Visual Predictive Checks for the Evaluation of the Hazard Function in Time-to-Event Analyses

22nd PAGE Meeting, Glasgow, UK

11-14 June 2013

Matt Hutmacher
Ann Arbor Pharmacometrics Group (A2PG)
Ann Arbor, MI, USA

The Data

Based on a real case study:

From hypothetical placebo arm

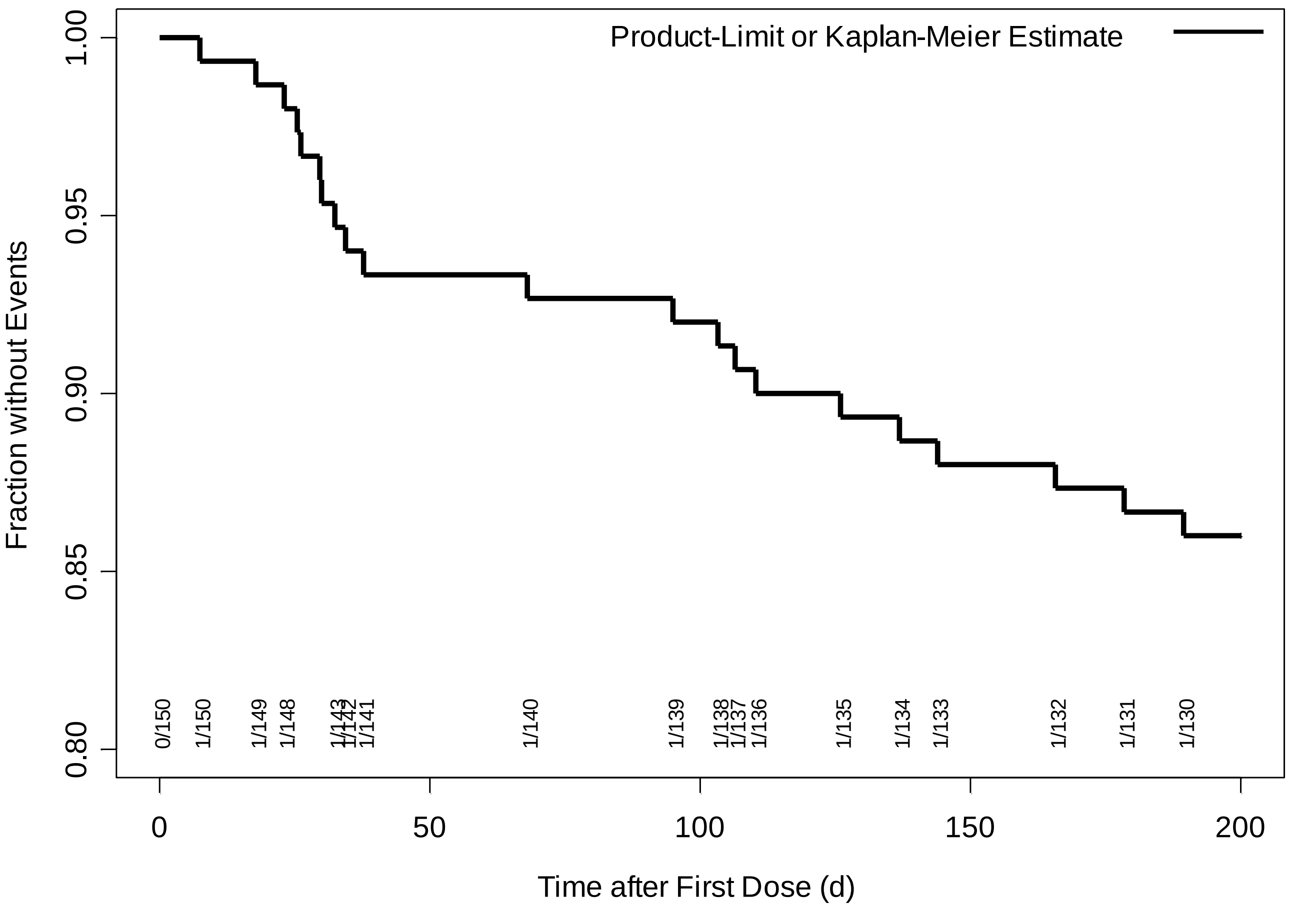
Trial duration = 200 days

Sample Size = 150

N=21 events were observed (only 1 event can occur)

Oddly – no censoring except due to study completion

The Data



Only 21 events → 14.0% failure rate at end of the study

Not likely much information to evaluate models

Or is there?

Fitting a Model

Important relationships:

$S(t) = Prob\{T > t\} = 1 - F(t)$ The survival function

$S(t) = exp[-H(t)]$ A function of the cumulative hazard, $H(t)$

$$H(t) = \int_0^t h(m)dt \rightarrow \boxed{h(t) = \frac{dH(t)}{dt}}$$

$$= \lim_{\Delta t \rightarrow 0} \frac{Prob\{t < T \leq t + \Delta t | T > t\}}{\Delta t}$$

The hazard is the derivative of $H(t)$

Fit Gompertz model to the data:

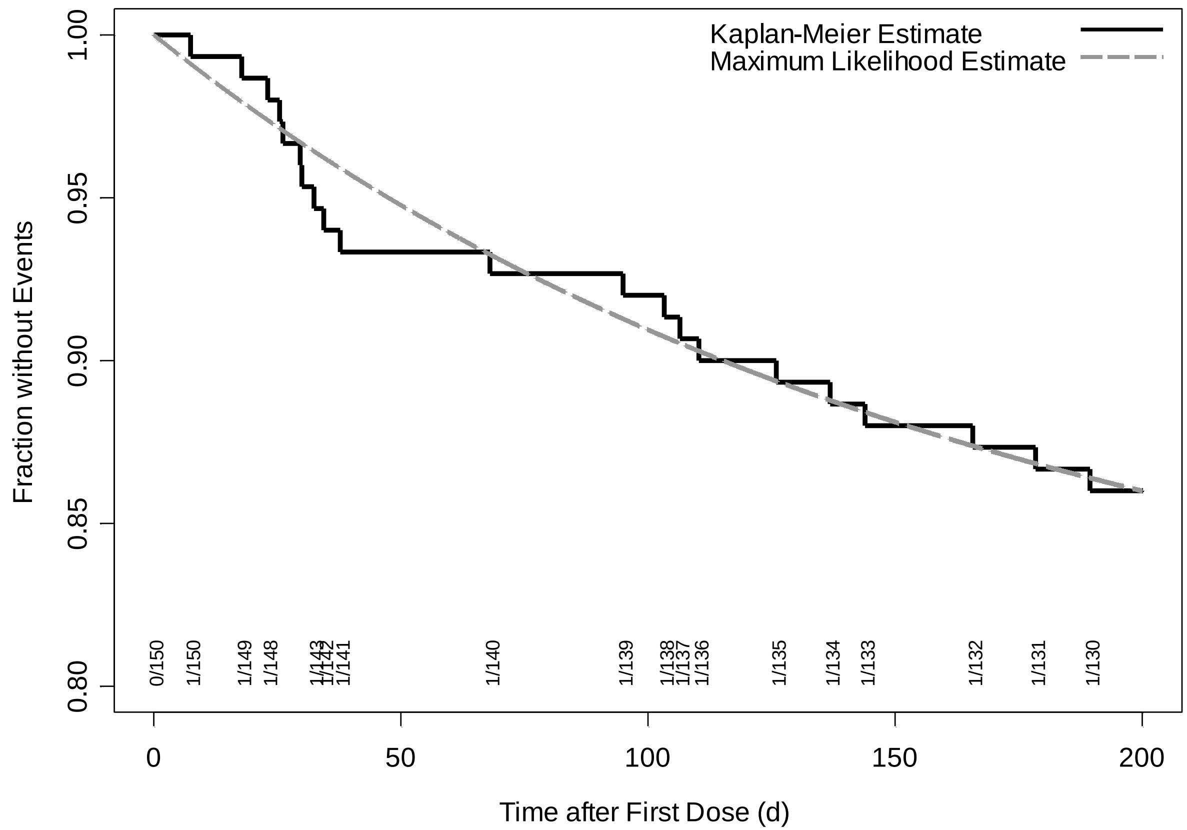
$$\log_e h(t) = \beta_0 + \beta_1 \cdot t$$

