

Regression Analysis of Recurrent Event Data

Jerry Lawless
jlawless@uwaterloo.ca

Talk at SickKids Hospital
November 23, 2009

Some examples involving recurrent events
Models and features of recurrent event processes
Some methods of analysis
Illustrations
Some complications for analysis

1. Examples Involving Recurrent Events

Mammary Tumors in Rats (Gail et al., 1980)

- Carcinogenicity experiment in which 48 rats are exposed to a carcinogen
- 23 assigned to a treatment group and 25 to a control group
- Careful examination over 122 days to detect development of new tumors
- Multiple tumors could develop

Figure 1: Timeline plots for data from Gail et al. (1980)

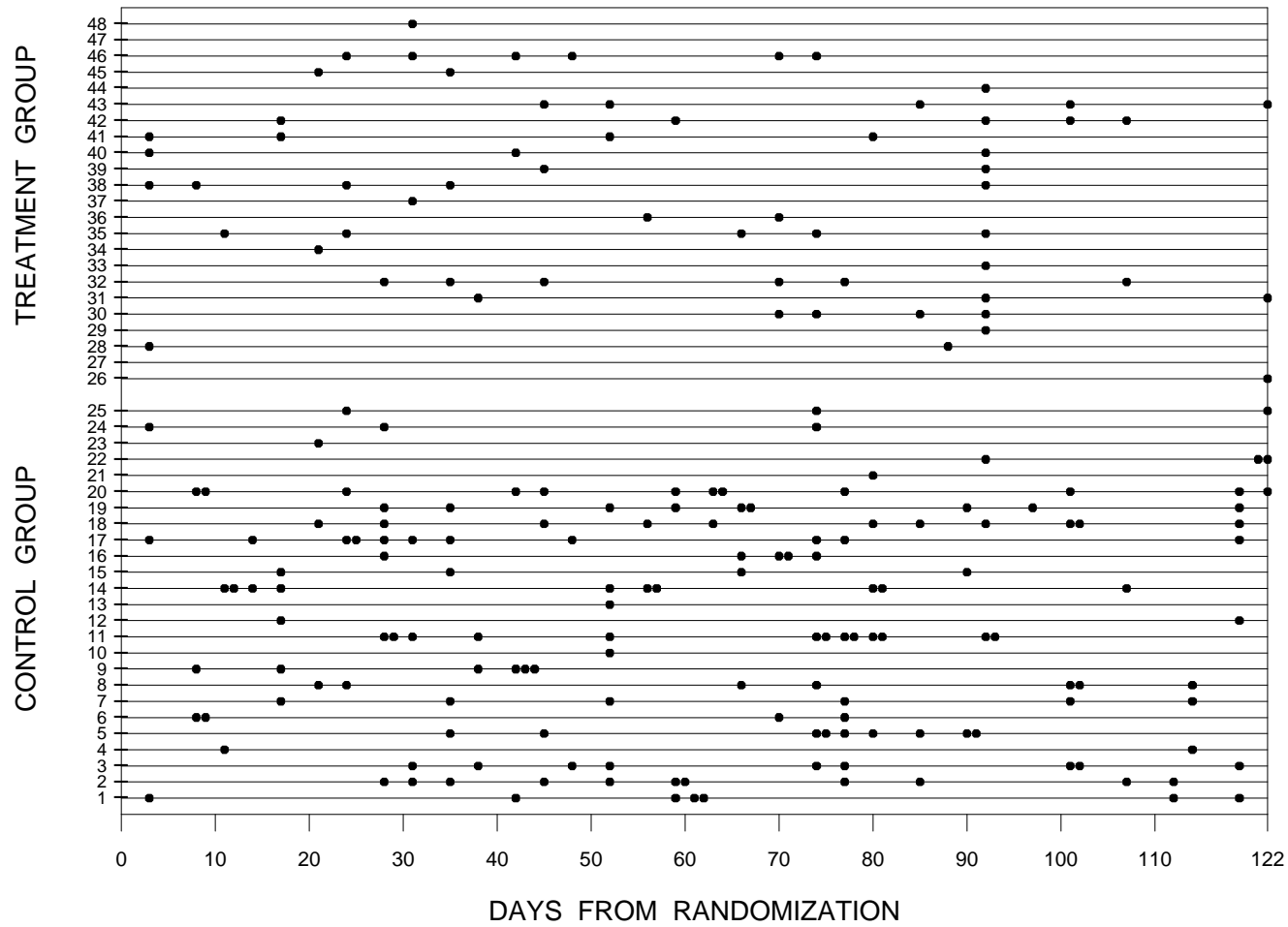
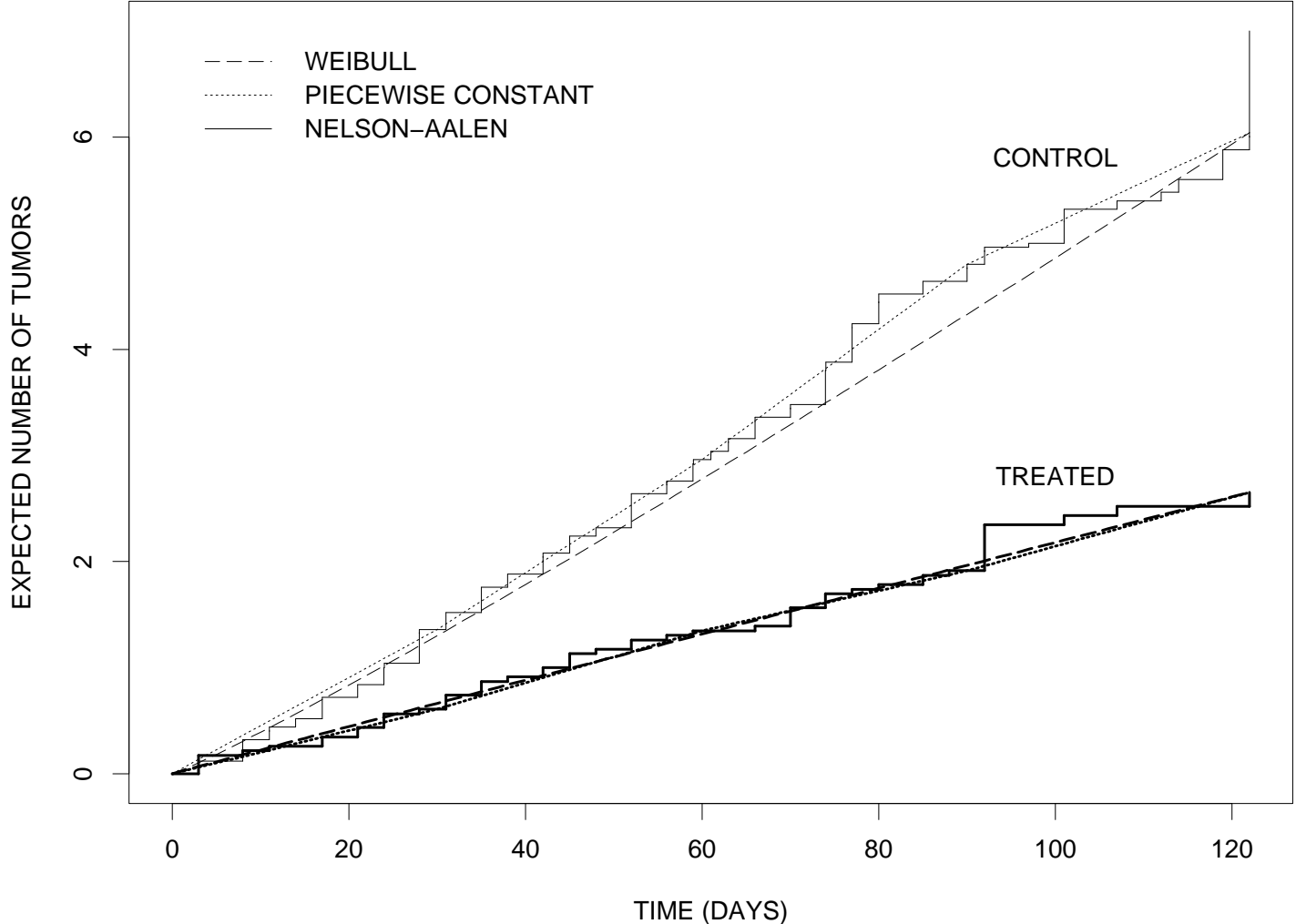


