{1,eeg}$	-0.006 (0.448)
	$\kappa_{1,eeg\times trt}$	-0.582 (0.228)
	$\kappa_{1,t-c\times eeg}$	0.624 (0.458)
	$\kappa_{1,2^\circ t-c\times eeg}$	1.378 (0.478)
	$\kappa_{1,partial\times eeg}$	reference
p_{2i}	$\kappa_{2,0}$	1.037 (0.099)
-Log-likelihood (d.f.)		10855 (1302)

Individuals presenting with partial seizures pre-randomisation typically have the highest seizure

rate, mean 6 per year, with those experiencing tonic-clonic seizures only and secondary tonic-clonic seizures having statistically significantly lower rates (Table 2).

Table 2: The expected pre-randomisation seizure rate per unit time and the corresponding expected yearly seizure rate.

Seizure Type	$\widehat{\lambda}_i$ (95% C.I.)		Expected yearly rate (95% C.I.)	
Tonic-Clonic	0.0055	(0.005,0.006)	2	(1.8,2.2)
2° Tonic-Clonic	0.008	(0.007,0.009)	3	(2.6,3.3)
Partial	0.016	(0.013,0.019)	6	(4.7,6.9)

Stepwise backwards elimination concluded that the optimal joint model included treatment policy, seizure type, EEG outcome and their interactions for the change in seizure rate at randomisation and following a first seizure post-randomisation. The optimal joint model also included treatment policy, seizure type, EEG outcome and its interactions with treatment policy and seizure type when determining the proportion immune from seizure recurrence post-randomisation. The cure rate for a second post-randomisation seizure was not associated with any covariates.

Table 3 shows that those individuals with a normal EEG and presenting with partial seizures can expect to have a cure rate of around 30%, irrespective of treatment policy, whilst those with tonic-clonic seizures can expect to have a cure rate of around 50%. For those with an abnormal EEG, higher cure rates are observed for those randomised to immediate treatment, rather than deferred. Recall that Warrell *et al.* [29] put seizure recurrence after a single untreated seizure at around 80%, but Berg and Shinnar [2] stated that seizure recurrence was 50%. The results that we have observed in Table 3 may provide an explanation for the difference in these values. The overall cure rate associated with time from first to second seizure is 26%.

Table 3: The expected cure rates associated with the post-randomisation times to first seizure.

Seizure Type	$1-\widehat{p}_{1i}$ (95% C.I.) cure rate			
	Abnormal EEG			
	Immediate		Deferred	
Tonic-Clonic	0.518	(0.45,0.59)	0.360	(0.20,0.43)
2° Tonic-Clonic	0.389	(0.31,0.48)	0.250	(0.19,0.33)
Partial	0.487	(0.36,0.62)	0.332	(0.22,0.46)
	Normal EEG			
	Immediate		Deferred	
Tonic-Clonic	0.528	(0.47,0.59)	0.511	(0.45,0.57)
2° Tonic-Clonic	0.584	(0.51,0.65)	0.568	(0.50,0.64)
Partial	0.345	(0.20,0.52)	0.330	(0.19,0.50)

Table 4 shows the maximum likelihood estimates for the changes in seizure rate following randomisation, and the changes in rate following first post-randomisation seizure. We again see that treatment policy does not appear to have a statistically significant effect on seizure rate for those individuals with a normal EEG. Additionally, for those individuals with an abnormal EEG, immediate treatment is favoured.

Table 4: The expected change in seizure rates following randomisation and following the first post-randomisation seizure.

Seizure Type	$\widehat{\psi}_{1i}$ (95% C.I.) first seizure rate change			
	Abnormal EEG			
	Immediate		Deferred	
Tonic-Clonic	0.347	(0.26,0.47)	0.738	(0.57,0.95)
2° Tonic-Clonic	0.326	(0.24,0.44)	0.786	(0.58,1.06)
Partial	0.522	(0.31,0.88)	0.695	(0.41,1.18)
	Normal EEG			
	Immediate		Deferred	
Tonic-Clonic	0.582	(0.45,0.75)	0.683	(0.53,0.87)
2° Tonic-Clonic	0.467	(0.34,0.65)	0.622	(0.46,0.84)
Partial	0.522	(0.27,1.00)	0.384	(0.20,0.73)
	$\widehat{\psi}_{2i}$ (95% C.I.) second seizure rate change			
	Abnormal EEG			
	Immediate		Deferred	
Tonic-Clonic	3.806	(2.43,5.97)	1.554	(1.00,2.41)
2° Tonic-Clonic	1.835	(1.17,2.88)	0.870	(0.56,1.36)
Partial	2.754	(1.17,6.49)	2.047	(0.98,4.26)
	Normal EEG			
	Immediate		Deferred	
Tonic-Clonic	1.688	(1.13,2.51)	1.373	(0.93,2.02)
2° Tonic-Clonic	2.727	(1.69,4.41)	2.577	(1.67,3.97)
Partial	3.138	(1.28,7.70)	4.649	(1.63,13.27)

