Time to Event Tutorial

Nick Holford
Dept Pharmacology & Clinical Pharmacology
University of Auckland, New Zealand

Marc Lavielle
INRIA Saclay, France

Outline

• The Hazard: Biological basis for survival

• Types of Event and their Likelihood
  » Exact time
  » Right censored
  » Interval censored
  » Count data

• Joint Modelling of Continuous and Event Data

How Not to Understand Time to Event


Relative Risk=0.7 (0.58-0.8 95%CI)

This landmark study led to the introduction of statins with a major impact on cardiovascular morbidity and mortality worldwide. However, this Kaplan-Meier plot shows that statins don’t seem to have any effect on survival until at least a year after starting treatment. As far as I know there has never been any good explanation of why the benefits of statins are so delayed but when properly analysed this kind of survival data can describe the time course of hazard and give a clearer picture of how long it takes for statins to be effective.
Why do women live longer than men?

Life is hazardous

The hazard describes the death rate at each instant of time. The shape of the hazard function over the human life span has the shape of a bathtub. US mortality data shows the hazard at birth falls quickly and eventually returns to around the same level by the age of 60. The hazard is approximately constant through childhood and early adolescence. The onset of puberty and subsequent life style changes (cars, drugs,...) adopted by men increases the hazard to a new plateau which lasts for 10 to 20 years. It would require a time varying model to describe how development (children) and ageing (adults) are associated with changes in death rate.
Why Pharmacokinetists are Time to Event Experts

- What is an elimination rate constant?
  - Proportionality factor relating elimination to amount of drug
    \[ \text{RateOut} = k \cdot \text{Amount} \]

- What is a hazard?
  - Proportionality factor relating death rate to number of people still alive
    \[ \text{RateOut} = h \cdot N_{ALIVE} \]

- Everything you know about elimination rate constants applies to hazards!

The elimination rate constant is the hazard of a molecule ‘dying’. Elimination rate constants and hazards always have units of 1/time. Unlike most drugs the hazard is not usually constant (‘first-order elimination’) but may change with time (‘time dependent clearance’) or with the number of people (‘concentration dependent clearance’).

PK and Survival

<table>
<thead>
<tr>
<th>Drug</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate of loss</td>
<td>RateOut = \frac{dA}{dt} = k \cdot A</td>
</tr>
<tr>
<td>Non-people alive</td>
<td>RateOut = \frac{dN}{dt} = \lambda \cdot N</td>
</tr>
<tr>
<td>Hazard</td>
<td>k_d</td>
</tr>
<tr>
<td>Integral</td>
<td>AUC</td>
</tr>
<tr>
<td>Non-parametric</td>
<td>Kaplan-Meier</td>
</tr>
<tr>
<td>Time Course</td>
<td>C(t) = \exp(-k_d \cdot t), S(t) = \exp(-\lambda \cdot t)</td>
</tr>
</tbody>
</table>

The event rate is frequently scaled to a standard number of persons e.g. death rates per 100,000 people. Hazard models are more typically scaled to a single person. Pharmacokinetic models are scaled to the dose. In this example a unit dose is assumed for the time course of concentration.

Some examples of baseline hazard functions

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Hazard Function $\lambda(t)$</th>
<th>Survivor Function $P(T&gt;t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>$\lambda(t) = \beta_0$</td>
<td></td>
</tr>
<tr>
<td>Gompertz</td>
<td>$\lambda(t) = \beta_0 \cdot e^{\lambda \cdot t}$</td>
<td></td>
</tr>
<tr>
<td>Weibull</td>
<td>$\lambda(t) = \beta_0 \cdot e^{\lambda \cdot t^{\alpha}}$</td>
<td></td>
</tr>
</tbody>
</table>

The hazard function is associated with a distribution of event times. Some common distributions have names e.g. Gompertz (one of the first mathematicians to explore survival analysis). Standard baseline hazard functions used by statisticians are typically chosen for their mathematical simplicity rather than any biological reason. (comment from Marc: not true and not relevant at all)

The biology of event time distributions is largely based on descriptive and empirical approaches. However, the hazard is the way to introduce biological mechanism in order to aid understanding of the variability of time to event distributions.

