Slide 1

Time to Event Analysis

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

Slide 2

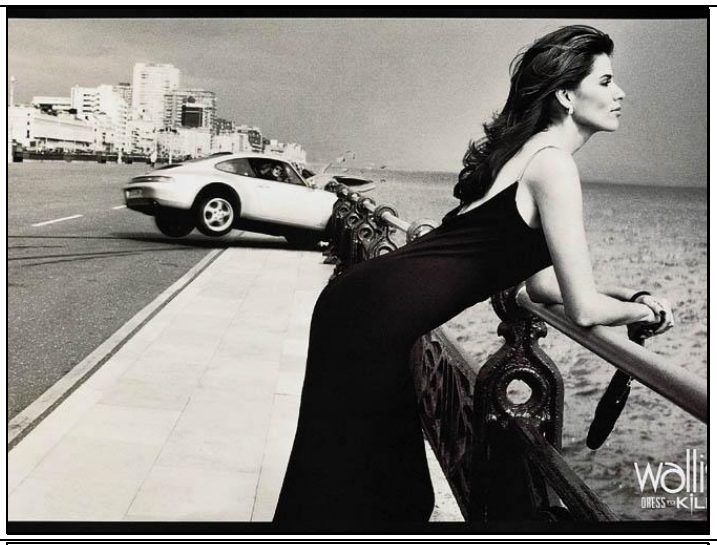
Outline

- The Hazard: Biological basis for survival
- The Pharmacokinetics of Survival
- Joint Models of PKPD and Time to Event
  - » Continuous
  - » Right censored
  - » Interval censored
- An Example (fracture risk reduction)

Slide 3

Why do women live longer than men?

Slide 4



http://www.allowe.com/Humor/whymendieyounger.htm

Slide 5

Survival in a Bathtub

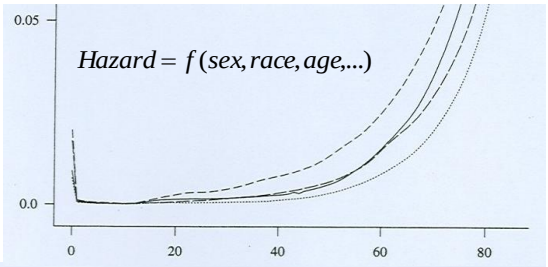


Figure 2.6 *Hazard functions for all cause mortality for the US population in 1989. White males (———); white females (·····); black males (- - - - -); black females (—— ——).*

"… a bathtub-shaped hazard is appropriate in populations followed from birth."
Klein, J.P., and Moeschberger, M.L. 2003. *Survival analysis: techniques for censored and truncated data*. New York: Springer-Verlag.
http://en.wikipedia.org/wiki/Bathtub_curve "The bathtub curve"

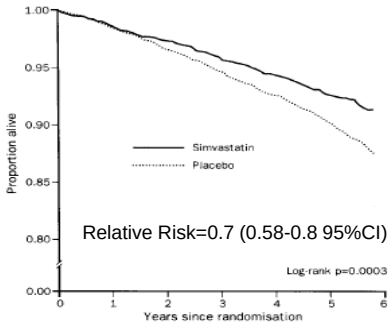
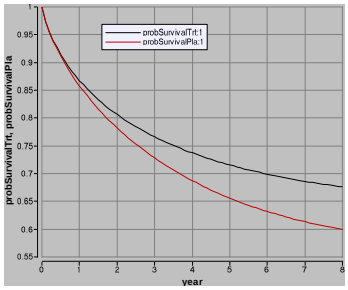
The hazard describes the death rate at each instant of time. The shape of the hazard function over the human life span has the shape of a bathtub.
US mortality data shows the hazard at birth falls quickly and eventually returns to around the same level by the age of 60. The hazard is approximately constant 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.

