GET MORE INFORMATION FROM RECURRENT EVENTS DATA

Wayne B. Nelson, consultant

WNconsult@aol.com

ABSTRACT. Recurrent events data arise in biomedical, engineering, sociological, business, and many other applications. Despite over 25 years of literature articles and books on analysis of such data, some current medical studies of recurrent diseases still use only each patient's first recurrence time and ignore subsequent ones, thus losing information. This article provides an introduction to modern analysis of recurrent events data and gives new results on how much added information and accuracy are gained by the use of all recurrences.

1. INTRODUCTION

Motivation. I suffered three episodes of blood clots (thromboses) while taking Coumadin, a standard treatment. My last episode included dangerous embolisms in my lungs and required six days in the hospital. Recently my doctor suggested that I take Xarelto, which was recently approved to prevent thromboses. Concerned, I searched the Web and found detailed information on Xarelto, including clinical studies of its effectiveness compared to Coumadin. Although blood clots are recurrent events, the data analyses of those studies used only patients' first occurrence time in the study and ignored later recurrences, obviously a loss of information.

Background. The practice of using only a patient's first occurrence time started before recurrent-events methods were widely known and uses well-known survival data methods. Considerable journal literature on recurrent events methods goes back at least 25 years, for example, Nelson (1988). Also, books for practitioners have been available for years, for example, Nelson (2003) and Cook and Lawless (2007), which also survey the literature. Nevertheless, the wasteful practice of using only the first occurrence time persists, perhaps due to analysts' ignorance of this specialized topic. To increase awareness, this article also presents basic models and analyses for recurrence data in Sections 2 and 3, where readers need only basic statistics.

Overview. Section 2 presents the basic nonparametric and Poisson models for recurrent events. Section 3 shows how much information is lost when only first occurrence times are used. Section 4 gives derivations of the results of Section 3. Those who are familiar with basic recurrent events methodology may wish to skip to Section 3.

2. MODELS

Overview. This section presents the basic nonparametric model for recurrent events, key concepts, and the Poisson model and its properties.

2.1 Nonparametric Model

Purpose. This section presents basic concepts for recurrent events data: cumulative history functions, the nonparametric population model, population distributions, the population mean cumulative function (MCF), and recurrence rate.

Data. Typical recurrence data are the patient's exact time under treatment at each recurrence and how long the patient was observed, the censoring time. Such data are called exact with right censoring. For example, Figure 1.1 is a time-line display of Byar

bladder tumor data (Nelson, 2003, p. 8) for 48 Placebo patients. Here the length of a line shows a patient's months under study until censored, and each x shows a recurrence time. More generally, recurrence time data may be interval data and may have right, left, or gap censoring, as described by Nelson (2003, Ch. 1) and Cook and Lawless (2007).

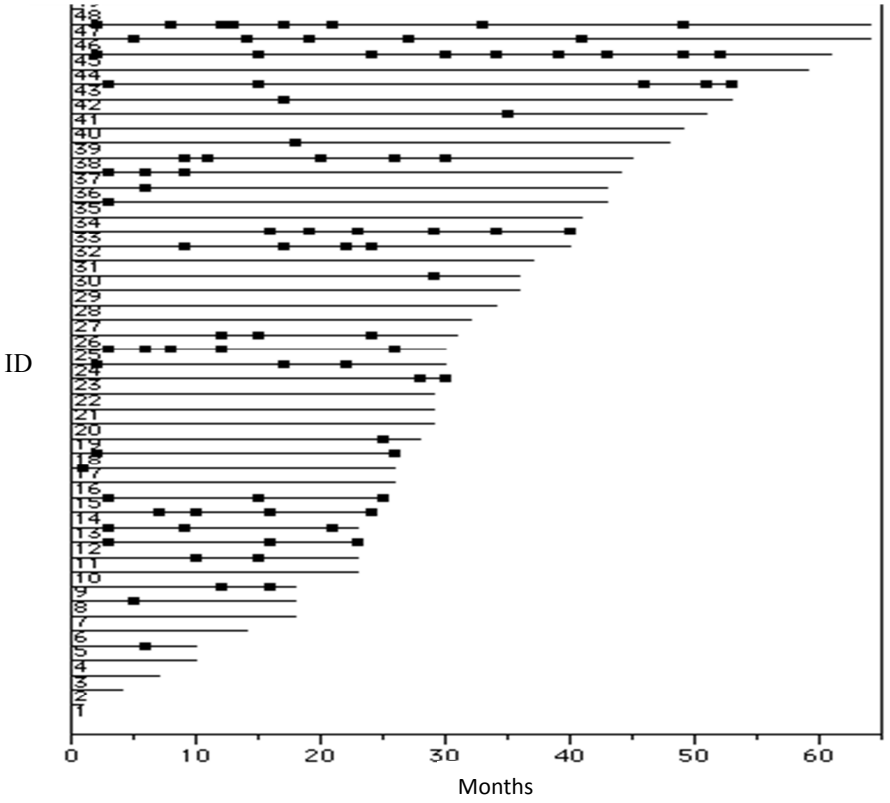

Figure 1.1 Display of tumor recurrences x for Placebo patients.