Figure 2. Average number of tumors up to time t on study



Pulmonary Exacerbations in Cystic Fibrosis (Fuchs et al., 1994)

- patients with cystic fibrosis are at risk of pulmonary exacerbations, which arise from fluid accumulation in lungs
- Fuchs et al. (1994) conducted a randomized trial of a treatment rhDNase, intended to reduce exacerbations, versus placebo
- 321 and 324 randomized to rhDNase and placebo respectively
- followup occurred over roughly 169 days

Table 1: Distribution of the numbers of exacerbations by treatment group for subjects in the rhDNase study

Number of Exacerbations	Number of Patients	
	Placebo Group	rhDNase Group
0	185	217
1	97	65
2	24	30
3	13	6
4	4	3
5	1	0

Asthma Prevention Trial (Duchateau et al., 2003)

- Six-month old children at high risk of asthma
- Randomized to a treatment or control and followed for 18 months
- Similar to rhDNase study
- In both studies a subject who has an event (infection or attack) is not considered at risk for a new event until the effects of the preceding event have cleared

Cerebrospinal Fluid Shunt Failures (Tuli et al., 2000)

- Clinical data on recurrent failures of shunts that drain excess cerebrospinal fluid for 839 children with hydrocephalus (SickKids)
- Failure necessitates (partial) shunt replacement
- Failures classified according to cause: obstruction (70%), infection (15%), other (15%)
- Followup ranges from 1 to 11 years
- Numerous potential risk factors (etiology, age, shunt type, ...)
- Number of failures per patient ranges from 0 (386 patients) to 11 or more (14 patients)

2. Models and Features of Recurrent Event Processes

- Modeling and analysis of data should be guided by objectives
- Common objectives include
 - understanding and describing features of recurrent event processes (patterns over time, probabilities, dynamics)
 - Comparison of groups of processes (e.g. clinical trials)
 - determining the effects of fixed or time-varying covariates
- Notation:
 - $N_i(t)$ = number of events in $(0, t)$ for subject i
 - $0 \leq T_{i1} < T_{i2} \leq \dots \leq T_{in_i}$: times of events
 - x_i or $x_i(t)$: covariates; followup periods $(0, \tau_i]$

Features of recurrent event processes

- mean functions: $\mu_i(t) = E\{N_i(t)\}$
 - corresponding rate function is $\rho_i(t) = \mu'_i(t)$
- gap times distribution

$$W_{ij} = T_{ij} - T_{i,j-1} = \text{times between events } (T_{i0} = 0)$$

Are W_{i1}, W_{i2}, \dots independent, given covariates? (Rarely)

How to characterize distribution of (W_{i1}, W_{i2}, \dots) ?

- intensity function: probability of a new event at time t , given previous history of events (and covariates)

$H_i(t)$ = history of events and covariates up to t

$$\lambda_i(t|H_i(t)) = \lim_{\Delta t \downarrow 0} \frac{\Pr(N_i(t + \Delta t) - N_i(t) = 1 | H_i(t))}{\Delta t}$$

Covariate effects

- In each approach multiplicative models (especially log linear models) are very useful

rate functions: $\rho_i(t) = \rho_0(t) \exp(\beta' x_i(t))$

- if x_i are fixed, $\mu_i(t) = \mu_0(t) \exp(\beta' x_i)$

- called Andersen-Gill (AG) model

gap times: use Cox proportional hazards model for W_{ij} 's

intensity functions: use either

(a) $\lambda_i(t|H_i(t)) = \lambda_0(t) \exp(\beta' z_i(t))$

(b) $\lambda_i(t|H_i(t)) = \lambda_0(B(t)) \exp(\beta' z_i(t))$

where $B(t) = t - T_{N(t-)} =$ time since last event and $z_i(t)$ includes $x_i(t)$ and components based on $H_i(t)$