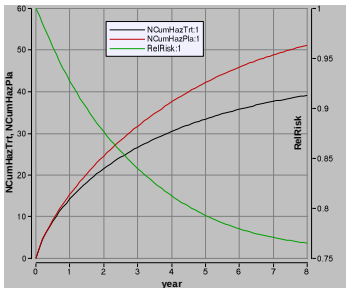
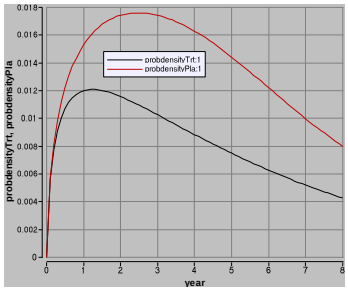
Slide 6

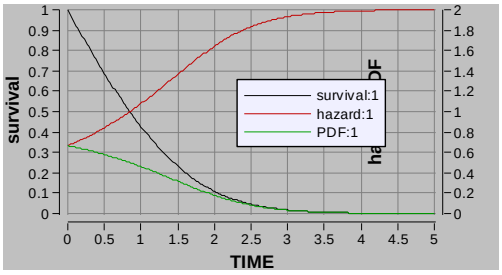
Survival and PK

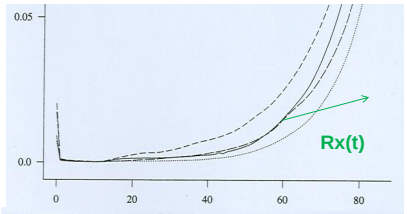
	Drug	Events
Rate of loss N=people alive A=molecules remaining	$\frac{dA}{dt} = -k_{el} \cdot A$	$\frac{dN}{dt} = -\lambda \cdot N$
Hazard	k_{el}	λ
Integral	AUC	Cumulative Hazard
Non-parametric	Non-compartmental	Kaplan-Meier
Time Course	$C(t) = \exp(-k_{el} \cdot t)$	$S(t) = \exp(-\lambda \cdot t)$

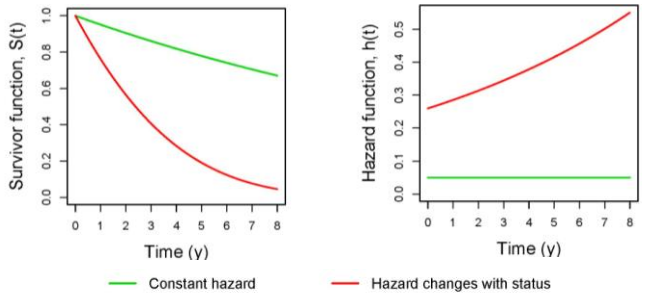
The event rate is frequently scaled to a standard number of persons e.g. death rates per 100,000 people.
Hazard models are more typically scaled to a single person.
Pharmacokinetic models are scaled to the dose. In this example a unit dose is assumed for the time course of concentration.

Slide 7	**The Kaplan-Meier Plot and Survivor Function**  Scandinavian Simvastatin Survival Study Group. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). Lancet. 1994;344:1383-89. 	
Slide 8	**Hazard and Survival** $$hazard = h(t)$$ $$CumulativeHazard(t) = \int_0^t h(t)$$ $$Survivor(t) = \Pr(T > t) = e^{-CumulativeHazard(t)}$$ $$pdf(t) = Survivor(t) \cdot h(t)$$ 	The probability of a having an event at a particular time can be predicted by describing the hazard for the event. Hazard is the instantaneous rate of the event. As time passes the cumulative hazard predicts the risk of having the event over the interval 0-t. The hazard model can be of any form but the hazard cannot be negative. The risk is the cumulative hazard. It is obtained by integrating hazard with respect to time. The probability of survival (not having the event) can be predicted from the cumulative hazard. This is called the survivor function. The probability of having an event at a particular time is predicted by the probability density function (pdf(t)). The pdf can be calculated from the survivor function and hazard at that time. The cumulative density is the integral of the pdf.
Slide 9	**Survivor Function**  	beta0=1 betaWB=1 betaTRT=-logn(2) hazpla=if (time>0) then beta0*exp(betaWB*logn(time)) else 0 haztrt=if (time>0) then beta0*exp(betaWB*logn(time)+betaTRT) else 0 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

Slide 10	**Cumulative Hazard and Relative Risk**  	
Slide 11	**Probability Density Function**  	
Slide 12	**Pharmacokinetic Survival Non-Linear** $\text{hazard}(t) = \text{Rateconstant}(t) = (\text{Vmax}/\text{V})/(\text{Km} + \text{C}(t))$ $\text{Risk}(t) = \text{CumConstant}(t) = \int_0^t \text{Rateconstant}(t)$ $\text{Survival}(t) = C(t) = e^{-\text{CumConstant}(t)}$ 	If the hazard varies with time (or a function of time, such as concentration) then exactly the same relationships between hazard, risk and survival exist as for the constant hazard case. It is possible to make non-linear pharmacokinetic model predictions of the amount eliminated and the time course of concentration using survival analysis functions.