Cumulative history functions. Figure 1.2 shows a typical sample cumulative history function of a population unit; it is the time-line data (on the horizontal axis) depicted as a step function for the *cumulative number* of recurrences over time. In most parametric models, like the Poisson process, the increments/steps are all 1, as simultaneous events are impossible in such models. More generally for counts of events, increments can be any integer 1, 2, 3, ⋯. For example, births data could contain twins, triplets, etc. The general model here also allows for negative increments. For example, the number of people or production units in a waiting line can increase and decrease

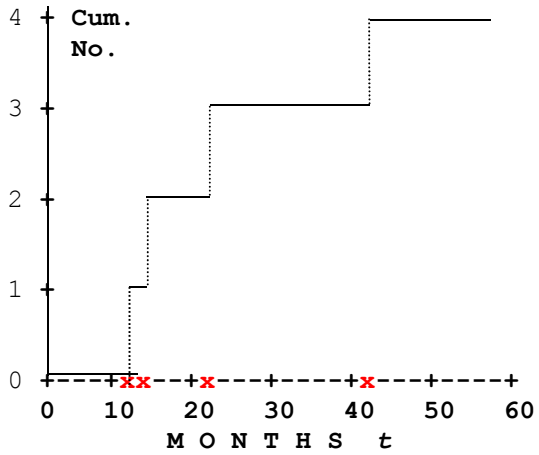

Figure 1.2 Cumulative history function for the *number* of recurrences.

Model. The *nonparametric population model* is the population of the uncensored cumulative history functions of a target population of patients. That is, the model is a population of step functions. Such a population of cumulative functions is depicted in Figure 1.3 as continuous curves, because a staircase functions would coincide and be unreadable.

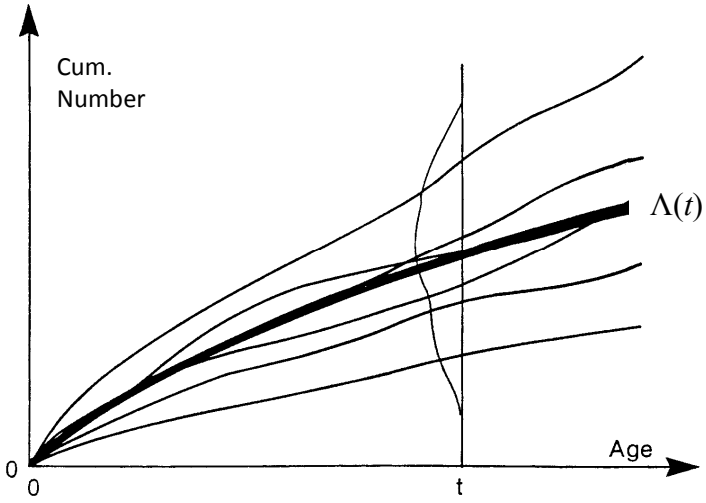

Figure 1.3 Population of cumulative histories for the number of recurrences.

MCF. At any age t, there is a population distribution of the cumulative number events as shown in Figure 1.3. The mean $\Lambda(t)$ of this distribution is a function of t. $\Lambda(t)$ is called the *mean cumulative function* (MCF) for the *number* of events. $\Lambda(t)$ is also called the cumulative intensity function. This function can be regarded as the population "mean curve," as it is the pointwise average of all population cumulative values at age t. For a finite population, $\Lambda(t)$ is a step function with many small steps, one for each event in the population. For many applications, the mean curve is regarded as continuous. $\Lambda(t)$ provides most of the information sought from recurrence data.

Applications. Figure 1.4 shows the MCF estimate (middle line) for recurrences of bladder tumors under Placebo treatment from the Byar data. Figure 1.5 from Prof. Richard Cook shows MCF estimates for recurrences of herpes episodes under two therapies with Valtrex: 1) taken only during episodes and 2) taken daily to suppress episodes; Cook and Lawless (2007, Sec. 5.3.4) provide more detail. Both figures use all recurrences. Theory for the MCF estimates requires that sample units are obtained with true random sampling of the population. Nelson (1988,2003) and Cook and Lawless (2007) show how to calculate a sample estimate $\Lambda^*(t)$ of the population $\Lambda(t)$ and plot it as in Figures 1.4 and 1.5. These sample MCFs are near linear. In most engineering and other applications, an MCF estimate is usually curved. The values of upper and lower curves at time t are approximate 95% confidence limits for $\Lambda(t)$. These limits are given by Nelson (1995,2003) and Cook and Lawless (2007). Commercial software that provide such plots include SAS Proc Reliability, JMP, Minitab, ReliaSoft RDA, and SuperSMITH Visual.