- (a) is called a modulated Poisson process and (b) a modulated renewal process

Random effects

- Sometimes a fairly simple model describes event occurrence for individuals, but there is excess variation in the rate or intensity functions across individuals
- Random effects are useful in modeling this “unobserved heterogeneity”
- Andersen-Gill (Poisson) models are especially useful: take

$$\rho_i(t) = u_i \rho_0(t) \exp(\beta' x_i)$$

where u_i is an unobserved random variable with $E(u_i) = 1$ and $\text{Var}(u_i) = \phi$. We usually assume some specific distribution for the u_i ; gamma and log-normal are most common.

3. Some Methods of Analysis

AG (Poisson) Analysis

- Based on Poisson process assumptions
- $\lambda_i(t|H_i(t)) = \rho_i(t) = \rho_0(t) \exp(\beta' x_i(t))$ (1)
R/SPlus - coxph SAS - phreg
- Data go in “counting process” format, with a new line for each period at risk of a new event (see below)
- Important extra feature: even if $\lambda_i(t|H_i(t))$ is not of the form (1) but $\rho_i(t)$ is, then the AG analysis is still valid provided (1) length of followup is independent of event occurrence and (b) robust variance estimators for $\hat{\beta}$ and $\hat{\mu}_0(t)$ are used.

- R/SPlus: use “cluster(id)” option with coxph
SAS: use “covs(aggregate)” option with phreg
- Important extra feature (R/SPlus):
coxph will fit AG models with random effects, as on slide 11; use “frailty(id)” option

Gap Time Analysis

- Use conditional Cox models for successive gap times W_{ij} : model $\Pr(W_{ij}|z_{ij})$ where z_{ij} includes covariates and previous gap times
- Can extend to include random effects (frailties)
- Naive “working independence” analysis for gap times is not valid unless the $W_{ij}(j = 1, , 2, \dots)$ are independent, which is rare

Intensity-Based Analysis

- Formulate modulated Poisson (Markov) or renewal (semi-Markov) processes for $\lambda_i(t|H_i(t))$, as on slide 10
- Allows a lot of flexibility in modeling dependence on previous event history
- Many analyses can be implemented using Cox model software.

General Remark

- Gap time and intensity-based analyses are useful for examination of recurrent event dynamics. They are not well suited to simple comparison of treatment groups, whereas (robust) AG methods are.

4. Illustrations

Rat Mammary Tumor Data

Data for the first five rats in the treated group are displayed below in “counting process” format (Therneau and Grambsch, 2000).

```
> rats[1:12, ]
  id start stop status rtrunc tstatus enum trt
1  1     0 122     1    NA        1     1   1
2  2     0 122     0    NA        1     1   1
3  3     0   3     1    NA        1     1   1
4  3     3  88     1    NA        2     2   1
5  3    88 122     0    NA        2     3   1
6  4     0  92     1    NA        1     1   1
7  4    92 122     0    NA        2     2   1
8  5     0  70     1    NA        1     1   1
9  5    70  74     1    NA        2     2   1
10 5    74  85     1    NA        2     3   1
11 5    85  92     1    NA        2     4   1
12 5    92 122     0    NA        2     5   1
```

- One or more lines per subject; separate line for each new interval “at risk” (of a new event)

- “start”, “stop” give start and end of an at risk interval; “status” is 1 if an event occurs at “stop” and 0 if this is the end of followup (no event)
- Columns rtrunc and tstatus are used with parametric models (fitted using censorReg in SPlus) but are not needed for coxph
- Code for coxph is shown below; SAS phreg code is available on the web site for Cook and Lawless (2007) at

www.stats.uwaterloo.ca/cook-lawless/book.shtml

- Model checking tools are available in coxph and phreg

The data for the first three rats in the treated group are displayed below in the so-called "counting process" format.

```
> rats[1:5, ]
      id start stop status  enum trt
1  1  1     0  122     1     1  1
2  2  2     0  122     0     1  1
3  3  3     0   3     1     1  1
4  4  3     3  88     1     2  1
5  5  3    88  122     0     3  1
```