Following a first seizure post-randomisation we see that, in general, seizure rates increase. We observe that those individuals allocated to immediate treatment see more of an increase in seizure rate following a first seizure post-randomisation than those allocated to deferred treatment. A possible explanation for this may be that those allocated to deferred treatment may subsequently be started on a course of AEDs following a seizure post-randomisation, bringing them in line with those allocated to immediate treatment. Also note that these individuals typically experience the smallest reduction in seizure rate following randomisation.

As an illustration, consider a person presenting with tonic-clonic seizures only pre-randomisation, with an abnormal EEG and randomised to deferred treatment. Their expected pre-randomisation seizure rate per unit time, $\widehat{\lambda}_i$, is 0.0055, equating to a seizure approximately every 182 days (Table

2). Table 3 tells us that they have a 64% $(1-\widehat{p}_{1i})$ chance of having a seizure post-randomisation, and if they were to have a seizure, we would expect this to be at $(\widehat{\lambda}_i\widehat{\psi}_{1i})^{-1} = 246$ days. If they have a seizure post-randomisation, they have a 74% $(1-\widehat{p}_{2i})$ chance of experiencing a second, and we would expect this second seizure to occur at $(\widehat{\lambda}_i\widehat{\psi}_{1i}\widehat{\psi}_{2i})^{-1} = 158$ days.

Standard software in R [23] allows the fitting of various parametric mixture models, including the Log-logistic, for the estimation of cure rates. A Log-logistic mixture model, with shape γ and scale $1/\mu_i$, that incorporates a cure fraction $(1-p_i)$, was applied to both times to first seizure and from first to second sezure post-randomisation. Table 5 shows that the cure rates associated with times to first seizure post-randomisation are dependent on seizure type, EEG outcome and the logarithm of the pre-randomisation seizure rate for each individual. For the times from first to second post-randomisation, no covariate effects were found to be statistically significant in determining the proportion cured.

Model checking has been carried out in the form of diagnostic plots that considered how well fitted estimates from the joint model and Log-logistic with a cure fraction modelled the distribution of the survival curves [24]. A method for testing goodness-of-fit of parametric distributions to survival data was developed by Maller and Zhou [19], which is a variant to the method devised by Filliben [11] for testing the normality of uncensored data. The test for censored survival times considers the hypothesis $H_0 : F = \widehat{F}$, where $\widehat{F}$ is some specified distribution function. When censoring is present, a plot of $\widehat{F}$ against the Kaplan-Meier estimate, $\widetilde{F}$, under H_0, should produce a near-straight line with slope close to 1. The full joint model was found to be at least as good as the Log-logistic model that included a cure rate in estimating the survivor functions for the post-randomisation survival times when these methods for model checking were applied. Additionally, Cowling [7] carried out a simulation study of power which indicated that the joint modelling strategy provides more precise estimates of treatment effects than standard parametric models do.

6 Discussion

The inclusion of additional information in our joint model of pre-randomisation event counts and post-randomisation survival times resulted in an increase in power over standard parametric survival methods, which consequently meant that statistically significant covariate effects could be affirmed. Previously analyses of the MESS data concluded that the risk of seizure recurrence increased with the number of seizures pre-randomisation and an abnormal EEG, and that immediate treatment

Table 5: Estimated regression coefficients for the Log-logistic cure rate model for time to first seizure and time from first to second seizure.