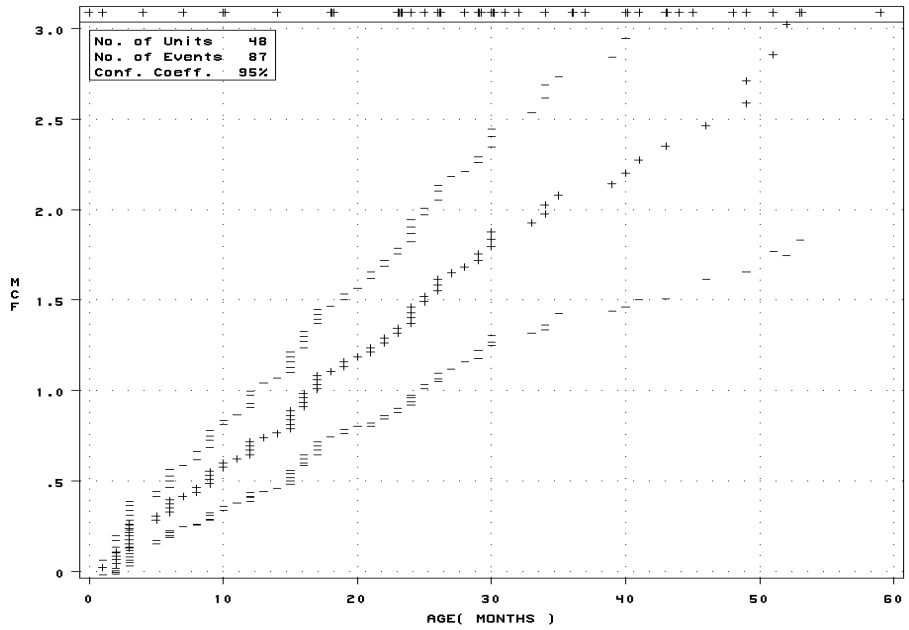

Figure 1.4 Sample MCF of tumor recurrences with Placebo.

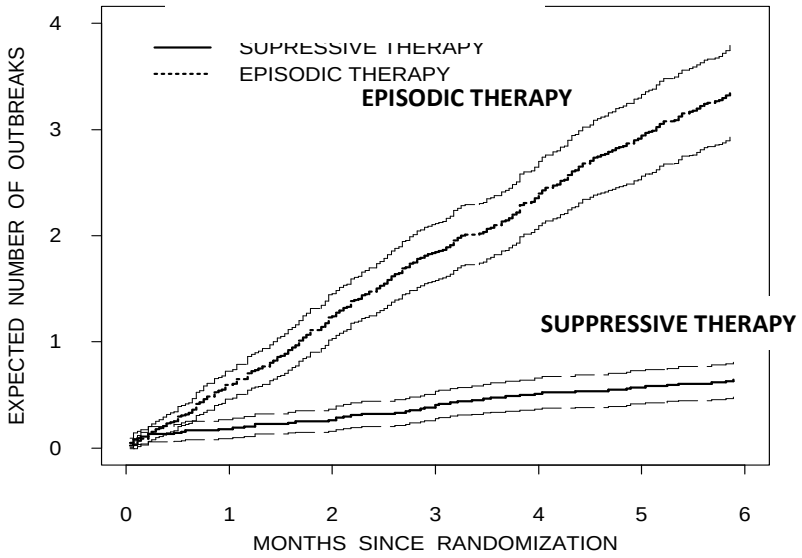

Figure 1.5 Sample MCFs of herpes episodes under two therapies.

Recurrence rate. Usually one assumes that $\Lambda(t)$ is continuous and has a derivative

$$\lambda(t) \equiv d\Lambda(t)/dt \,, \tag{1.1}$$

which is called the population instantaneous *recurrence rate* or *intensity function*. It is called "instantaneous" because it depends on the age t. The word "instantaneous" is often omitted but implied. $\lambda(t)$ is the mean number of recurrences per month per population unit, for example, the mean number of herpes recurrences per month per patient. In medicine, reliability, and other applications, one usually wants to know if the recurrence rate increases or decreases over time. Figure 1.3 shows an MCF with a decreasing recurrence rate (derivative) over time. Figures 1.4 and 1.5 have a constant recurrence rate, which is common in medical applications. Then patients do not get better or worse under a treatment, which just reduces the recurrence rate.

Censoring. For the population model, a cumulative history function of a patient is regarded as potentially observable over any time range of interest. In contrast, history functions of sample patients are during the study, and later yet unobserved recurrences are said to be censored. Censoring is a property of the data and its collection, not of the population model. Throughout it is assumed that the censoring is random (noninformative), that is, statistically independent of subsequent recurrences. In some applications, the histories of some units terminate before their planned censoring times. For example, some patients may die or vanish before a study ends. Such histories may not be randomly censored. Then the model and theory here must be modified to handle such termination. For example, Wang et al. (2001) deal with such informative censoring.

Cost/value histories. In this nonparametric model, the increment (step) of a cumulative history function can be a cost or "value" of each observed event. Any measure of "value" of an event may be used, for example, cost of treatment or days in hospital. Also, more generally, history functions may have negative increments, for example,

withdrawals from a bank account or customer returns of purchases. Nelson (1988, 2003, Ch. 2) provides a population model and data analyses for the cumulative *cost* or *value* of recurrences. An MCF for the cost of recurrences is denoted by $M(t)$ to distinguish it from an MCF $\Lambda(t)$ for the number of recurrences.

Types of events. Such history functions may include all types of events, selected types of events, or a single type of event. The choice of what to include depends on the application and the desired information. In practice, one may analyze a data set any number of times using different choices of the types of events and their values. For example, in a Xarelto study, minor and major bleeding are types of events. Nelson (2003, Ch. 6) gives models, data analyses, and applications for such data.