The Weibull distribution is traditionally written as a power function of time. It can be reparameterized (as shown here) to show it’s close connection to the exponential distribution (when $\beta_1$ is zero) and the Gompertz distribution (ln(time) instead of time). (comment from Marc: "technical")
Comment of little interest for this tutorial
Note that the Weibull has the often non-biological property of a zero hazard when time is zero. (Comment from Marc: not true and not relevant)

The explanatory variable function is quite empirical. This form is used because there are some simple solutions for integrating the hazard and the exponential form ensures that the hazard is always non-negative.

The Cox proportional hazards model is a semi-parametric version of this parametric model. The Cox model does not estimate \( \lambda(t) \) but assumes it is similar for all cases of the explanatory variables. (Comment from Marc: this remark is incorrect and should be replaced by “Sir David Cox observed that if the proportional hazards assumption holds (or, is assumed to hold) then it is possible to estimate the effect parameter(s) without any consideration of the hazard function”.)

The coefficients of the exponential function are convenient for describing how the hazard varies with the explanatory variable. Exponentiation of the coefficient gives the hazard ratio for the effect of the explanatory variable.

Proportional hazards model

\[
\lambda(t) = \lambda_0(t) \cdot e^{\beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \ldots + \beta_n \cdot x_n}
\]

\( \lambda_0(t) \) : baseline hazard function,
• parametric (constant, Weibull, Gompertz, …)
• non-parametric (Cox model)
\( x_1, x_2, \ldots, x_n \) independent variables (covariates)

Exponentiation of the explanatory variable function ensures non-negative hazards

Example of proportional hazards model

\[
\lambda(t) = \lambda_0(t) \cdot e^{\beta_1 \cdot x_1 + \beta_{\text{SEX}} \cdot \text{SEX} + \ldots + \beta_n \cdot x_n}
\]

If the \( \text{SEX} \) is 0 for females and 1 for males and the value of \( \beta_{\text{SEX}} \) is 0.693 then the hazard ratio for men is 2 (compared to women).
Hazard and Survival

<table>
<thead>
<tr>
<th>Hazard function</th>
<th>$\lambda(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulative hazard function</td>
<td>$\Lambda(a,b) = \int_a^b \lambda(t) dt$</td>
</tr>
<tr>
<td>Survival function</td>
<td>$P(T &gt; t) = e^{-\Lambda(t_0,t)}$</td>
</tr>
<tr>
<td>Probability density function</td>
<td>$p(t) = \lambda(t)e^{-\Lambda(t_0,t)}$</td>
</tr>
<tr>
<td>Cumulative distribution function</td>
<td>$P(T &lt; t) = \int_0^t p(s) ds$</td>
</tr>
</tbody>
</table>

$t_0$: start of the experiment

Marc: I removed the word 'relative' before likelihood in the definition of pdf. The pdf IS the likelihood. There is nothing 'relative'.

Hazard is the instantaneous rate of the event. The hazard model can be of any form but the hazard cannot be negative. As time passes the cumulative hazard predicts the risk of having the event over the interval 0-t. The risk in any interval a-b is obtained by integrating hazard with respect to time over this interval a-b. In case of multiple events, the risk in interval a-b is the expected number of events in this interval. The probability of survival (not having the event) can be predicted from the cumulative hazard. This is called the survivor function. The probability density function (pdf) describes the likelihood for this random event to occur at a given time. It can be calculated from the survivor function and hazard at that time. The cumulative distribution function, i.e. $P(T < t)$, is the integral of the pdf between 0 and t.

Likelihood of a single event

1) Exact time of event

Single event observations (e.g. death) have just one observation event. The likelihood of a single event is the pdf. Note that this is not the probability of the event at that time.

$t_0=0 \quad x \quad T=a$
Likelihood of a single event

2) Right censored event

If the event is not observed at the end of the experiment, it is “right-censored”: it will (maybe) occur after $t_{\text{end}} = a$.

The likelihood of this right-censored event is $P(T > a)$, i.e. the survivor function computed at time $t=a$.

Assume now that the only information available is that the event occurred in an interval $a-b$; this is called an “interval censored event”.

The likelihood of this interval censored event is the probability that the event occurred between $a$ and $b$.

- A first approach for computing this probability $P(a<T<b)$ decomposes this probability as follows:

$$P(a<T<b) = P(T<b) - P(T<a) = 1 - \exp(-\Lambda(a,b)) - 1 + \exp(-\Lambda(0,a))$$

This first approach is only valid for single events and cannot be extended to repeated time to events (RTTE).