Fitting a Model

Fit using maximum likelihood

$$\hat{h}_{ML}(t) = e^{(\hat{\beta}_0 + \hat{\beta}_1 \cdot t)}$$

$$\hat{S}_{ML}(t) = e^{-\int_0^t \hat{h}_{ML}(m)dm}$$

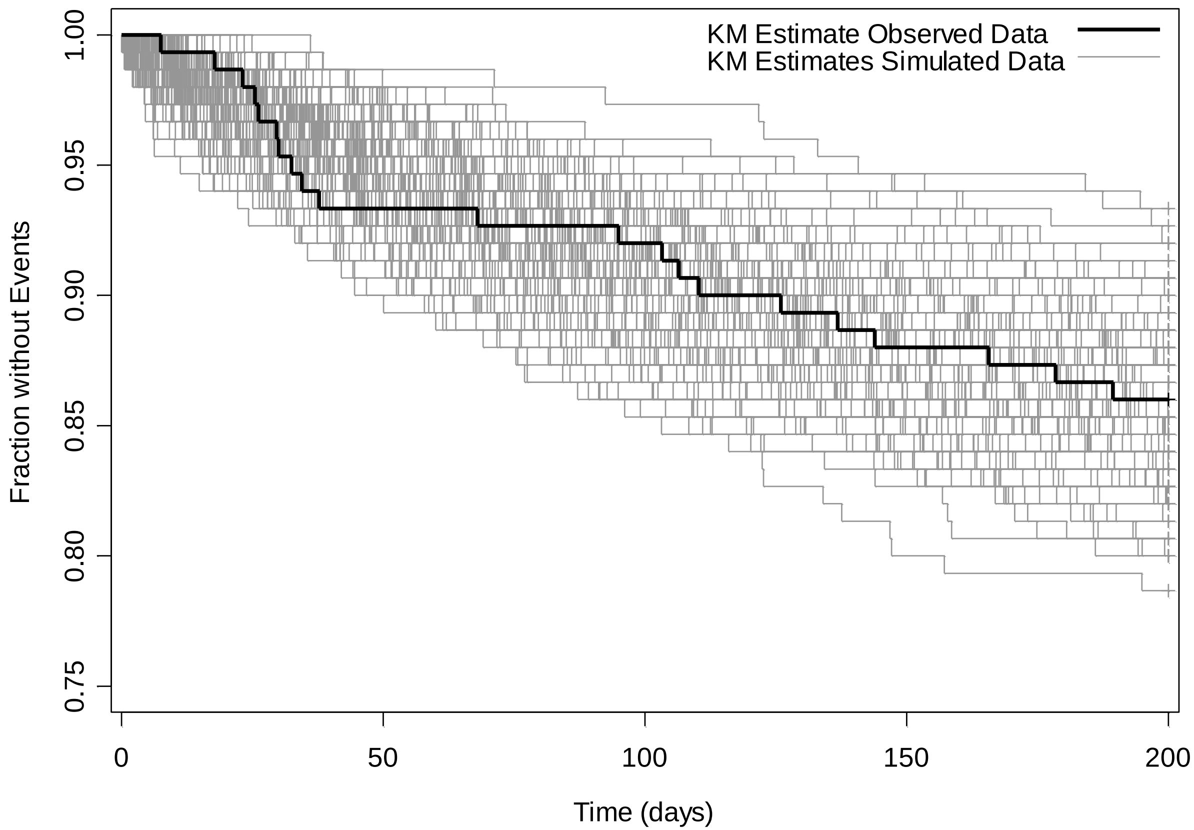


VPC using Kaplan-Meier*

Generate 200 replicates from the model

Calculate KM curves for each replicate

Overlay



Kaplan and Meier (1958)

VPC using Kaplan-Meier

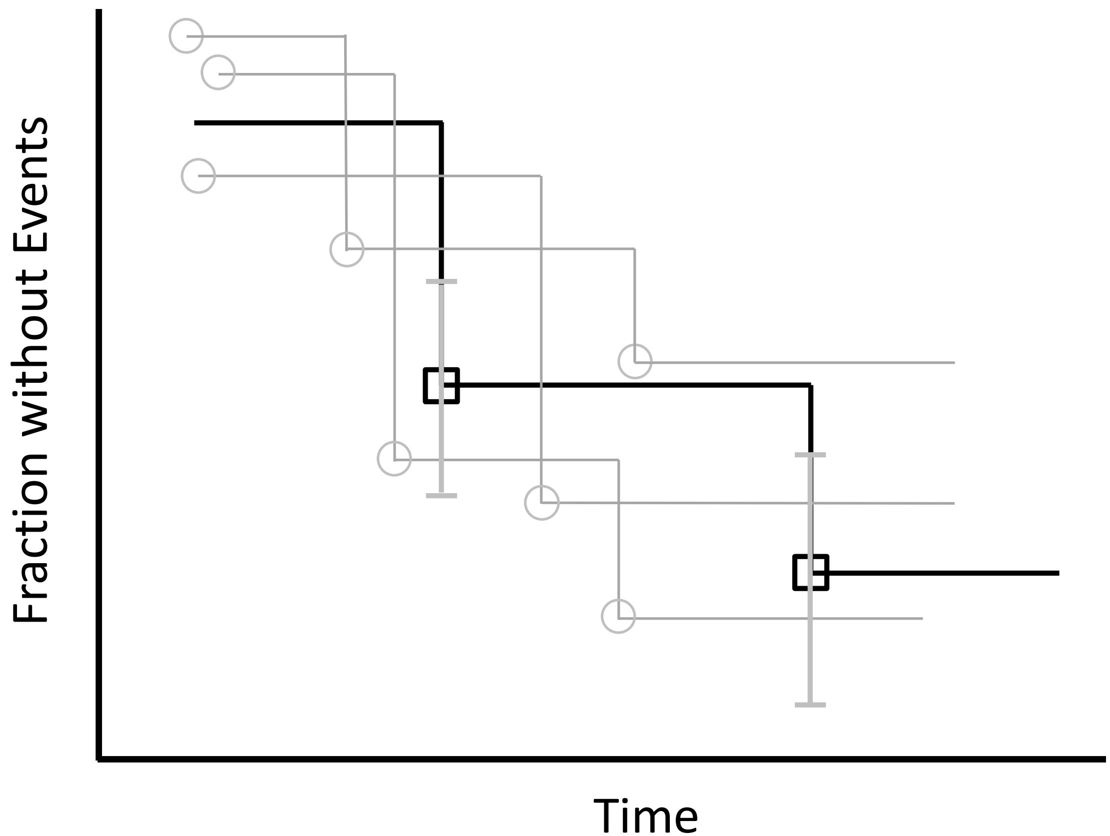
Let the event times be:
$t_1 \leq t_2 \cdots t_{n-1} \leq t_n$

Would like X% pointwise prediction intervals (PI) at each $S(t_i)$
{although any t works}

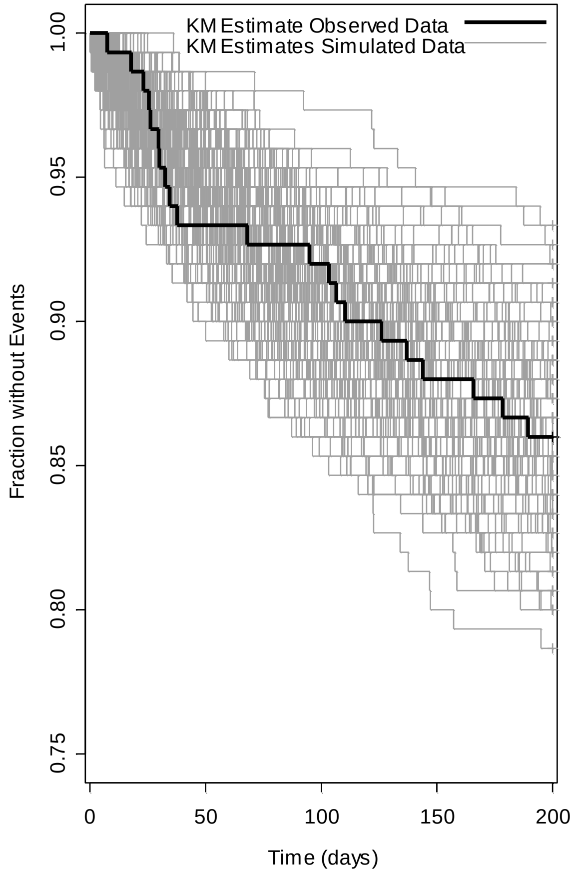
However:
$t_i \neq t_i^{(s)}$, where $t_i^{(s)}$ is the VPC simulated event time