Robust Semiparametric Analysis

```
coxph(Surv(start,stop,status) ~ trt + cluster(id),
      data=rats, method="breslow")
n= 254
      coef exp(coef) se(coef) robust se      z      p
trt -0.815774  0.442297 0.151836  0.19809 -4.11819 3.8186e-05

      exp(coef) exp(-coef) lower .95 upper .95
trt  0.442297    2.26092  0.299985  0.652122

Likelihood ratio test= 31.69 on 1 df, p=1.81146e-08
Wald test              = 16.96 on 1 df, p=3.8186e-05
Score (logrank) test = 30.54 on 1 df, p=3.26554e-08, Robust = 11.2
p=0.000816617
```

(Note: the likelihood ratio and score tests assume independence of observations within a cluster, the Wald and robust score tests do not).

Frailty model

- Robust AG analysis is for a model with rate function

$$\rho_i(t) = \rho_0(t) \exp(\beta x_i)$$

where $x_i = I$ (treatment)

- The fact the robust $se(\hat{\beta})$ is bigger than the standard AG (Poisson) $se(\hat{\beta})$ suggests unobserved heterogeneity
- Another approach is to use a random effects model with intensity function

$$\lambda_i(t|u_i) = u_i \lambda_0(t) \exp(\beta x_i)$$

Using `coxph` with the “`frailty(id)`” option gives results for $u_i \sim \text{Gamma}$ (mean = 1, variance = ϕ). This gives $\hat{\beta} = -0.816$ and $se(\hat{\beta}) = 0.211$, close to the robust analysis

rhDNase and Recurrent Pulmonary Infections

- Data are mentioned in Section 1 and Therneau and Grambsch (2000); available on Cook and Lawless (2007) website
- Subjects with cystic fibrosis were randomized to receive either a daily aerosol form of rhDNase ($n = 321$) or a placebo ($n = 324$)
- Responses were pulmonary exacerbations (infections) observed over a followup period of (approximately) 169 days
- Key question: effectiveness of rhDNase in lowering the incidence of infection
- How to use data after first event?
- Complication: subjects not at risk while an infection is being treated

Table 2: Number of Exacerbations Per Subject During Study

Number of Exacerbations	Placebo	rhDNase
0	185	217
1	97	65
2	24	30
3	13	6
4	4	3
5	1	0
	324	321

- Conditional analysis: look at W_1 and $W_2|W_1$
 - include forced expiratory volume (FEV) as a baseline covariate
 - periods where infections are treated with antibiotics are excluded

We show below part of a dataframe containing the data

- Gap times (or censoring times) are given by `time2 - time1`
- Lines with “`etype=2`” correspond to intervals where a person was taking antibiotics for an infection, and so were **not at risk for a new infection**. These lines have been omitted below.
- `enum`= the cumulative number of “**state**” changes (i.e. the start or end of an exacerbation); `enum1`= the cumulative number of **exacerbations**

Obs	id	trt	fev	fev2	time1	time2	status	etype	enum	enum1	enum2	fevc
1	493301	1	28.8	28.1	0	168	0	1	1	1	0	-32.2778
2	493303	1	64.0	63.0	0	169	0	1	1	1	0	2.9222
3	493305	0	67.2	68.7	0	65	1	1	1	1	0	6.1222
4	493305	0	67.2	68.7	75	168	0	1	3	2	1	6.1222
5	493309	1	57.6	56.5	0	168	0	1	1	1	0	-3.4778
6	493310	0	57.6	56.3	0	171	0	1	1	1	0	-3.4778
7	493311	1	25.6	25.3	0	166	0	1	1	1	0	-35.4778
8	493312	0	86.4	85.4	0	168	0	1	1	1	0	25.3222
9	493313	0	32.0	32.4	0	90	1	1	1	1	0	-29.0778
10	493313	0	32.0	32.4	104	166	0	1	3	2	1	-29.0778

Table 3: Fitted models for W_1 and for W_2 given W_1 .