	Regression Coefficient	Estimates (standard errors)			
		Time to First Seizure		Time from First Second Seizure	
	γ	2.133	(0.038)	1.823	(0.042)
μ_i	θ_0	-0.925	(1.031)	-0.213	(1.345)
	θ_{trt}	-1.925	(1.325)	-1.545	(1.683)
	θ_{t-c}	-0.052	(0.629)	-2.218	(0.871)
	$\theta_{2^\circ t-c}$	0.158	(0.645)	-1.073	(0.860)
	θ_{par}	reference		reference	
	$\theta_{\ln(rate)}$	0.913	(0.198)	0.514	(0.272)
	θ_{eeg}	-0.966	(1.297)	-1.595	(1.583)
	$\theta_{t-c\times trt}$	-0.905	(0.656)	1.040	(0.912)
	$\theta_{2^\circ t-c\times trt}$	-0.747	(0.640)	0.478	(0.905)
	$\theta_{par\times trt}$	reference		reference	
	$\theta_{\ln(rate)\times trt}$	-0.484	(0.249)	-0.128	(0.336)
	$\theta_{eeg\times trt}$	-1.117	(0.348)	0.507	(0.477)
	$\theta_{t-c\times eeg}$	-0.697	(0.707)	0.550	(0.927)
	$\theta_{2^\circ t-c\times eeg}$	-0.442	(0.726)	0.259	(0.925)
	$\theta_{par\times eeg}$	reference		reference	
	$\theta_{\ln(rate)\times eeg}$	0.026	(0.244)	-0.166	(0.320)
p_i	κ_0	3.145	(0.885)	1.434	(0.191)
	κ_{t-c}	-0.679	(0.623)		
	$\kappa_{2^\circ t-c}$	-1.099	(0.625)		
	κ_{par}	reference			
	$\kappa_{\ln(rate)}$	0.390	(0.140)		
	κ_{eeg}	-0.650	(0.684)		
	$\kappa_{t-c\times eeg}$	1.464	(0.772)		
	$\kappa_{2^\circ t-c\times eeg}$	2.352	(0.834)		
	$\kappa_{par\times eeg}$	reference			
$-$Log-likelihood (d.f.)		1979	(1313)	1178	(1319)

increased time to first and second seizures [20, 16]. These findings are consistent with our analysis of the MESS data. Nonetheless, neither of these analyses considered differences between types of epileptic seizures, or interactions between the covariates, which we have found to be statistically significant in determining underlying seizure rates and post-randomisation seizure rate modifiers.

It was proposed that there may be a proportion of the MESS data not susceptible to post-randomisation seizure recurrence. For this reason, cure rate models were considered for the post-randomisation survival times, as an alternative to standard survival distributions. The intial joint model assumed that, post-randomisation, an individuals' seizure rate remained constant. This assumption was relaxed to allow the seizure rate to change both at randomisation, and following a first post-randomisation seizure.

Those individuals with a normal EEG, presenting with partial seizures can expect to have a cure rate of around 30%, irrespective of treatment policy, whilst those with tonic-clonic seizures can expect to have a cure rate of around 50%. For those with an abnormal EEG, higher cure rates are observed for those individuals randomised to immediate treatment rather than deferred. The cure rate for a second post-randomisation seizure was found to be dependent on an intercept term only. Treatment policy was not statistically significant in determining the estimate of the change in seizure rate following randomisation for those individuals with a normal EEG. For those individuals with an abnormal EEG, immediate treatment was favoured. Following a first seizure post-randomisation we observed a general increase in seizure rates, with those allocated to deferred treatment having the smallest increase in seizure rate following a first seizure post-randomisation. Possible explanations for this observation were discussed.

In addition to the diagnostic plots which we examined, a simulation study of power, discussed in Cowling [7], that assessed a joint model for event counts and post-randomisation times to first seizure, provided more precise estimates of treatment effects than standard parametric survival models. The methodology presented here belongs to the accelerated failure time family of distributions. Misspecification of parametric accelerated failure time models has been considered by Hutton and Monaghan [14], who concluded that accelerated failure time models are robust to model misspecification because of their log-linear form and are more robust than proportional-hazards models.

Joint modelling strategies have been used by Rogers *et al.* [26] to assess the risk of tonic-clonic seizures in patients with a history of partial seizures. Here, pre-randomisation seizure rates and post-randomisation times to first seizure of any type and first tonic-clonic seizure were analysed using joint models that incorporated cure fractions. It was concluded, using these methods, that patients presenting with a history of only partial seizures are at a low risk of subsequent tonic-clonis seizures. It was also found that the effects of the AEDs used in the MESS study are greater for tonic-clonic seizures than they are for partial seizures.

The statistical models that have been developed in this paper can be used to analyse general recurrent event data in the form of gap times, with associated baseline count data. Further work may consider the use of a Poisson distribution for the pre-randomisation event counts that is truncated at zero and allows for the excess of ones observed in the MESS dataset. The performance of the joint model was compared with the best fitting survival model. This was a visual comparison of diagnostic plots which did not, however, formally assess the performance of the full joint model in its own right.