Not assumed. Note that this nonparametric model does not use a parametric stochastic process to generate cumulative history functions and entails no assumptions about the history functions; anything can physically produce them. Nelson (2003) gives the few assumptions for this nonparametric model and describes added assumptions of other models.

2.2 Poisson Model

Overview. The well-known Poisson process is the simplest parametric model for the cumulative number Y of recurrences in an observed time t. It can be a model for a single unit or a population of units all with the same constant recurrence rate $\lambda > 0$. defined below. Reviewed below are a definition of the Poisson process, its density, mean, and interarrival times. For more details, see Nelson (2003, Chap. 8).

Definition. The Poisson model is usually defined by the following properties:

1> The potential random number Y of recurrences in any time interval $(0,t)$ is unlimited.

2> The chance of a recurrence is the same for each point in time with a constant recurrence rate λ defined below. In practice, few populations have a constant recurrence rate. Then the widely-used Poisson model may serve as a first approximation.

3> The numbers of recurrences in any number of nonoverlapping intervals are all statistically independent (called the "independent increments" assumption). This is not so if population units have different recurrence rates λ_i. In practice, few populations have independent increments. In medical studies, patient λ_i usually differ.

4> All population units have the same constant recurrence rate λ.

Then the random number Y of recurrences in any period $(0,t)$ has a Poisson distribution (given below) with mean $\Lambda(t) = \lambda t$. The main goal of Poisson analysis is to accurately estimate λ for a single treatment or for treatments being compared. Also, this simple model is used to build more realistic models.

Density and mean. The Poisson density function $f(y)$ for the probability of y recurrences between time 0 and time t is

$$f(y) \equiv \Pr\{Y=y\} = (1/y!)(\lambda t)^y \exp(-\lambda t), \quad y = 0, 1, 2, \cdots. \tag{2.1}$$

In medicine, the dimensions of $\lambda > 0$ are recurrences per unit "time" per patient. Some authors use a single parameter $\mu = \lambda t$, which is the expected (mean) number of recurrences per patient up to time t. The mean of Y is $EY = \lambda t$, and the variance is $V(Y) = \lambda t$.

MCF. The Poisson mean cumulative function (MCF) for the number of recurrences up to time is $EY = \Lambda(t) = \lambda t$. Figure 2.1 depicts the Poisson process and shows that $\Lambda(t)$ is a straight line through the origin. It also shows the Poisson density (histogram) of the cumulative number of population recurrences at two ages. In biomedical work, $\Lambda(t)$ is often called the cumulative intensity function. These near linear MCF estimates in Figures 1.4 and 1.5 suggest that the Poisson process is a plausible first-approximation; there the slope of an MCF plot is the recurrence rate.

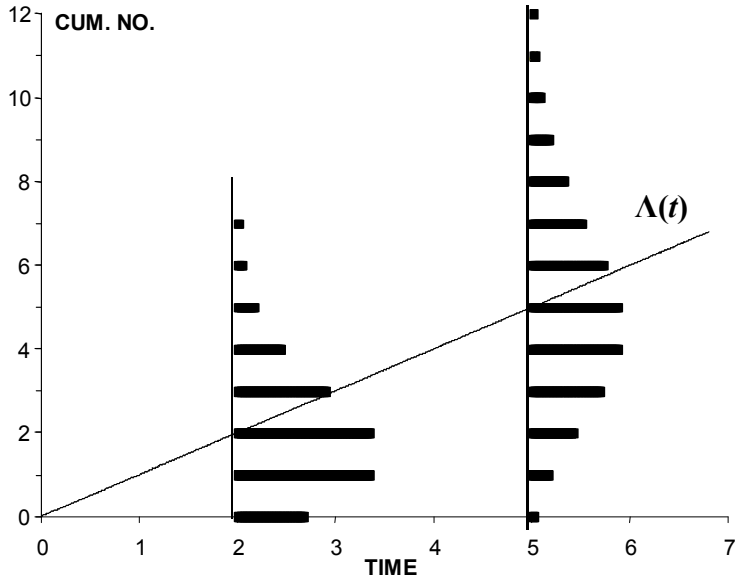

Figure 2.1 Poisson process depicted.

Times between recurrences. A patient′s times, $D_1, D_2, D_3, \cdots$ between successive recurrences are called *interarrival times*. For a Poisson process with recurrence rate λ, these D_j are statistically independent and from an exponential distribution with hazard rate λ. That is,

$$\Pr\{D_j \le t\} = 1 - \exp(-\lambda t), \; j = 1, 2, 3, \cdots. \qquad (2.2)$$

Sums of counts. The sum of statistically independent Poisson counts has a Poisson distribution. This fact is used to pool data from n patients as follows. Suppose that the i-th Poisson count Y_i has recurrence rate λ_i and observation time τ_i, $i = 1, 2, \cdots, n$. Then the sum $Y = Y_1 + Y_2 + \cdots + Y_n$ has a Poisson distribution with mean

$$\mu = \lambda_1 \tau_1 + \lambda_2 \tau_2 + \cdots + \lambda_n \tau_n \, . \qquad (2.3)$$

When $\lambda_1 = \lambda_2 = \cdots = \lambda_n = \lambda$, a common rate, Y has a Poisson distribution with mean $\mu = \lambda\tau_0$, where $\tau_0 = \tau_1 + \tau_2 + \cdots + \tau_n$ is the total observed time. This fact is used to combine recurrence data on patients, which are presumed to have the same λ here.