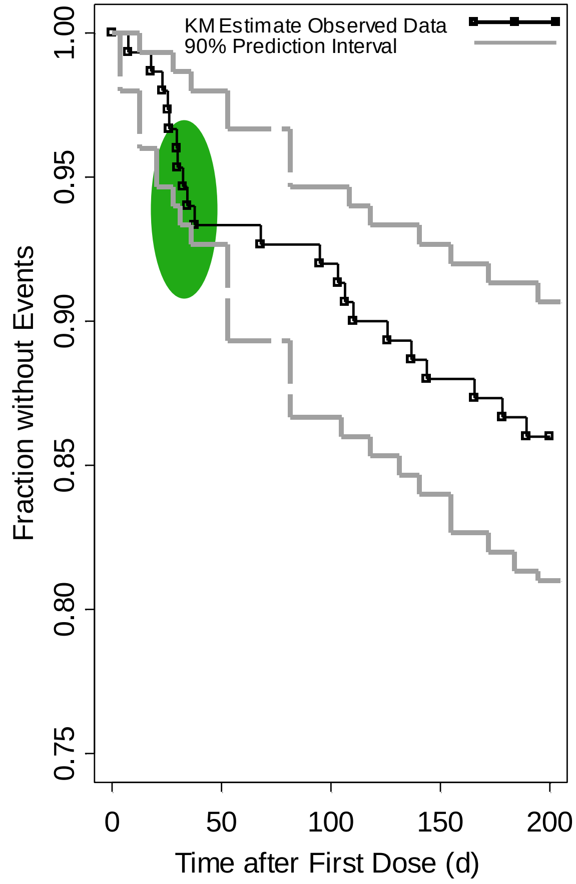
So, use KM interpolation

Compute, eg, 90% PI



VPC using Kaplan-Meier





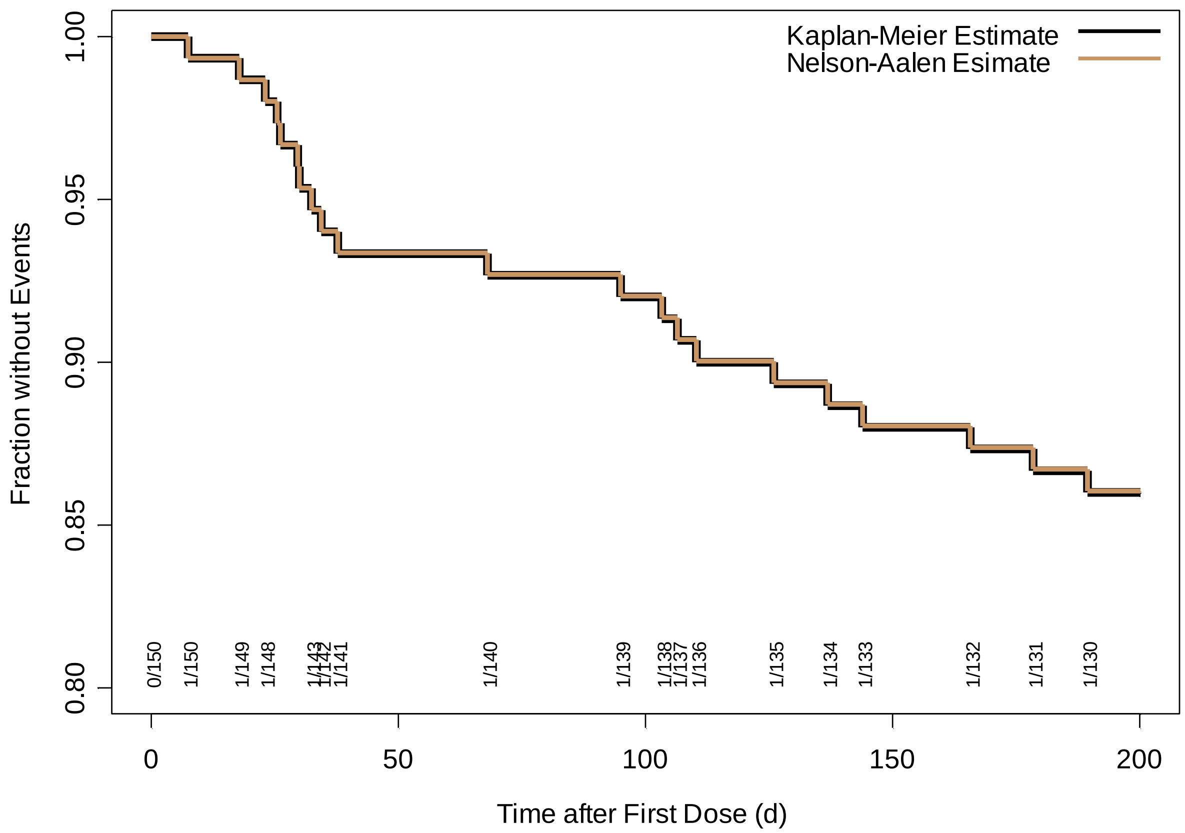
The Nelson-Aalen* Estimator

Let, at time t_i:
d_i = the number of events
n_i = the number at risk

$$\hat{H}_{NA}(t) = \sum_{i:t_i<t} \frac{d_i}{n_i}$$

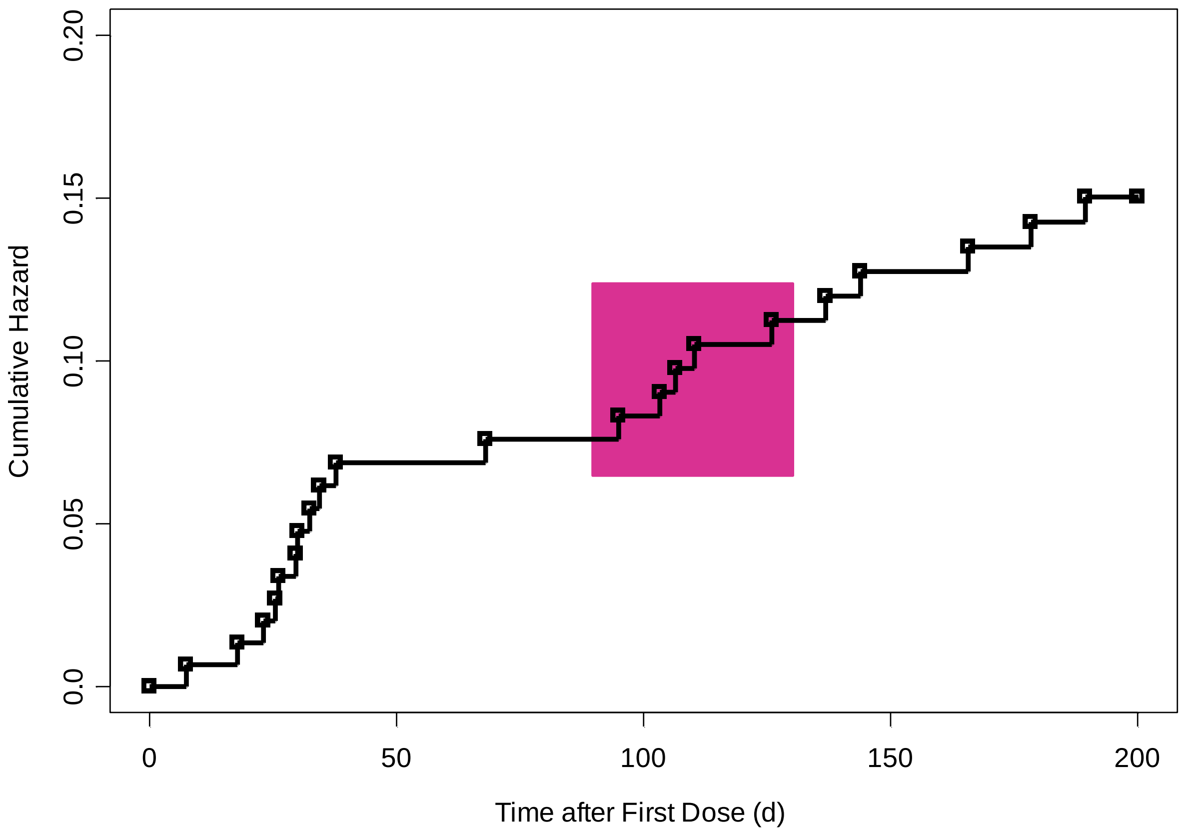
$$\hat{S}_{NA}(t) = exp\left(-\hat{H}_{NA}(t)\right)$$

$$\hat{S}_{KM}(t) = \prod_{i:t_i<t} \left(1 - \frac{d_i}{n_i}\right)$$



*Nelson (1972)
*Aalen (1978)