Slide 13	**Non-linear PK and Survival** dose=100 v=1 vmax=100 km=50 rateconstant=(vmax/v)/(km+conc) init(conc)=dose/v d/dt(conc)= -rateconstant*conc init(cumconstant)=0 d/dt(cumconstant)=rateconstant survival=dose*exp(-cumconstant) Time / Conc / Survival 0 100 100 1 42.6303 42.6303 2 10.8858 10.8858 3 1.76793 1.76793 4 0.246655 0.246655 5 0.033524 0.033524 6 0.00454 0.00454 7 6.14E-04 6.14E-04 8 8.32E-05 8.32E-05 9 1.13E-05 1.13E-05 10 1.52E-06 1.52E-06 	
Slide 14	**Distribution of Survival Times Michaelis-Menten Elimination**  $Survival(t) = e^{-\int_0^t hazard(t)}$ PDF(t)=Survival(t)·hazard(t) 	A useful view of survival is to look at the probability density function for the survival times.
Slide 15	**Baseline Hazard Functions** $h(t) = \lambda$ Exponential $h(t) = \lambda \cdot \gamma \cdot t^{\gamma-1}$ $h(t) = \lambda \cdot \gamma \cdot e^{(\gamma-1)\cdot\ln(t)}$ $h(t) = \beta_0 \cdot e^{\beta_1 \cdot \ln(t)}$ Weibull $h(t) = \lambda \cdot e^{\beta_1 \cdot t}$ Gompertz Parametric: Can be used for simulation and time varying hazards h(t) = Similar for everyone but undefined Cox Proportional Non-Parametric: No good for simulation. Tricky with time varying hazards. 	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 chosen for their mathematical simplicity rather than any biological reason. The biology of event time distributions is largely based on descriptive and empirical approaches. However, the hazard is the way to introduce biological mechanism and understanding the variability of time to event distributions.

Slide 16	**Explanatory Variable Functions** $h(t) = \beta_0 \cdot e^{\beta_1 \cdot x_1 + \beta_2 \cdot x_2 \ldots + \beta n \cdot x_n}$ Includes exponential, Weibull, Gompertz as special cases $h(t) = \beta_0 \cdot e^{f(B,X)}$ Linear and non-linear functions of explanatory variables 	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.
Slide 17	**Exponential Coefficients and the Hazard Ratio** $h(t) = \beta_0 \cdot e^{\beta_1 \cdot x_1 + \beta_{SEX} \cdot SEX \ldots + \beta n \cdot x_n}$ If the explanatory variable is 0 for females and 1 for males and the value of β_{SEX} is 0.693 then the hazard ratio for men is 2 (compared to women). 	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.
Slide 18	**How can the effect of treatment Rx(t) be described?**  **Figure 2.6** *Hazard functions for all cause mortality for the US population in 1989. White males (———); white females (·····); black males (- - - - -); black females (— — —).* h(t) = f(sex, race, age(t), **Rx(t)**,…) 	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.

Slide 19	**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** $h(t)=\beta_0$ $h(t)= \beta_0 \cdot exp(\beta_{status} \cdot status(t))$  	
Slide 20	**Joint Models** ● Basic concept Compute LIKELIHOOD for ANY kind of response ● Predict likelihood of time of event instead of the time of event ● All types of response can be combined 	Any kind of response, continuous or non-continuous, can be combined in a NONMEM model by using the joint likelihood computed for each observation. It is up to the user to compute the likelihood (or -2LL) for any non-continuous responses. In NONMEM VI the F_FLAG variable is used to distinguish the type or prediction being made. In NONMEM V user written CCONTR is used to tell NONMEM how to compute the likelihood (or -2LL) for continuous responses.
Slide 21	**Applications** ● Continuous Response » Standard PKPD ● Non-continuous Response » Binary Response – Awake or Asleep » Ordered Categorical Response – Neutropenic adverse event type » Count Response – Frequency of epileptic seizures » Time to Event – Bone fracture » Dropout – Missing data ● 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.

Slide 22

NONMEM

- NONMEM Maximises the Likelihood
- $ESTIMATION

  `METHOD=CONDITIONAL LAPLACIAN LIKE` or `-2LL`

NONMEM maximises the likelihood when estimating parameters.