Gap Time	Parameter	Cox PH		Log-normal AFT	
		EST.	S.E.	EST.	S.E.
W_1	β_{10} (intercept)	-	-	5.40	0.11
	β_{11} (trt)	-0.38	0.13	0.43	0.14
	β_{12} (FEV)	-0.021	0.003	0.022	0.003
	σ_1			1.45	0.07
W_2	β_{20} (intercept)	-	-	3.21	0.49
	β_{21} (trt)	0.36	0.23	-0.23	0.21
	β_{22} (FEV)	0.001	0.005	-0.005	0.005
	$\beta_{23}(w_1 \text{ or } \log w_1)^\dagger$	-0.014	0.004	0.42	0.13
	σ_2	-	-	1.23	0.11

$^\dagger w_1$ for PH model and $\log w_1$ for AFT model.

- Results are summarized above for Cox PH and log-normal models.
- Note strong effect of treatment (and FEV) on time W_1 to first infection
- Effects of treatment, FEV not significant in regression of $W_2|W_1$, but W_1 is highly significant
- Such conditional analyses are not good ways to assess treatment effects

Further Analysis

- difficult to assess persistence of rhDNase treatment effect with conditional models; joint models $[W_1, W_2, \dots | x]$ are better
 - coxph with “frailty(id)” and “strata(enum1)”

Consider also four modulated Markov intensity models; write

- $\lambda_i(t)$ for $\lambda(t|H_i(t))$
- $Y_i(t) = I(\text{at risk of exacerbation at time on study } t)$

Model 1: $\lambda_i(t) = Y_i(t)\lambda_0(t) \exp(\beta_1 x_{i1} + \beta_2 x_{i2})$ where $x_{i1} = I(\text{received rhDNase})$, $x_{i2} = \text{centered FEV}$

Model 2: $\lambda_i(t|u_i) = u_i Y_i(t)\lambda_0(t) \exp(\beta_1 x_{i1} + \beta_2 x_{i2})$ where u_i is gamma with $E(u_i) = 1$

Model 3: $\lambda_i(t) = u_i Y_i(t)\lambda_0(t) \exp\{\beta_1 x_{i1} I(t \leq 80) + \beta'_1 x_{i1} I(t > 80) + \beta_2 x_{i2}\}$

Model 4: $\lambda_i(t) = Y_i(t)\lambda_0(t) \exp\{\beta_1 x_{i1} I(t \leq 80) + \beta'_1 x_{i1} I(t > 80) + \beta_2 x_{i2} + \beta_3 I(N_i(t-) > 0)\}$

- These allow consideration of treatment in terms of time on study

Table 4: Estimates (S.E.) for Models 1-4

Covariate or Parameter	Model 1	Model 2	Model 3	Model 4
Treatment	-0.29 (.11)	-0.31(.13)	-	-
Treatment ($t \leq 80$)	-	-	-0.51(.18)	-0.43(.16)
Treatment ($t > 80$)	-	-	-0.16(.16)	-0.06(.14)
FEV \div 10	-0.17(.02)	-0.19(.03)	-0.19(.03)	-0.15(.02)
$I(N(t-) > 0)$	-	-	-	0.73(.13)
Variance	-	0.94(48.0*)	0.94(48.0*)	-

*LR Statistic (l.d.f.) for testing $\phi = 0$

- Indication of either heterogeneity across subjects (Models 2, 3), or of event dependence (Model 4)
- Indication of diminishing treatment effect with time on study, BUT number of subjects with more than 1 event is rather small; longer followup would be helpful
- Models 3, 4 fit well; they give similar conclusions but different representations of event dynamics (heterogeneity vs. event-dependence) and of treatment effect

5. Some Complications for Analysis

- Terminating events
- Event-dependent loss to followup
 - See Cook et al. (JASA, 2009, pp. 60-75), bone mets trial
 - Cerebrospinal fluid shunt failures (Lawless et al., 2001)
- Multiple event types
- Observational data with varying time origins across individuals
- Retrospective data: selection effects, missing data