Cumulative Hazard



$$\widehat{H}_{NA}(t) = \sum_{i:t_i<t} \frac{d_i}{n_i}$$

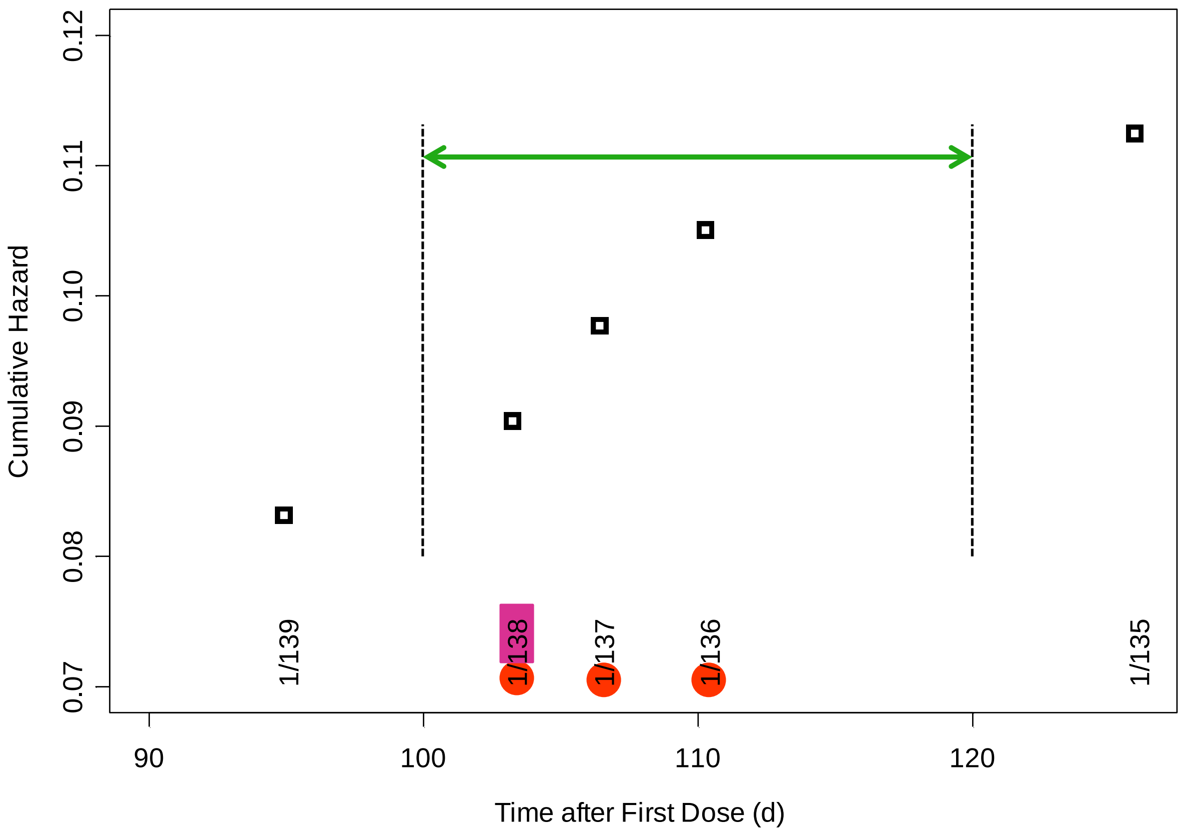
$$H(t) = \int_0^t h(m)dt$$

$$\rightarrow h(t) = \frac{dH(t)}{dt}$$

Binned Hazard Estimator (1)

$$\hat{h}_{BH1}(t=100) = \frac{\#\ events\ (100 \le t < 120)}{\#\ at\ Risk\ (t=100)} \times \frac{1}{\Delta t} = \frac{3}{138} \times \frac{1}{20} = 0.00109^{*}$$

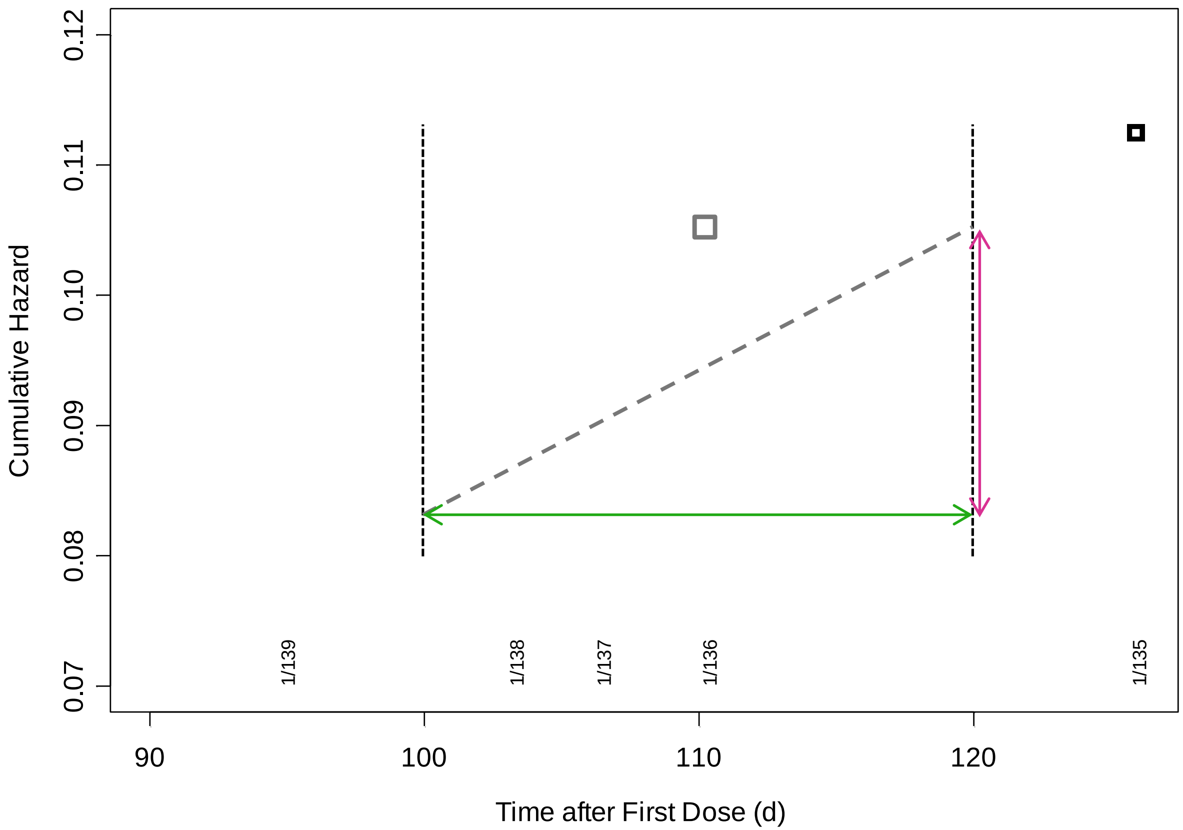
*Liu and Van Ryzin (1985)
A histogram estimator



Binned Hazard Estimator (2,3)