Other models. The simple Poisson model is adequate for showing that use of only first occurrences loses information. One could also show this with more complicated models, for example, regression models with covariates. Also, for example, the observed numbers Y_i could be Poisson where patients have differing λ_i, instead of a common λ. Cook and Lawless (2007, p. 79) present such a model where the λ_i are from a gamma distribution.

3. RESULTS

Overview. The goal of an analysis with the Poisson model is to obtain an accurate estimate of λ. When only each patient's first occurrence time is used, there is a loss of information and accuracy of the maximum likelihood (ML) estimator for λ. This section evaluates this loss. Two typical types of censoring are considered:

1) Equal. All n patients are studied for the same time τ, a common censoring time. Then the total time under study of all n patients is $n\tau$. Figure 1.3 depicts this situation if the cumulative history functions there are now regarded as sample data. Cook and Lawless (2007, Sec. 1.2) present such data on cancer in mice.

2) Staggered. The n patients enter the study at a constant rate, and there is a uniform distribution of length of time under study. This is called (linearly) staggered entry, Figure 1.1 depicts patients' censoring times τ_i that are near uniformly distributed from time 0 to $2\tau = 64$ months; there the average censoring age is $\bar{\tau} = \tau = 32$ months. Here, too, the total time under study of all n patients is $n\tau$ for comparison purposes.

The following results are derived in Section 4.

All recurrences. When patient i has Y_i Poisson recurrences in observed (censoring) time τ_i, $i = 1, 2, \cdots, n$, then the ML estimate of λ using All recurrences is

$$\hat{\lambda}_A = Y / \tau_0, \tag{3.1}$$

where $Y = Y_1+Y_2+\cdots+Y_n$ is the total number of recurrences and $\tau_0 = \tau_1+\tau_2+\cdots+\tau_n$ is the total observed time. The Fisher information for $\hat{\lambda}_A$ is

$$I(\hat{\lambda}_A) = \tau_0 / \lambda \, . \tag{3.2}$$

In a sense, the information is proportional to the sample size, since the asymptotic (large Y) variance of $\hat{\lambda}_A$ is

$$V(\hat{\lambda}_A) = 1 / I(\hat{\lambda}_A) = \lambda / \tau_0 \, . \tag{3.3}$$

These Poisson formulas apply to *any* censoring times $\tau_1, \tau_2, \cdots, \tau_n$, in particular, to equal and staggered censoring times.

Equal censoring times. When all $\tau_i = \tau$ and only each patients' first occurrence time is used, the data consist of the random number r of observed first times $t_1, t_2, \cdots, t_r$ and $(n-r)$ censoring times all equal to τ. Such data are a singly censored random sample from an exponential distribution with hazard rate λ. Nelson (2004, Sec. 8.1) gives the following results. The ML estimate of λ is

$$\hat{\lambda}_E = r / [t_1+t_2+\cdots+t_r + (n-r)\tau] \, . \tag{3.4}$$

The Fisher information for $\hat{\lambda}_E$ is

$$I(\hat{\lambda}_E) = (n/ \lambda^2) [1-\exp(-\lambda\tau)] \, . \tag{3.5}$$

The asymptotic (large r) variance of $\hat{\lambda}_{1E}$ is

$$V(\hat{\lambda}_E) = 1/I(\hat{\lambda}_E) = (\lambda^2/n) / [1-\exp(-\lambda\tau)] \, . \tag{3.6}$$

The fractional gain in information using all recurrences instead of just the first is

$$\text{GainE} = [I(\hat{\lambda}_A) - I(\hat{\lambda}_E)] / I(\hat{\lambda}_E) = [V(\hat{\lambda}_E)/V(\hat{\lambda}_A)] - 1 = \{\lambda\tau / [1-\exp(-\lambda\tau)] \} - 1. \tag{3.7}$$

GainE is also the effective gain in sample size when all recurrences are used. GainE (as a percentage) is plotted versus $\mu = \lambda\tau$ in Figure 3.1.