- A second approach for computing this probability $P(a<T<b)$ decomposes the information $a<T<b$ into two successive observations:

  - At time $a$, the event was not observed yet: we know that $T>a$. Then, the first component of the likelihood is the probability $P(T>a) = \exp(-\Lambda(0,a))$.

  - At time $b$, the event was observed: we know that $T<b$, given the previous information that $T>a$. Then, the second component of the likelihood is the conditional probability $P(T<b|T>a)$, i.e. the cumulative distribution function computed on the interval $a-b$: $1 - \exp(-\Lambda(a,b))$.

Then,

$$P(a<T<b) = P(T>a) \times P(T<b|T>a) = \exp(-\Lambda(0,a)) \times (1 - \exp(-\Lambda(a,b)))$$
We will see that this second approach can easily be extended to repeated time to events (RTTE).

Encoding Single Events

- **Exact time of event**: $DV=1$
  
  - $T=a$

- **Right censored event**: $DV=0$
  
  - $T > a$

- **Interval censored event**: $DV=0$ for $T > a$, $DV=2$ for $T < b$

Usually, $DV=1$ is used for an exact time event and $DV=0$ for a right censored event. In the case of an interval censored event, we need an additional coding for the end of the interval. We will use $DV=2$ in this tutorial.

A record at time=0 is needed to define when the hazard integration starts. Remark: the MDV data item is required by NONMEM: it is a reminder that that the interval censored event computes the likelihood from two observation events (MDV=0). This MDV column is not required by MONOLIX since the information given by this column already exists in the DV column.

<table>
<thead>
<tr>
<th>ID</th>
<th>TIME</th>
<th>DV</th>
<th>MDV</th>
<th>Comment</th>
<th>Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>Start observing</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>1</td>
<td>0</td>
<td>Exact Time Event</td>
<td>$\lambda(50)e^{-\lambda(50)}$</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Start observing</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>Censored Event</td>
<td>$e^{-\lambda(100)}$</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Start observing</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>Start Event Interval</td>
<td>$e^{-\lambda(55)}$</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>2</td>
<td>0</td>
<td>End Event Interval</td>
<td>$1 - e^{-\lambda(55)}$</td>
</tr>
</tbody>
</table>
Estimation of the parameters of any hazard model can be done using this kind of code. It uses ADVAN6 to integrate the hazard and obtain the cumulative hazard. This can be used with the hazard at the time of the event to calculate the likelihood of right censored, exact time and interval censored events. Note that the likelihood for an individual is the product of each of the contributions. This is important for interval censored events which are described by the likelihood of the right censoring event at the start of the interval (DV.EQ.0) and the interval censored event at the end of the interval (DV.EQ.2). Random effects on hazard model parameters (e.g. BASHAZ and BETACP) are not estimable with single events.

This code will be implemented in MONOLIX 4.0. A beta version will be available and presented during PAGE 2011.

Repeated event observations (e.g. seizures) have several observation events.
Extension to repeated events

The likelihood of the observations is the joint probability:

\[ L(t) = \prod_{k=1}^{K} \int_{t_{k-1}}^{t_k} \prod_{i=1}^{n} \left( \int_{t_{i-1}}^{t_i} p(t_{i-1}, t_i, t_{i+k-1}) dt_i \right) \prod_{i=1}^{n} (1 - \lambda(t_{i-1}, t_i, t_{i+k-1})) dt_i \]

A careful calculation of the likelihood of repeated events is not straightforward… but is possible!

The same formulas used for exact times of event and right censored events can be used for repeated events.
Extension to repeated events

2) Interval censored events

For each interval, we have to compute 2 likelihoods: the likelihood when the interval starts and the likelihood when the interval ends.