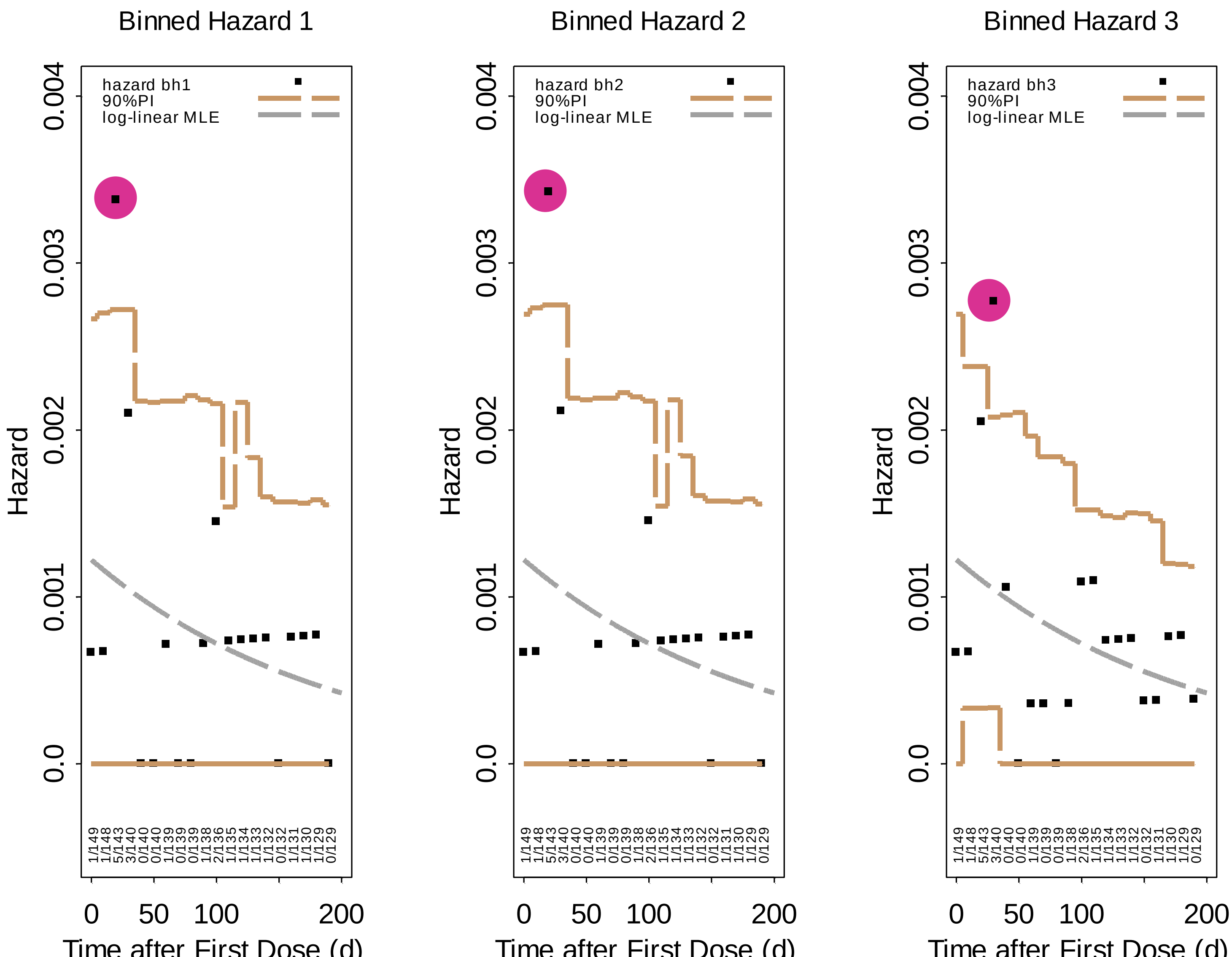
$$\hat{h}_{BH2}(t=100) = \frac{\hat{H}_{NA}(120) - \hat{H}_{NA}(100)}{\Delta t} = \frac{0.105 - 0.083}{20} = 0.00109\ *$$

$$\hat{h}_{BH3}(t=100) = \frac{\hat{H}_{NA}(120) - \hat{H}_{NA}(80)}{\Delta t} = \frac{0.105 - 0.076}{40} = 0.00073$$



*Equal to $\hat{h}_{BH2}$ when only 1 event and 0 censoring in an interval

VPC of the BH Estimators



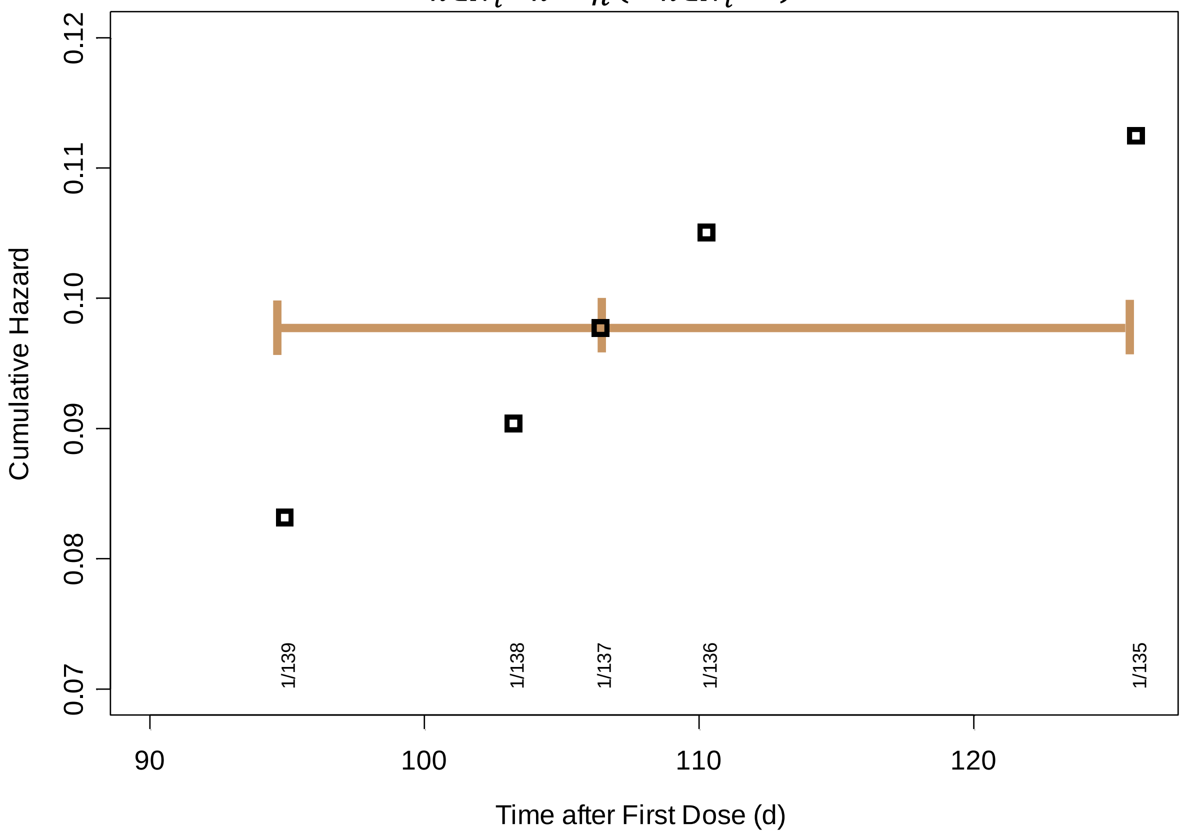
Running Lines Smoother

$$N_i = \left\{ max\left(i - \frac{[\phi n] - 1}{2}, 1\right) \dots i-1, i, i+1, \dots, min\left(i + \frac{[\phi n] - 1}{2}, 1\right) \right\}$$

ϕ = span

$[.]$ = integer function

$$\hat{h}_{RL}(t = 106) = \frac{\sum_{k \in N_i} \widehat{H}_{AN}(t_k) \cdot t_k - \frac{1}{n} \sum_{k \in N_i} \widehat{H}_{AN}(t_k) \cdot \sum_{k \in N_i} t_k}{\sum_{k \in N_i} t_k^2 - \frac{1}{n}\left(\sum_{k \in N_i} t_k\right)^2} = 0.00097$$