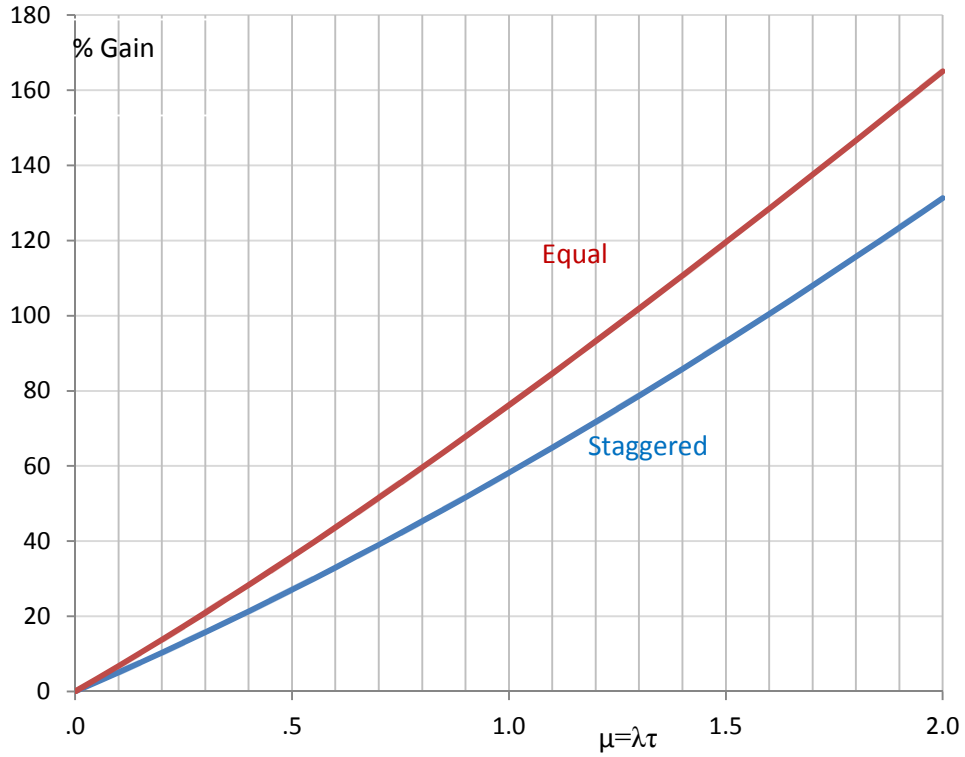

Figure 3.1 % Gain in information (effective sample size) using all recurrences.

Staggered censoring times. Such censoring time are uniformly distributed from 0 to 2τ , and the average censoring time is τ. When only each patient's first occurrence time is used, the data consist of the random number r of observed first times $t_1, t_2, \cdots, t_r$ and $(n-r)$ censoring times $\tau_1, \tau_2, \cdots, \tau_{n-r}$ of patients with no recurrence. Such data are a multiply censored random sample from an exponential distribution with hazard rate λ. The ML estimate of λ is

$$\hat{\lambda}_S = r \,/\, [t_1+t_2+\cdots+t_r + \tau_1+\tau_2+\cdots+\tau_{n-r}] \,. \tag{3.8}$$

The Fisher information for $\hat{\lambda}_S$ is

$$I(\hat{\lambda}_S) = (n\,/\lambda^2)\ \{1-(2\lambda\tau)^{-1}[1-\exp(-2\lambda\tau)]\}. \tag{3.9}$$

The asymptotic (large r) variance of $\hat{\lambda}_S$ is

$$V(\hat{\lambda}_S) \;=\; 1/\,I(\hat{\lambda}_S) \;=\; (\lambda^2/n)\,/\,\{1-(2\lambda\tau)^{-1}[1-\exp(-2\lambda\tau)]\}. \tag{3.10}$$

The fractional gain in information using all recurrences instead of just the first is

$$\begin{aligned} \text{GainS} &= [I(\hat{\lambda}_A)-I(\hat{\lambda}_S)\,]\,/\,(\hat{\lambda}_S) = [V(\hat{\lambda}_S)/V(\hat{\lambda}_A)] - 1 \\ &= \lambda\tau\{1-(2\lambda t)^{-1}[1-\exp(-2\lambda\tau)]\}^{-1} - 1. \end{aligned} \tag{3.11}$$

GainS is also the effective gain in sample size when all recurrences are used. GainS as a percentage is plotted versus $\mu = \lambda\tau$ in Figure 3.1.

Interpretation. Examination of Figure 3.1 shows:

- When the mean number of recurrences per patient is $\mu = \lambda\tau = 0.5$, the added information (sample size) is 27% for staggered censoring times and 36% for equal censoring times.
- When $\mu = \lambda\tau = 1.0$, the added information (sample size) is 58% for staggered censoring times and 76% for equal censoring times.

For such large percentages and the high cost of a study, it is clear that all recurrences should be used, not just the first ones. Of course, if $\mu = \lambda\tau$ is small, the additional information is small. For example, if $\mu = 0.1$, the added information is 5% for equal censoring and 7% for staggered censoring – both negligibly small.

Xarelto study. Xarelto studies are reported on Drugs.com (2013). Pooled data from the Einstein DVT and PE studies are

- $n = 4130$ patients ,
- $\bar{\tau} = 208$ days mean duration of treatment,
- $Y' = 40$ patients had major bleeding,
- $Y'' = 1169$ patients had minor bleeding.

Here $\tau_0 = 4130$x208 days. For major bleeding, the recurrence rate estimate is $\lambda' = Y'/\tau_0 = 40/(4130\text{x}208) = 0.00009685$ recurrences per patient per day (0.035 or 3.5% per year). Then $\lambda'\bar{\tau} = [40/(4130\text{x}208)]\ 208 = 40/4130 = 0.00969$, and the information gain (or sample size) using all recurrences is GainE = 0.49% for equal censoring and GainS = 0.65% for staggered censoring. These gains are negligible, because there are so few recurrences. For minor bleeding the recurrence rate estimate is $\lambda'' = Y''/\tau_0 = 1169/(4130\text{x}208) = 0.00136$. Then $\lambda''\bar{\tau} = 1169/4130 = 0.283$. The information gain (or sample size) using all recurrences is GainE = 14.8% for equal censoring and GainS = 19.7% for staggered censoring. These gains are not negligible. In the acyclovir study (Figure 2.3), censoring times were essentially equal, and all recurrences were used. If only the first occurrence were used for the suppressive therapy where $\lambda\tau \approx 0.6$, the gain using all recurrences would have been GainE = 33%.