Hence, methods of model checking could also be investigated further.

Acknowledgements

The authors thank Tony Marson and David Chadwick for providing the MRC Multicentre Trial in Early Epilepsy and Single Seizures. We thank these investigators further for their additional advice. The authors would also like to thank Karla Hemming for her statistical input in the early stages of the methodological development. I thank the Engineering and Physical Research Council for funding my PhD, which allowed me to carry out this work.

References

[1] M. Akacha and N. Benda, *The impact of dropouts on the analysis of dose-finding studies with recurrent event data*, Statistics in Medicine 29 (2010), pp. 1635–1646.

[2] A.T. Berg and S. Shinnar, *The risk of seizure recurrence following a first unprovoked seizure: A quantitative review*, Neurology 41 (1991), pp. 965–972.

[3] B. Chandra, *First seizure in adults: to treat or not to treat*, Clinical Neurology and Neurosurgery 94(Suppl.) (1992), pp. S61–S63.

[4] O.C. Cockerell, A.L. Johnson, J.W.A.S. Sander, Y.M. Hart, and S.D. Shorvon, *Remission of epilepsy: results from the national general practice study of epilepsy*, The Lancet 346 (1995), pp. 140–144.

[5] R.J. Cook and J.F. Lawless, *The Statistical Analysis of Recurrent Events*, Statistics for Biology and Health Vol. , Springer, 2007.

[6] R.J. Cook and W. Wei, *Conditional analysis of mixed poisson processes with baseline counts: implications for trial design and analysis*, Biostatistics 4 (2003), pp. 479–494.

[7] B.J. Cowling, *Survival models for censored point processes*, PhD Thesis, Department of Statistics, University of Warwick, Coventry, 2003.

[8] B.J. Cowling, J.L. Hutton, and J.E.H. Shaw, *Joint modelling of event counts and survival times*, J. R. Statist. Soc. C 55 (2006), pp. 31–39.

[9] P.J. Diggle, I. Sousa, and A.G. Chetwynd, *Joint modelling of repeated measurements and time-to-event outcomes: The fourth Armitage lecture*, Statistics in Medicine 27 (2008), pp. 2981–2998.

[10] L. Duchateau, P. Janssen, I. Kezic, and C. Fortpied, *Evolution of recurrent asthma event rate over time in frailty models*, J. R. Statist. Soc. C 52 (2003), pp. 355–363.

[11] J.J. Filliben, *The probability plot correlation coefficient test for normality*, Technometrics 11 (1975), pp. 111–117.

[12] M. Greenwood and G.U. Yule, *An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurence of multiple attacks of disease or of repeated accidents*, Journal of the Royal Statistical Society 83 (1920), pp. 255–279.

[13] P. Hougaard, M. Lee, and G.A. Whitmore, *Analysis of overdispersed count data by mixtures of poisson variables and poisson processes*, Biometrics 53 (1997), pp. 1225–1238.

[14] J.L. Hutton and P.F. Monaghan, *Choice of parametric accelerated life and proportional hazards models for survival data: asymptotic results*, Lifetime Data Analysis 8 (2002), pp. 375–393.

[15] ILAE Commission on Antiepileptic Drugs, *Considerations on designing clinical trials to evaluate the place of new antiepileptic drugs in the treatment of newly diagnosed and chronic patients with epilepsy*, Epilepsia 39 (1998), pp. 799–803.

[16] L. Kim, A. Johnson, A. Marson, and D. Chadwick, *Prediction of risk of seizure recurrence after a single seizure and early epilepsy: further results from the MESS trial*, The Lancet Neurology 5 (2006), pp. 317–322.

[17] D.Y. Lin, W. Sun, and Z. Ying, *Nonparametric estimation of the gap time distributions for serial events with censored data*, Biometrika 86 (1999), pp. 59–70.

[18] R.J.A. Little and D.B. Rubin, *Statistical Analysis with Missing Data*, Wiley series in Probability and Statistics Vol. , John Wiley and Sons, 2002.

[19] R. Maller and X. Zhou, *Survival Analysis with Long-Term Survivors*, Wiley series in Probability and Statistics Vol. , John Wiley and Sons, 1996.

[20] A. Marson, A. Jacoby, A. Johnson, L. Kim, C. Gamble, and D. Chadwick, *Immediate versus deferred antiepileptic drug treatment for early epilepsy and single seizures: a randomised control trial*, The Lancet 365 (2005), pp. 2007–2013.