<table>
<thead>
<tr>
<th>ID</th>
<th>TIME</th>
<th>DV</th>
<th>MDV</th>
<th>Comment</th>
<th>Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>t₀</td>
<td>-</td>
<td>1</td>
<td>Start observing</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>t₁</td>
<td>0</td>
<td>0</td>
<td>Start Event Interval</td>
<td>$e^{-A(t_0, t_1)}$</td>
</tr>
<tr>
<td>1</td>
<td>t₂</td>
<td>2</td>
<td>0</td>
<td>End Event Interval</td>
<td>$A(t_1, t_2)$</td>
</tr>
<tr>
<td>1</td>
<td>t₃</td>
<td>0</td>
<td>0</td>
<td>Start Event Interval</td>
<td>$e^{-A(t_2, t_3)}$</td>
</tr>
<tr>
<td>1</td>
<td>t₄</td>
<td>2</td>
<td>0</td>
<td>End Event Interval</td>
<td>$A(t_3, t_4)e^{-A(t_3, t_4)}$</td>
</tr>
<tr>
<td>1</td>
<td>t₅</td>
<td>0</td>
<td>0</td>
<td>Right Censored Event</td>
<td>$e^{-A(t_4, t_5)}$</td>
</tr>
</tbody>
</table>

Any kind of response, continuous or non-continuous, can be used for estimation by using the joint likelihood computed for each observation.

Extension to Joint Models

- Basic concept
  Compute LIKELIHOOD for ANY kind of response
  » Predict likelihood of an observation for a continuous variable (e.g. disease status)
  » Predict likelihood of time of event for time to event data

- All types of response can be combined
  » Continuous, categorical, count, time to event

Applications

- Continuous Response
  » Standard PKPD

- Non-continuous Response
  » Binary Response
    - Awake or Asleep
  » Ordered Categorical Response
  » Count Response
  » Frequency of epileptic seizures
  » Time to Event
    - Death
    - Dropout

- Joint Response
  » Continuous plus non-continuous

NONMEM (and many other parameter estimation procedures) uses the likelihood to guide the parameter search. The likelihood is the fundamental way to describe the probability of any observation given a model for predicting the observation. NONMEM shields us from the details for common PKPD models that use continuous response scales for the observation (e.g. drug concentration, effect on blood pressure).

A variety of non-continuous responses are widely used to describe drug effects – especially clinical outcomes. By computing the likelihood directly for each of these kinds of response we can ask NONMEM to estimate parameters for any mixture of response types.
Example of Joint Model: Disease Progress and Time Varying Hazard

1) Continuous biomarker

\[ f(t) = a + bt \]
\[ y(t) = f(t) + \varepsilon(t) \]

2) Time to event

\[ \lambda(t) = h \cdot e^{bf(t)} \]

Statistical model:
- IIV on \( a \) and \( b \)
- Treatment effect on \( b \)

This illustrates joint modelling for disease progress and an event. The event hazard depends on disease progress.

A differential equation is used to integrate the hazard.

An effect of treatment (TRT) is assumed to affect the intercept of the disease progress model which in turn influences the hazard of the event.

It is useful to be able to save the value of the cumulative hazard in order to calculate the likelihood of an interval censored event. In this example DV=0 is used to indicate the start of the interval censored event period and the cumulative hazard at this time is saved in the CHLAST variable.

The F_FLAG variable is used to tell NONMEM how to use the predicted \( Y \) value. F_FLAG of 0 is the default i.e. \( Y \) is the prediction of a continuous variable. F_FLAG of 1 means the prediction is a likelihood. F_FLAG of 2 means the prediction is \(-2\ln(\text{Likelihood})\).
### Slide 30

**Disease Progress and Time Varying Hazard**  
**MONOLIX 4**

**$DATA$** information in the dataset
- ID, TRT use= cov type =cat, TIME, DVID, DV, MDV

**$INDIVIDUAL$** distribution of the individual parameters
- default dist=log-normal, INTRI, SLOPE cov=TRT, BASHAZ iv=no, BETADP iv=no

**$EQUATION$**
- DISPRG= INTRI + SLOPE*T

**$EVENT$**
- lambda=BASHAZ*EXP(BETADP*DISPRG)

**$OBSERVATIONS$** distribution of the observations
- Biomarker type=continuous pred=DISPRG err=constant,
- Death type=event hazard=lambda

---

### Slide 31

**Extension to count data**

The exact times of event

\[
\begin{align*}
& t_0 & T_1 & T_2 & T_3 & T_4 & T_5 & T_6 & T_7 & T_8 & T_9 & T_{10} & T_{11} & T_{12} \\
\end{align*}
\]