There is a pair of default subroutines (CCONTR and CONTR) that are used to compute the Conditional CONTRibution to the likelihood.

The NM-TRAN $ESTIMATION record can have an option that allows the user to directly return the likelihood or -2LL instead of letting NONMEM compute it with its CCONTR and CONTR subroutines. This option always requires the CONDITIONAL method (FOCE) and is thought to work better if the LAPLACIAN option is used. NONMEM VI allows the user to perform joint modelling of different types of data by using the F_FLAG variable to indicate what is returned by the prediction for Y.

Slide 23

Likelihood for Continuous Variable

$$Like_{ij} = \exp\left[-0.5 \cdot \left(\left(\frac{Y_{OBS,ij} - Y_{PRED,ij}}{SD_{ij}}\right)^2 + \ln\left(SD^2_{ij} \cdot 2\pi\right)\right)\right]$$

For a continuous response type (the default for NONMEM) the likelihood is computed using the extended least squares objective function (ELS). Each observation (Yobs) and each prediction (Ypred), along with the predicted variance of the difference between Yobs and Ypred (Var) are used to compute a contribution to the ELS objective function. This contribution is summed over all subjects and all observations to compute -2 times the log of the likelihood. Actually its not quite -2LL but proportional to it. It is missing a constant (NOBS*Ln(2*Pi)) but this is not needed in order to minimize -2LL and obtain maximum likelihood estimates.

The likelihood of non-continuous types of event will depend on the type of event. The likelihood of an event at time t is given by the probability density function at that time.

Slide 24

Likelihoods for Survival

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

$$\Pr(T = T_i|\theta) = f(T_i|\theta)$$ = $S(T_i|\theta) * h(T_i)$

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

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

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

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

For an interval censored datum, such that the age at death is known to be greater than $T_{i,r}$ and less than $T_{i,l}$, we have

$$\Pr(T_{i,l} < T < T_{i,r}|\theta) = S(T_{i,l}|\theta) - S(T_{i,r}|\theta)$$

http://en.wikipedia.org/wiki/Survival_analysis

Slide 25

Time Varying Hazard (CP) With Right Censoring using NONMEM

```
$ESTIM MAXEVAL=9990 METHOD=COND
LAPLACE LIKE
$THETA
10  FIX     ; POP_CL
100 FIX     ; POP_V
0.1         ; POP_BETA
(0,0.01)    ; BASE
;Random effect for baseline hazard
$OMEGA 0.01 ; PPV_HAZ

$SUBR ADVAN=6 TOL=3
$MODEL
   COMP=(CENTRAL)
   COMP=(CUMHAZ)
$PK
   CL=THETA(1)
   V=THETA(2)
   BETA=THETA(3)
   BASHAZ=THETA(4)*EXP(ETA(1))
```

```
$DES
   DCP=A(1)/V
   DADT(1)=-CL*DCP
   DADT(2)=BASHAZ*EXP(BETA*DCP)

$ERROR
  CP=A(1)/V
  RISK=A(2)          ; cumulative hazard
  SURV=EXP(-RISK)

  IF (DV.EQ.0) THEN
     Y=SURV          ; right censored event
  ELSE
     HAZNOW=BASHAZ*EXP(BETA*CP)
     Y=SURV*HAZNOW ; Like of having an
                   ; event at time=TIME
  ENDIF
```


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. The cumulative hazard is used in $ERROR to calculate the probability of survival (SURV) i.e. the likelihoodof not having the event at a particular time (DV=0). If the event occurs (DV=1) at the time of the record then the likelihood is determined by the probability density function at that time i.e. S(t)*h(t).

Slide 26

Right Censored Data Single Event

#ID	TIME	DV	AMT	Comment
1	0	0	100	Dose
1	50	1	0	**Event**
2	0	0	100	Dose
2	75	1	0	**Event**
3	0	0	100	Dose
3	100	0	0	**Censored Event**
4	0	0	100	Dose
4	25	1	0	**Event**
5	0	0	100	Dose
5	100	0	0	**Censored Event**

Single event observations (e.g. death) have just one observation event.