4. DERIVATIONS

Overview. This advanced section provides derivations of Section 3 results. For each model and type of censoring, this includes (1) the sample (log) likelihood, (2) ML estimate of λ, (3) Fisher information, and (4) asymptotic variance. General ML theory and applications are presented, for example, by Nelson (1982, Ch. 8) and Lawless (2002).

All recurrences. For this case, patient i has Y_i recurrences in observed time τ_i, $i = 1, 2, \cdots, n$, where the Y_i are statistically independent Poisson random variables with a common recurrence rate λ. The sample likelihood is the product of the n Poisson probabilities of the Y_i ; namely,

$$L_A(\lambda) = (1/Y_1!)(\lambda\tau_1)^{Y1} \exp(-\lambda\tau_1)\ \ (1/Y_2!)(\lambda\tau_2)^{Y2} \exp(-\lambda\tau_2) \cdots (1/Y_n!)(\lambda\tau_n)^{Yn} \exp(-\lambda\tau_n). \quad (4.1)$$

The sample log likelihood is

$$£_A(\lambda) \ = \ \ln[L(\lambda)] \ = \ \Sigma_i \ln(Y_i!) + \Sigma_i\ Y_i \ln(\lambda\tau_i) - \lambda\ \Sigma_i\ \tau_i\ . \quad (4.2)$$

Sums Σ_i run from $i = 1$ to n. The ML estimate $\hat{\lambda}_A$ maximizes (4.1) or equivalently (4.2); namely,

$$\hat{\lambda}_A = Y / \tau_0 \tag{4.3}$$

where $Y = Y_1+Y_2+\cdots+Y_n$ and $\tau_0 = \tau_1+\tau_2+\cdots+\tau_n$. A ML estimate is usually obtained by setting the derivative of the log likelihood with respect to the parameter equal to zero and solving for the parameter estimate; that is, solve $\partial £_A(\lambda)/\partial\lambda = 0$. The Fisher information $I(\hat{\lambda}_A)$ is the expectation of the negative second partial derivative of $£_A(\lambda)$ with respect to λ; that is,

$$I(\hat{\lambda}_A) = E[-\partial^2 £_A/\partial\lambda^2] = E[\ \Sigma_i\ (Y_i / \lambda^2)] = \ \tau_0 / \lambda\ . \tag{4.4}$$

since $EY_i = \lambda\tau_i$. The asymptotic variance of $\hat{\lambda}_A$ is

$$V(\hat{\lambda}_A) = 1 / I(\hat{\lambda}_A)\ =\ \lambda / \tau_0\ . \tag{4.5}$$

These Poisson formulas apply to *any* censoring times $\tau_1, \tau_2, \cdots, \tau_n$, in particular, to equal and staggered censoring times. (4.3), (4.4), and (4.5) are (3.4), (3.5), and (3.6).

First occurrence only. When patient i is observed to time τ_i and only the first occurrence time T_i is used, the T_i come from an exponential distribution with probability density $f(T) = \lambda \exp(-\lambda T)$ and survival (censoring) probability $R(T) = \exp(-\lambda T)$. Then the data are a multiply censored sample from the exponential distribution. Nelson (2004, Sec. 8.1) derives the following results. For patient i, define the indicator variable

$d_i = 1$ if $T_i \le \tau_i$, a recurrence is observed,

$= 0$ if $T_i > \tau_i$, no recurrence is observed.

The sample likelihood is the probability of the observed sample data, namely,

$$L_1(\lambda) = \ \Pi_i\ \ [\lambda \exp(-\lambda T_i)]^{di}\ \ [\exp(-\lambda T_i)]^{1-\,di} \tag{4.6}$$

where the product Π_i runs over $i = 1$ to n. The log likelihood is

$$£_1(\lambda) = \ln[L_1(\lambda)] = \Sigma_i\ d_i\ [\ln(\lambda) - \lambda T_i] + \Sigma_i\ (1-d_i)\ (-\lambda T_i)\ . \tag{4.7}$$

where the sum Σ_i runs over $i = 1$ to n. The ML estimate $\hat{\lambda}_1$ maximizes (4.7) and is

$$\hat{\lambda}_1 = [\Sigma_i\ d_i\ ]\ /\ [\Sigma_i\ d_i\ Y_i\ +\ \Sigma_i\ (1-d_i)\ \tau_i\ ]\ =\ r\ /\ T_0\ ; \tag{4.8}$$

here r is the number of patients with a recurrence and T_0 is the sum of the r observed first occurrence times plus the sum of the $(n-r)$ censoring times of patients without a recurrence. The Fisher information is

$$I(\hat{\lambda}_1) = E[-\partial^2 £_1/\partial\lambda^2] = E[\Sigma_i\ (d_i / \lambda^2)\ ] = \ \Sigma_i\ [1-\exp(-\lambda\tau_i)] / \lambda^2\ . \tag{4.9}$$

where $Ed_i = 1-\exp(-\lambda\tau_i)$ is the probability of a first occurrence before time τ_i . The asymptotic (large r) variance of $\hat{\lambda}_1$ is

$$V(\hat{\lambda}_1) = 1 / I(\hat{\lambda}_1)\ =\ \lambda^2 / \Sigma_i\ [1-\exp(-\lambda\tau_i)]\ . \tag{4.10}$$

Results (4.6) through (4.10) apply to *any* censoring times $\tau_1, \tau_2, \cdots, \tau_n$, in particular, to equal and staggered censoring times as follows.