For example: selected ϕ such that $[\phi n] = 5$

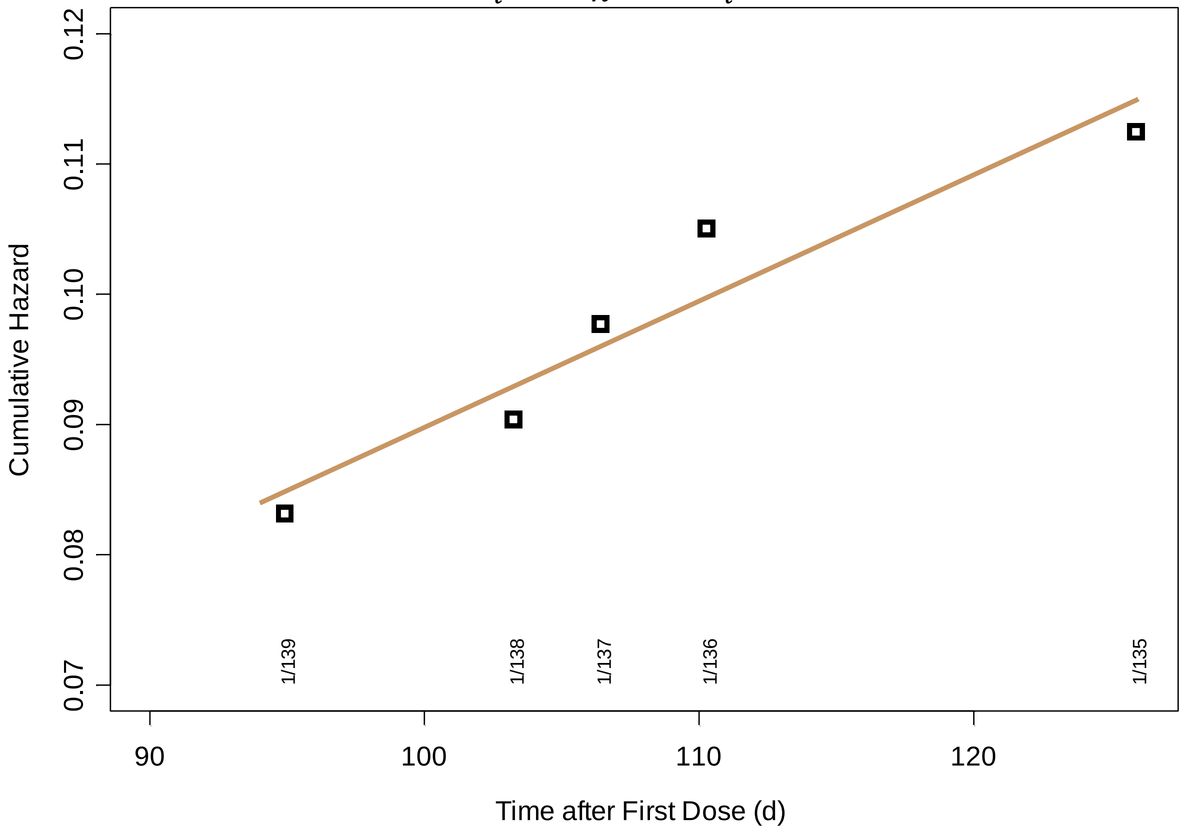
Running Lines Smoother

$$N_i = \left\{max\left(i - \frac{[\phi n]-1}{2}, 1\right) \ldots i-1, i, i+1, \ldots, min\left(i + \frac{[\phi n]-1}{2}, 1\right)\right\}$$

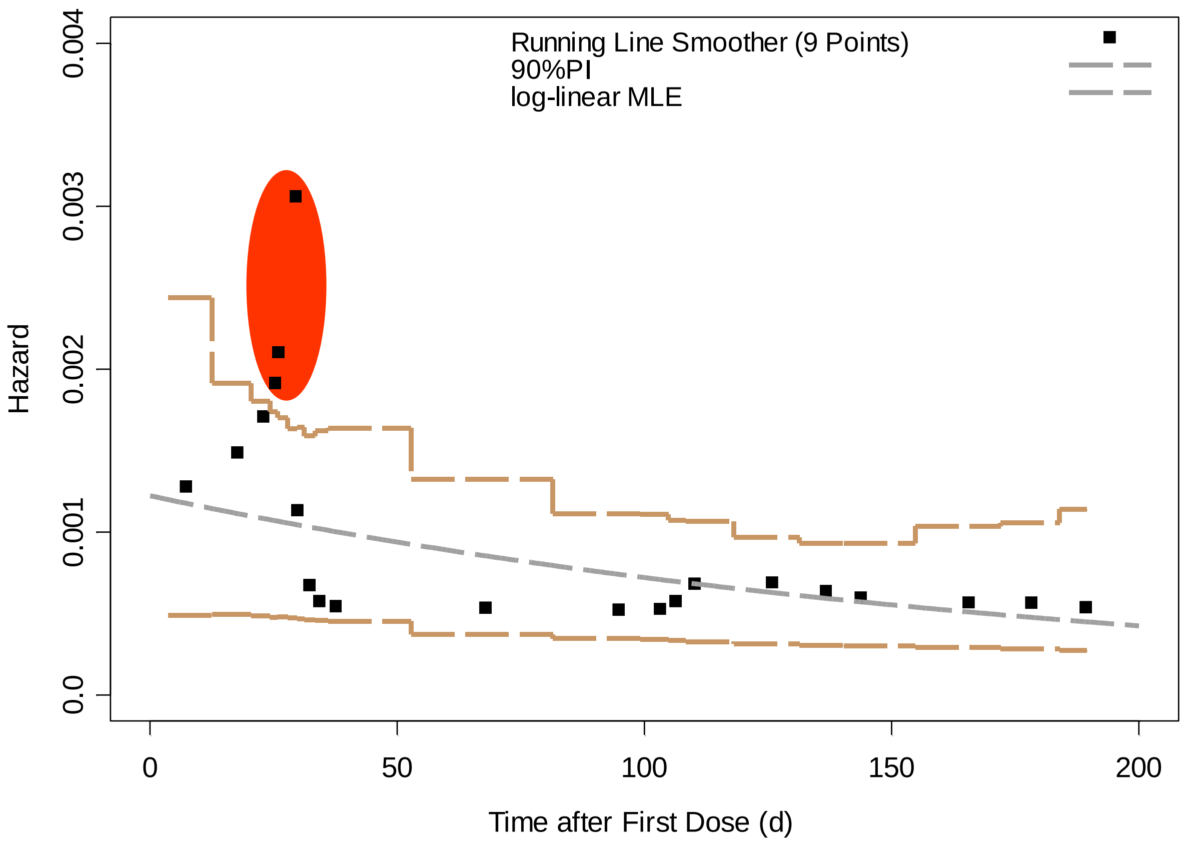
ϕ = span

$[.]$ = integer function

$$\hat{h}_{RL}(t=106) = \frac{\sum_{k\in N_i} \hat{H}_{NA}(t_k)\cdot t_k - \frac{1}{n}\sum_{k\in N_i} \hat{H}_{NA}(t_k)\cdot \sum_{k\in N_i} t_k}{\sum_{k\in N_i} t_k^2 - \frac{1}{n}\left(\sum_{k\in N_i} t_k\right)^2} = 0.00097$$



VPC using Running Line Smoother



Kernel Hazard Estimation

Can use kernel smoothing to ↓ variability of running line smoother

$$\hat{\hat{h}}_{RL}(t) = \frac{\sum_k K_b(t - t_k) \cdot \hat{h}_{RL}(t_k)}{\sum_k K_b(t - t_k)}, \qquad \tfrac{1}{b}K\left(\tfrac{t-t_k}{b}\right) \qquad \tfrac{1}{b}\int_{-\infty}^{\infty} K\left(\tfrac{t-t_k}{b}\right)dt = 1$$

After some research found discussed by Watson
and Leadbetter (1964)
(borrowing from the notation and exposition of Ramlau-Hansen [1983]):

$$\hat{h}_{SM}(t) = \int K_b(t - m)d\widehat{H}_{NA}(m) \qquad = \sum_{i=1}^{n} K_b(t - t_i)\frac{d_i}{n_i}$$

Kernel Hazard Estimation

Epanechnikov (1969):

$$K_b(t - t_i) = \frac{1}{b} \cdot \frac{3}{4}\left[1 - \left(\frac{t - t_i}{b}\right)^2\right] \cdot I_{\left(\left|\frac{t-t_i}{b}\right| \leq 1\right)} \qquad t \in (t_i - b, t_i + b)$$

Normal:

$$K_b(t - t_i) = \frac{1}{b\sqrt{2\pi}} \cdot exp\left[-\frac{1}{2}\left(\frac{t - t_i}{b}\right)^2\right] \qquad t \in (-\infty, \infty)$$

Lognormal:

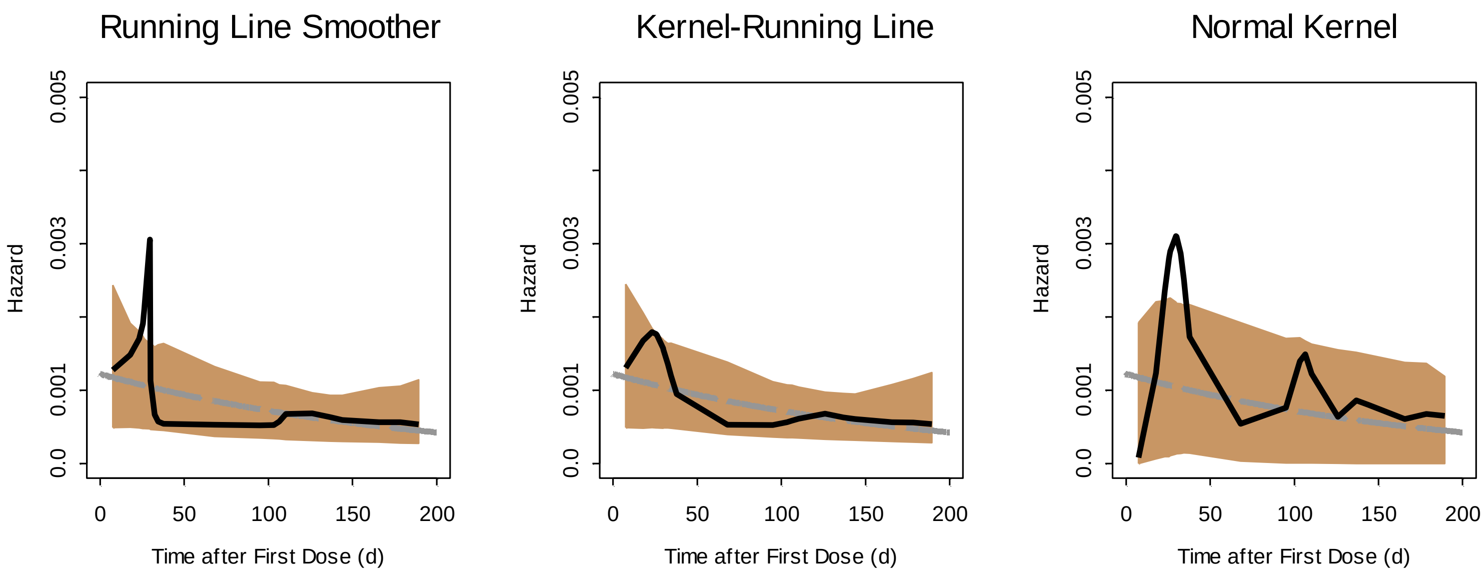
$$K_b(t - t_i) = \frac{1}{t \cdot b\sqrt{2\pi}} \cdot exp\left[-\frac{1}{2}\left(\frac{\log_e t - \log_e t_i}{b}\right)^2\right] \qquad t \in (0, \infty)$$