---

### Slide 32

**Extension to count data**

The exact times of event

\[
\begin{align*}
& t_0 & T_1 & T_2 & T_3 & T_4 & T_5 & T_6 & T_7 & T_8 & T_9 & T_{10} & T_{11} & T_{12} \\
\end{align*}
\]

are not observed …

\[
\begin{align*}
& t_0 & t_1 & t_2 & t_3 & t_4 & t_5 & t_6 & t_7 \\
\end{align*}
\]
Extension to count data

The exact times of event

\[ t_0 \quad t_1 \quad t_2 \quad t_3 \quad t_4 \quad t_5 \quad t_6 \quad t_7 \quad t_8 \quad t_9 \quad t_{10} \quad t_{11} \quad t_{12} \]

are not observed …

\[ t_0 \quad \times \quad \times \quad \times \quad \times \quad \times \quad \times \quad \times \quad \times \quad \times \quad \times \quad \times \quad \times \]

Only the number of events in each interval is observed

\[ t_0 \quad t_1 \quad t_2 \quad t_3 \quad t_4 \quad t_5 \quad t_6 \quad t_7 \]

The count data is a (non homogenous) Poisson process. The expected number of events in interval \([a, b]\) is the expected number of events in this interval: it is defined as the risk (the cumulative hazard) in this interval.

The expected number of events in interval \([a, b]\) is \( \Lambda(a, b) = \int_a^b \lambda(t) \, dt \)

The count data is a (non homogenous) Poisson process. The expected number of events in interval \([a, b]\) is \( \Lambda(a, b) = \int_a^b \lambda(t) \, dt \)

Unlike the previous examples the DV value is used to indicate the number of events in the interval. It does not indicate the event type (exact time, right, interval censored).
The severity of Parkinson’s disease is usually assessed by the Unified Parkinson’s disease response scale (UPDRS). The UPDRS score increases with time as the disease progresses. The disease status can be described by a model for disease progression (natural history) and the effects of treatment e.g. the use of levodopa (the mainstay of treatment) with or without deprenyl (a monoamine oxidase inhibitor commonly used as an adjunctive treatment). The hazard of a clinical outcome event e.g. death, can be described by a baseline hazard, \( h(t) \), and explanatory factors such as drug treatment and the time course of disease status. Other factors (age, sex, smoking, etc) are easily included in this kind of model.

The change of disease status, reflected by the time course of UPDRS, is the most important factor determining the hazard of clinical outcome events in Parkinson’s disease. The different shapes of the survival function for death, disability, cognitive impairment and depression reflect different contributions of disease status to the probability of not having had the event as time passes.
Backup Slides

Constant hazard $\lambda(t) = \lambda$

$T$ is a random variable with an exponential distribution:

$$P(T > t) = e^{-\lambda t}$$

The survival function of a constant hazard decreases exponentially to 0.

Constant hazard makes the very strong assumption of memoryless. The modeller should be aware of this strong assumption at the time to select a hazard function.

Consider for example that your event is the first passing of the viral load (HIV, HCV,...) under a given threshold (e.g. LOQ). Here, $t_0$ is the time when the active treatment starts. We assume that the initial viral load at $t_0$ is above this threshold. Then:

- the hazard is 0 at $t_0$ and increases with time
- if you know that you are still above the threshold after 6 months for instance, then this information will "modify" the distribution of your event time:

$$P(T > t+a | T>a) > P(T>t | T>0)$$

In other words, you are more likely to be a non-responder and the probability to reach the threshold decreases.

This is one of the many examples where a constant hazard is a very poor choice and
when alternative models (Weibull for instance) should be considered.

Slide 42

Parametric Regression
In Standard Packages

- Estimation of hazard parameters is done after transformation e.g. \( \ln(T) \)
- Explanatory variable model is then linear regression e.g. for Weibull

\[
\ln(T_i) = \frac{1}{\gamma} \sum_{j=1}^{n} (\lambda_j - \beta_1 \cdot x_{i1} - \beta_2 \cdot x_{i2} - \ldots - \beta_p \cdot x_{ip} + \epsilon_i)
\]

or more generally

\[
\ln(T_i) = \mu + \alpha_1 x_{i1} + \alpha_2 x_{i2} + \ldots + \alpha_p x_{ip} + \sigma \cdot \epsilon_i
\]

Note that covariates \((x_1 \ldots x_p)\) are usually assumed to be time invariant

Standard survival analysis is equivalent to non-compartmental PK. It is useful for description but ignores time variation.