[21] P. McCullagh and J.A. Nelder, *Generalised Linear Models*, Monographs on Statistics and Applied Probability Vol. , Chapman and Hall, 1989.

[22] T.K. Nayak, *Multivariate lomax distribution: properties and usefulness in reliability theory*, Journal of Applied Probability 24 (1987), pp. 170–177.

[23] Y. Peng, *gfcure*, (Available from http://post.queensu.ca/∼pengp/ software.html), downloaded 14 July 2009.

[24] J.K. Rogers, *Statistical models for censored point processes with cure fractions*, PhD Thesis, Department of Statistics, University of Warwick, Coventry, 2011.

[25] J.K. Rogers, J.L. Hutton, and K. Hemming, *Joint Modelling of Event Counts and Survival Times*, CRiSM Working Paper 44, University of Warwick, (Available from http://www2.warwick.ac.uk/fac/sci/statistics/crism/research/2009/ paper09-44), 2009.

[26] J.K. Rogers, J. Hutton, A.G. Marson, and D.W. Chadwick, *Assess the risk of subsequent tonic-clonic seizures in patients with a history of simple or complex partial seizures*, J. Neurol. Neurosurg. Psychiatry 83 (2012), pp. 803–809.

[27] C.M. Verity, G. Hosking, and D.J. Easter, *A multicentre comparitive trial of sodium valproate and carbamazepine in paediatric epilepsy*, Developmental Medicine and Child Neurology 37 (1995), pp. 97–108.

[28] M. Visser, *Nonparametric estimation of the bivariate survival function with an application to vertically transmitted AIDS*, Biometrika 83 (1996), pp. 507–518.

[29] D.A. Warrell, T.M. Cox, J.D. Firth, and E.J. Benz (eds.), Vol. , Oxford University Press, 2003.

7 Appendix A. The Lomax distrubution

We assume that the pre-randomisation event count, X_i, over a period u_i for individual i, $i = 1, \ldots n$, follows a Poisson distribution with rate $\lambda_i \nu_i$, with additional heterogeneity in the population being modelled through ν_i, assumed to follow a Gamma distribution with expectation 1 and variance $1/\alpha$. If the random effect term is integrated out of the joint density of X_i and ν_i, the resulting unconditional

density, $f_X(x_i;\lambda_i,u_i,\alpha)$, is simply the Negative Binomial, specified by the following equation:

$$f_X(x_i;\lambda_i,u_i,\alpha)=\frac{\Gamma(x_i+\alpha)}{x_i!\Gamma(\alpha)}\left(\frac{\lambda_i u_i}{\alpha+\lambda_i u_i}\right)^{x_i}\left(\frac{\alpha}{\alpha+\lambda_i u_i}\right)^{\alpha},$$

where $\lambda_i=\exp(\boldsymbol{\beta}_1'\boldsymbol{z}_{1i})$. Here $\boldsymbol{z}_{1i}$ is a vector of covariates for individual i, and $\boldsymbol{\beta}_1$ is a vector of regression coefficients, including an intercept term.

The response variables Y_{1i} and Y_{2i} are the post-randomisation times to first seizure and times from first to second seizure respectively. Both Y_{1i} and Y_{2i} will be independent given ν_i and Exponentially distributed with rate $\lambda_i\psi_i\nu_i$. The unconditional joint distribution of the Y_{ji}, $j=1,2$, obtained when the random effect term is integrated out of the joint density of Y_{1i}, Y_{2i} and ν_i, has the following density and survivor functions:

$$\begin{aligned}
f_{Y_1,Y_2}(y_{1i},y_{2i};\lambda_i,\psi_i,\alpha) &= \int_0^\infty f_{Y_1,Y_2|\nu}(y_{1i},y_{2i}\mid\nu_i;\lambda_i,\psi_i)g_\nu(\nu_i;\alpha)d\nu_i \\
&= \frac{\alpha+1}{\alpha}(\lambda_i\psi_i)^2\left\{1+\frac{\lambda_i\psi_i(y_{1i}+y_{2i})}{\alpha}\right\}^{-(\alpha+2)},
\end{aligned}$$

$$\begin{aligned}
S_{Y_1,Y_2}(y_{1i},y_{2i};\lambda_i,\psi_i,\alpha) &= \int_{y_{2i}}^\infty\int_{y_{1i}}^\infty f_{Y_1,Y_2}(u,v;\lambda_i,\psi_i,\alpha)du\,dv \\
&= \left\{1+\frac{\lambda_i\psi_i(y_{1i}+y_{2i})}{\alpha}\right\}^{-\alpha}.
\end{aligned}$$