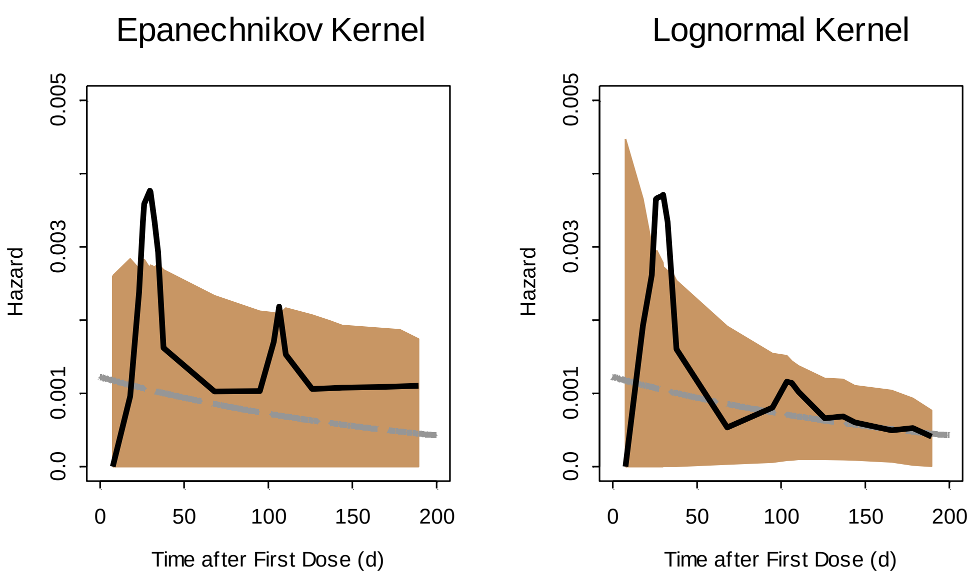
$b =$ bandwidth, a function of n – ie, $b = b_n$ such that $n \to \infty$, $b_n \to 0$.
Also, b_n can also be a function of time $b_n = b_n(t)$ to enhance estimation.

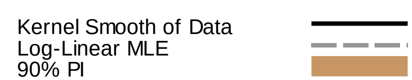
Bandwidth balances variance versus bias of estimate and can be estimated

Kernels can also be modified to minimize *edge effects* (Müller, 1991).