Equal censoring times τ. When all $\tau_i = \tau$, the previous ML results yield the following. Denote the r observed first occurrence times by $T_1, T_2, \cdots, T_r$. There are the $(n-r)$ censoring times τ of patients without recurrences. The ML estimate is

$$\hat{\lambda}_E = r / T_0 \tag{4.11}$$

where $T_0 = [T_1+T_2+\cdots+T_r + (n-r)\tau]$. The Fisher information is

$$I(\hat{\lambda}_E) = n\,[1-\exp(-\lambda\tau)] / \lambda^2 . \tag{4.12}$$

The asymptotic (large r) variance of $\hat{\lambda}_E$ is

$$V(\hat{\lambda}_E) = 1 / I(\hat{\lambda}_E) = (\lambda^2/n) / [1-\exp(-\lambda\tau)] . \tag{4.13}$$

These equations are (3.4), (3.5), and (3.6).

Staggered censoring times. From (4.8), the ML estimate for λ is

$$\hat{\lambda}_S = r / T_0 \tag{4.14}$$

where $T_0 = [Y_1+Y_2+\cdots+Y_r + (n-r)\,\tau]$. For linearly staggered censoring times, the τ_i are uniformly distributed from time 0 to time 2τ. Then the sum (4.9) can be approximated with uniform distribution of the τ_i with density $g(u) = 1/(2\tau)$, $0 \le u \le 2\tau$. Then the Fisher information of $\hat{\lambda}_S$ is obtained by replacing the sum in (4.9) with the integral

$$I(\hat{\lambda}_S) = n\int_0^{2\tau} (2\tau)^{-1}\{[1-\exp(-\lambda u)] / \lambda^2\}\,du = (n/\lambda^2)\,\{1+ (2\lambda\tau)^{-1}\,[\exp(-2\lambda\tau) -1]\} . \tag{4.15}$$

The asymptotic variance of $\hat{\lambda}_S$ is

$$V(\hat{\lambda}_S) = 1 / I(\hat{\lambda}_S) = (\lambda^2/n) / \{1+ (2\lambda\tau)^{-1}\,[\exp(-2\lambda\tau) -1]\}. \tag{4.16}$$

These equations are (3.8), (3.9), are (3.10).

Concluding remarks. This work has used a simple model, adequate for present purposes, to show how much information is lost when an analysis of recurrent events data uses only the first occurrence times. This work is intended to convince analysts to use the more informative methods for recurrence data in the references.

Acknowledgments. The author is much obliged to Prof. Jerry Lawless, Prof. Richard J. Cook, and Dr. Terry Therneau for their input on this article.

The author. Dr. Wayne Nelson is a leading expert on analysis of reliability and accelerated test data. He consults on applications, gives training courses for companies and professional societies, and works as an expert witness. For 24 years he consulted across the General Electric Co. and received the Dushman Award of GE Corp. R&D for developments and applications of product reliability data analysis. He was elected a Fellow of the Amer. Statistical Assoc. (1973), the Amer. Soc. for Quality (1983), the Institute of Electrical and Electronics Engineers (1988) for his innovative developments. He was awarded the 2003 Shewhart Medal and the 2010 Shainin Medal of ASQ and the 2005 Lifetime Achievement Award of IEEE for outstanding developments of reliability methodology and contributions to reliability education. He authored three highly regarded books *Applied Life Data Analysis* (Wiley 1982, 2004), *Accelerated Testing*

(Wiley 1990, 2004), *Recurrent Events Data Analysis* (SIAM 2003), two ASQ booklets, and 130 journal articles. He can be contacted at WNconsult@aol.com

REFERENCES

Cook, R.J. and Lawless, J.F. (2007), *The Statistical Analysis of Recurrent Events*, 420 pp., Springer, New York.

Drugs.com (2013), http://www.drugs.com/pro/xarelto.html .

Lawless, J.F. (2002), *Statistical Models and Methods for Lifetime Data*, 2nd ed., 630 pp., Wiley, New York.

Nelson, Wayne (1988), "Graphical Analysis of System Repair Data," *J. of Quality Technology* **20,** 24-35.

Nelson, Wayne (1995), "Confidence Limits for Recurrence Data -- Applied to Cost or Number of Product Repairs," *Technometrics* **37,** 147-157.

Nelson, Wayne B. (2003), *Recurrent Events Data Analysis for Product Repairs, Disease Recurrences, and Other Applications*, 151 pp., ASA-SIAM Series, SIAM, Philadelphia.

Nelson, W.B. (2004), *Applied Life Data Analysis*, paperback edition, 634 pp., Wiley, New York.

Wang, M.-C., Qin, J., and Chiang, C.T. (2001), "Analyzing Recurrent Events Data with Informative Censoring," *J. Amer. Statist. Assoc.* **96**, 1057-1065.