Slide 43

Distribution of Survival Times
Michaelis-Menten Elimination

A useful view of survival is to look at the probability density function for the survival times.

\[
\text{Survival}(t) = e^{-\int_0^t \text{hazard}(t) dt}
\]

\[
\text{PDF}(t) = \text{Survival}(t) \cdot \text{hazard}(t)
\]
How can the effect of treatment \( Rx(t) \) be described?

\[
h(t) = f(sex, race, age(t), Rx(t), \ldots)
\]

Standard survival analysis can include varying age implicitly. Adding time-varying covariates for survival analysis is harder to do because of the need to integrate the hazard. Drug treatments will often change with time and if expressed in terms of drug concentration the hazard could change in proportion to concentration after every dose.

Survivor Function

An example of how to simulate the time course of survivor function, cumulative hazard and pdf with a continuously time varying hazard using Berkeley Madonna code.

METHOD RK4

STARTTIME = 0
STOPTIME = 10
DT = 0.02

\( beta0 = 0.1 \)
\( betaStatus = 0.01 \)
\( S0 = 20 \)
status = S0 + 12 * time

\( hazpla = beta0 * \exp(betaStatus * S0) \)
\( haztrt = beta0 * \exp(betaStatus * status) \)

\( init(cumpla) = 0 \)
\( d/dt(cumpla) = hazpla \)
\( survpla = \exp(-cumpla) \)

\( init(cumtrt) = 0 \)
\( d/dt(cumtrt) = haztrt \)
\( survtrt = \exp(-cumtrt) \)

\( pdfpla = survpla * hazpla \)
\( pdftrt = survtrt * haztrt \)
Cumulative Hazard and Relative Risk

<table>
<thead>
<tr>
<th>Time (y)</th>
<th>Cumulative Hazard</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>2.5</td>
</tr>
<tr>
<td>10</td>
<td>2.5</td>
<td></td>
</tr>
</tbody>
</table>

Relative Risk
- Constant Hazard
- Time varying hazard

Probability Density Function

<table>
<thead>
<tr>
<th>Time (y)</th>
<th>S(t) * h(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.15</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>0.05</td>
</tr>
<tr>
<td>6</td>
<td>0.0</td>
</tr>
<tr>
<td>8</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Time (y)
- Constant Hazard
- Time varying hazard

Hazard models link disease progress and clinical outcome probability

Survivor Function
\[ S(t) = \Pr(T \geq t) = e^{-\int_0^t h(t)} \]

Hazard Function
\[ h(t) = \beta_0 \exp(\beta \cdot \text{status}(t)) \]

Hazard (1/y)

<table>
<thead>
<tr>
<th>Time (y)</th>
<th>Hazard</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>0.02</td>
</tr>
<tr>
<td>6</td>
<td>0.01</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
</tr>
<tr>
<td>10</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Survivor Function

<table>
<thead>
<tr>
<th>Time (y)</th>
<th>Survivor Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>0.7</td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
</tr>
<tr>
<td>6</td>
<td>0.1</td>
</tr>
<tr>
<td>8</td>
<td>0.0</td>
</tr>
<tr>
<td>10</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Time (y)
- Constant hazard
- Hazard changes with status

©NHG Holford, M. Lavielle 2011, all rights reserved.
Likelihoods for Survival

An alternative way of describing the likelihoods in terms of the survivor function and hazard function alone.

For an uncensored datum, with \( T_i \) equal to the age at death, we have

\[
\Pr(T = T_i | \theta) = f(T_i | \theta) \cdot h(T_i) \cdot S(T_i) \cdot h(T_i) \]

For a left censored datum, such that the age of death is known to be less than \( T_i \), we have

\[
\Pr(T < T_i | \theta) = F(T_i | \theta) = 1 - S(T_i) \]

For a right censored datum, such that the age of death is known to be greater than \( T_i \), we have

\[
\Pr(T > T_i | \theta) = 1 - F(T_i | \theta) = S(T_i) \]

For an interval censored datum, such that the age of death is known to be greater than \( T_{i,l} \) and less than \( T_{i,u} \), we have

\[
\Pr(T_{i,l} < T < T_{i,u} | \theta) = S(T_{i,u} | \theta) - S(T_{i,l} | \theta) \]