Kernel Hazard Estimation

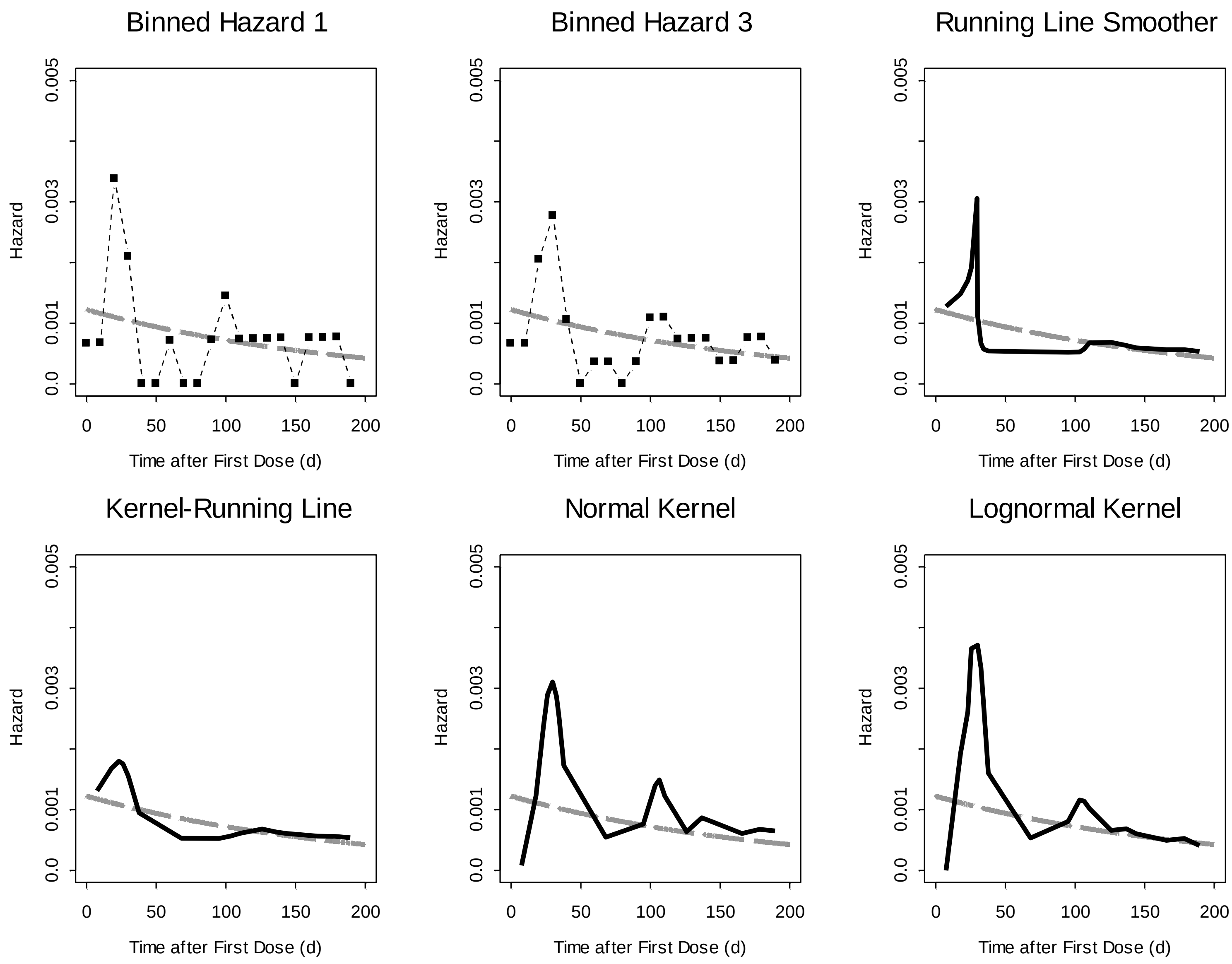




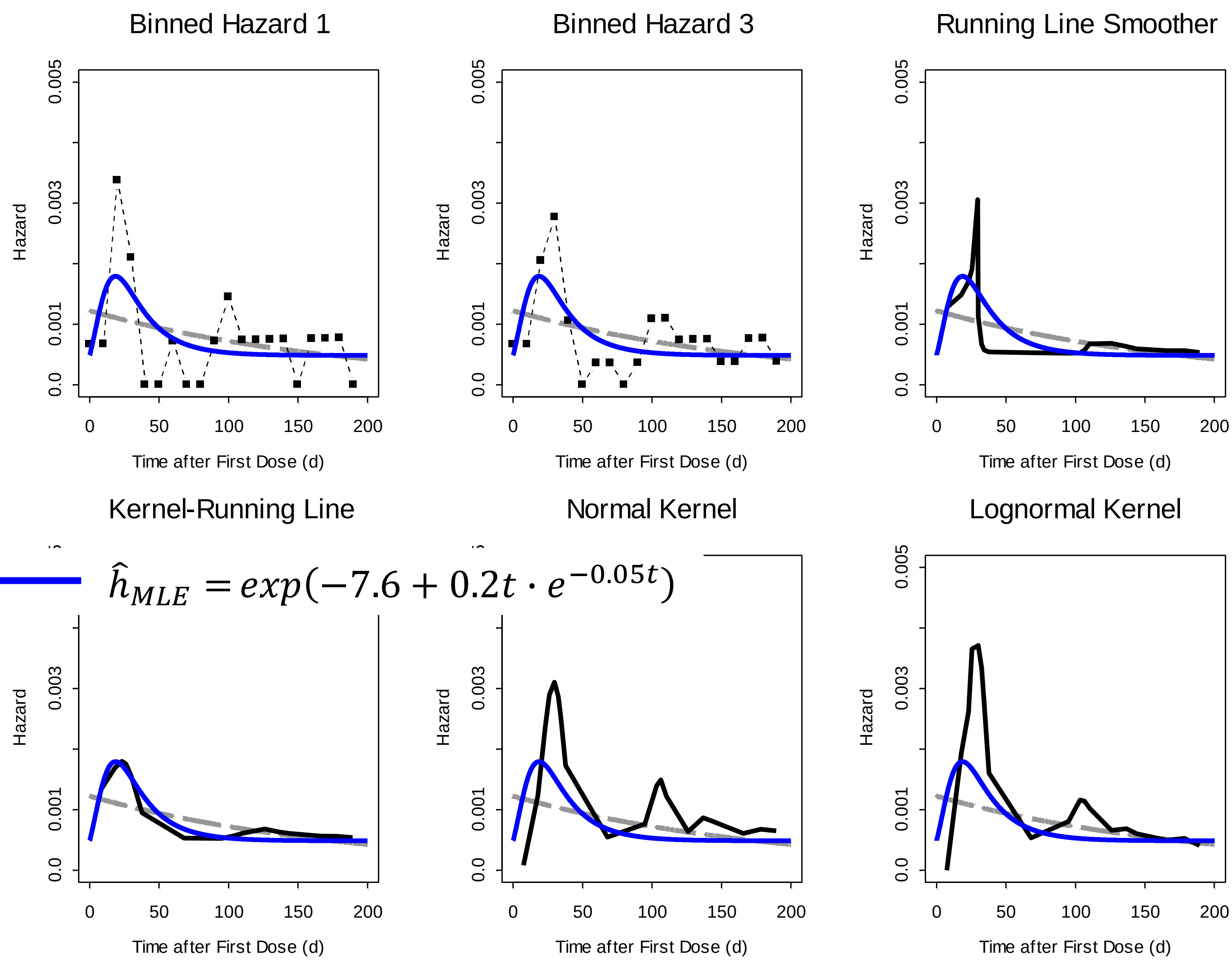


Predictions at observed event times – hazard smooth can be computed at any time

Overall Comparison

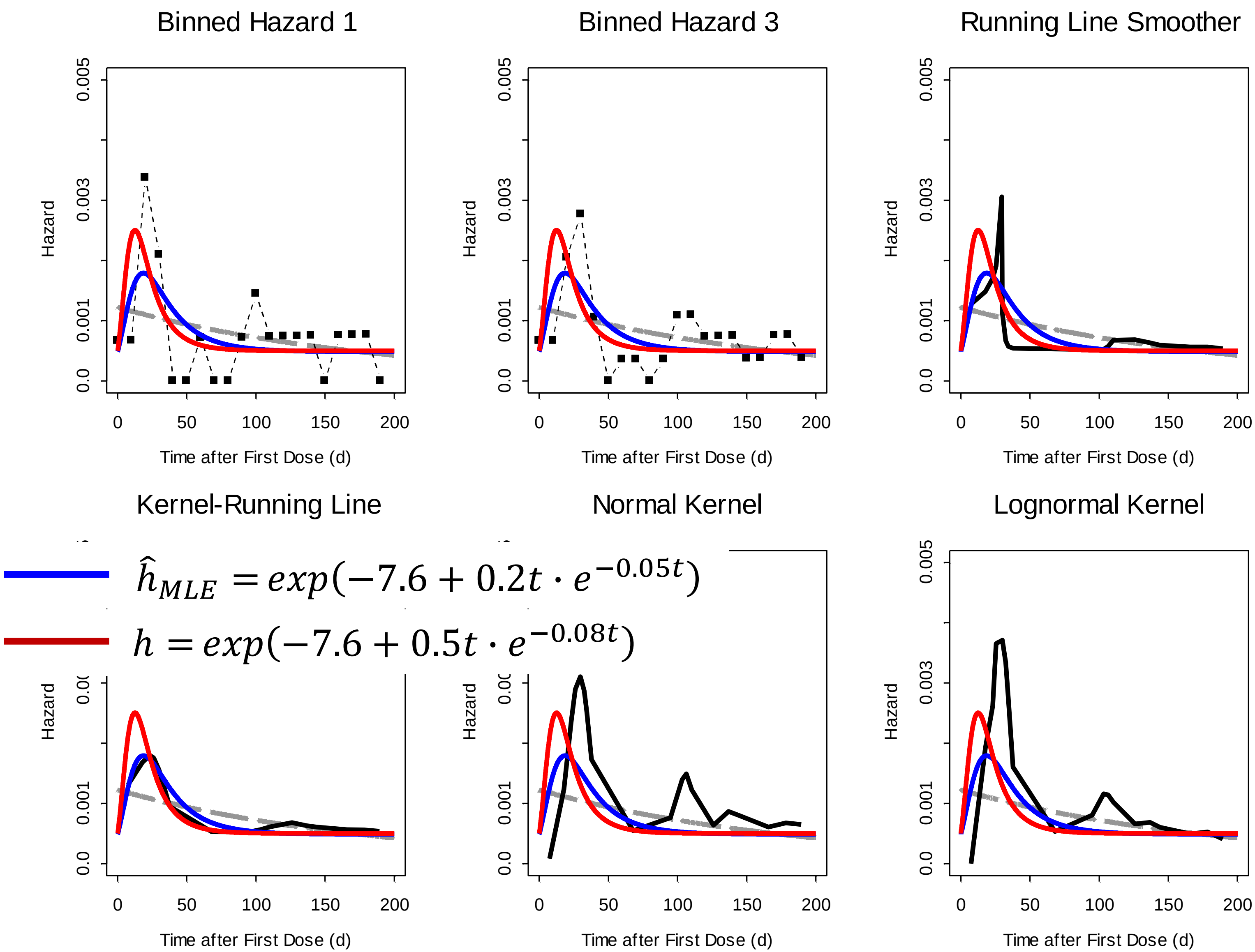


Overall Comparison – the Reveal



$$\hat{h}_{MLE} = exp(-7.6 + 0.2t \cdot e^{-0.05t})$$

Overall Comparison



Discussion

Pharmacometric analysis of TTE data is increasing

Evidenced by:

Holford N, Lavielle M. A tutorial on time to event analysis for mixed effect modellers. PAGE 20 (2011) Abstr 2281 [www.page-meeting.org/?abstract=2281]

Hu C, Szapary P, Yeilding N, Zhou N. Informative dropout and visual predictive check of exposure-response modeling of ordered categorical data. PAGE 20 (2011) Abstr 1991 [www.page-meeting.org/?abstract=1991]

Frobel AK, et. al. A time-to-event model for acute rejections in paediatric renal transplant recipients treated with ciclosporin A. PAGE 21 (2012) Abstr 2374 [www.page-meeting.org/?abstract=2374]

Etc…

Tendency to use Cox PH modeling, but it targets *relative* risk, not ***absolute*** risk. Therefore, information is lost on how to dose the patient. Establishing absolute risk provides another dimension besides ***Dose*** for optimizing therapy → ***TIME***

Discussion

$S(t) = exp\left[-\int_0^t h\,(m)dm\right]-$ it is the exponentiated negative AUC of the hazard, so the changes may not be as dynamic, also the VPC of the hazard is *more direct* to what we are *modeling* and it reflects changes in *risk over time*

It is hoped that these techniques will elicit increased numbers of parametric time-to-event analyses and the pursuit of estimating absolute risk

Thank You!

References:

1. Aalen OO (1978) Nonparametric inference for a family of counting processes. *Ann Statist* 6:701-726
2. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53:457-481
3. Liu RYC, Van Ryzin J (1985) A histogram estimator of the hazard rate with censored data. *Ann Statist* 13:592-605
4. Müller HG (1991) Smooth optimum kernel estimators near endpoints. *Biometrika* 78:521-530
5. Nelson W (1972) Theory and application of hazard plotting for censored failure data. *Technometrics* 14:945-965
6. Ramlau-Hansen H (1983) Smoothing counting process intensities by means of kernel functions. *Ann Statist* 11;453-466
7. Watson GS, Leadbetter MR (1964) Hazard analysis I. *Biometrika* 51;175-184