Each of the Y_{ji} also have univariate Lomax marginal distributions, each with density:

$$f_{Y_j}(y_{ji};\lambda_i,\psi_i,\alpha)=\frac{\lambda_i\psi_i}{(1+\lambda_i\psi_i y_{ji}/\alpha)^{\alpha+1}},\qquad j=1,2.$$

8 Appendix B. The EM algorithm

The EM algorithm is a general iterative algorithm for maximum likelihood estimation, in the presence of missing data [18]. This technique is based on a somewhat ad-hoc idea of: (1) Replacing missing values by estimated values, (2) Estimating parameters, (3) Re-estimating the missing values according to the new parameter estimates, (4) Re-estimating the parameters, iterating until convergence.

Suppose that the complete data are Y, with associated density $f(Y;\theta)$. We write $Y=(Y_{obs},Y_{mis})$, where Y_{obs} represents the observed part of Y, and Y_{mis} denotes the missing part. The objective is to

maximise the ignorable likelihood,

$$L_{ign}(\theta \mid Y_{obs}) = \int f(Y_{obs}, Y_{mis}; \theta) dY_{mis},$$

with respect to θ.

The EM algorithm comprises an E-step and an M-step. The E-step finds the conditional expectation of the missing data, given the observed data and the current estimated parameters. These expectations are then substitutes for the missing data. In reality, the EM algorithm doesn't substitute the missing values themselves, but regards the functions of Y_{mis} appearing in the complete-data log-likelihood, as missing. The subsequent M-step simply finds the maximum likelihood parameter estimates, as if there were no missing data.

The algorithm is formally defined as:

1. Choose initial value $\theta^{(0)}$, set t=0.

2. E-step: calculate

$$\begin{aligned} \mathfrak{Q}(\theta, \theta^{(t)}) &= \mathbb{E}[\ln(L(\theta \mid Y_{obs}, Y_{mis}))] \\ &= \int \ell(\theta \mid Y) f(Y_{mis} \mid Y_{obs}, \theta = \theta^{(t)}) dY_{mis} \end{aligned} \tag{3}$$

   (at this stage $\theta^{(t)}$ is fixed and $\mathfrak{Q}(\theta, \theta^{(t)})$ is a function of θ).

3. M-step: find $\theta^{(t+1)}$ which maximises $\mathfrak{Q}(\theta, \theta^{(t)})$ as a function of θ.

4. Set $t = t + 1$ and go to step 2.

The parameters to be maximised are $\theta = \{\alpha, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2\}$. The observed data are comprised of $Y_{obs} = \{X, Y_1, Y_2, Z_1, Z_2, Z_3, W_1, W_2, \delta_1, \delta_2\}$ and $Y_{mis} = \{Q\}$ is the missing data. To implement the EM-algorithm, we first carry out the E-step and calculate $\mathfrak{Q}(\theta, \theta^{(t)}) = \mathbb{E}[\ln(L(\theta \mid Y_{obs}, Y_{mis}))]$. Completion of the E-Step requires the derivation of an expression for $\mathbb{E}(q_i \mid Y_{obs}, \theta^{(t)})$, the expected value of q_i, given the observed data and the current values of the parameters of interest. Clearly, as q_i is an indicator variable, we subsequently have $\mathbb{E}(q_i \mid Y_{obs}, \theta^{(t)}) = \mathbb{P}(q_i = 1 \mid Y_{obs}, \theta^{(t)})$. Recall that $q_i = 1$ corresponds to an individual having a post-randomisation time to first seizure that is censored, but exists. We can now formulate an expression for $\mathbb{E}(q_i \mid Y_{obs}, \theta^{(t)})$ as

$$\mathbb{E}(q_i \mid Y_{obs}, \theta^{(t)}) = \mathbb{P}(q_i = 1 \mid Y_{obs}, \theta^{(t)})$$

$$
\begin{aligned}
&= p_{1i} R_{Y_1}(y_{1i}; \lambda_i, \psi_{1i}, \alpha) \\
&= p_{1i} \left(1 + \frac{\lambda_i \psi_{1i} y_{1i}}{\alpha}\right)^{-\alpha},
\end{aligned}
$$

where $R(\cdot)$ is the proper survivor function for the susceptible proportion.