Slide 27

Interval Censored Time to Event

IF (DV.EQ.0) ; right censored event at t

$$Like_{NoEvent} = S(t) = e^{-\int_0^t h(t)}$$

ELSE ; event between $t_{lastobs}$ and t_{end}

$$Like_{Event} = S(t_{lastobs}) - S(t_{end}) = e^{-\int_{t=0}^{t=t_{lastobs}} h(t)} - e^{-\int_{t=0}^{t=t_{end}} h(t)}$$

Using NONMEM nomenclature, DV is a the observed event state. If it is 0 it means the subject has not had the event. If it is 1 then the subject had an event at some time between the last observation the end time when it is known they have had an event since the last observation time.
At the time a subject is known to have dropped out (DV=1) the likelihood of dropping out in the interval between the last observed time and the end time is given by the difference in the survivor function at the last observed time and the survivor function at the end time.
These two survivor functions can be computed from cumulative hazards from 0 to the last observed time and from 0 to the end time.

Slide 28

Disease Progress and Time Varying Hazard with Interval Censoring (NM7 or NMVI)

```
$INPUT ID TRT TIME DV DVID
$ESTIM MAX=9990 SIG=6 NOABORT
METHOD=CONDITIONAL LAPLACE
$SUBR ADVAN=6 TOL=6
$MODEL
COMP=(CUMHAZ)

$PK
IF (NEWIND.LE.1) SRVZ=1 ; Survival(0) is 1
BSHZ=THETA(1) ; Baseline hazard
BETADP=THETA(2) ; Disease progress hazard
EFFECT=TRT*THETA(3)
INTRI=(THETA(4)+EFFECT)*EXP(ETA(1))
SLOPI=THETA(5)*EXP(ETA(2))

$DES
DISPRG=INTRI + SLOPI*T
EXPHAZ=EXP(BETADP*DISPRG)
DADT(1)=BSHZ*EXPHAZ ; h(t)
```

```
$ERROR
CHZT=A(1) ; Cum hazard at this time
IF (DVID.NE.2) THEN
   F_FLAG=0 ; CONTINUOUS ELS
   Y=INTRI + SLOPI*TIME + ERR(1); Biomarker
ENDIF
SRVT=EXP(-CHZT) ; Survival at t
IF (DVID.EQ.2.AND.DV.EQ.0) THEN
   F_FLAG=1 ; LIKELIHOOD
   Y=SRVT   ; Like no event
ENDIF
IF (DVID.EQ.2.AND.DV.EQ.1) THEN
   F_FLAG=1 ; LIKELIHOOD
   Y=SRVZ-SRVT ; Like event
ENDIF
IF (DVID.EQ.3) THEN ; Last obs before event
   SRVZ=SRVT ; remember survival
ELSE
   SRVZ=SRVZ ; keep NM-TRAN happy
ENDIF
```


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.

When using NMVI it is possible to save the value of the survivor function even if the hazard involves a random variable (i.e. an ETA is used in the computation). The same method can be used with NMV if there is no random variable in the hazard. In this example DVID=3 is used to indicate the start of the interval censored event period and the survivor function (SRVT) is saved in the variable SRVZ. Note to keep NM-TRAN happy it is necessary to explicitly include the assignment of SRVZ for records with DVID not equal to 3.

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(Likelihood).

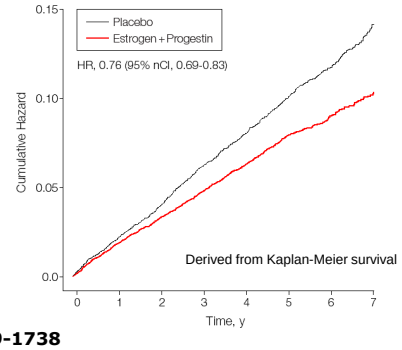
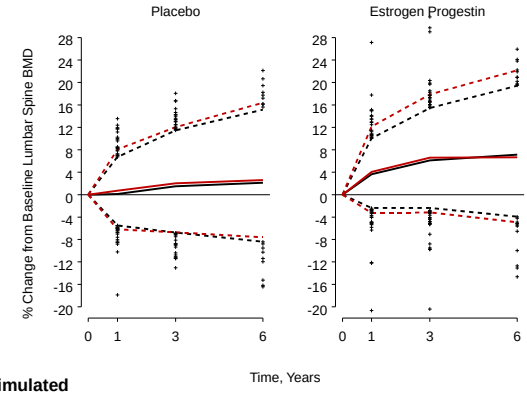
Slide 29

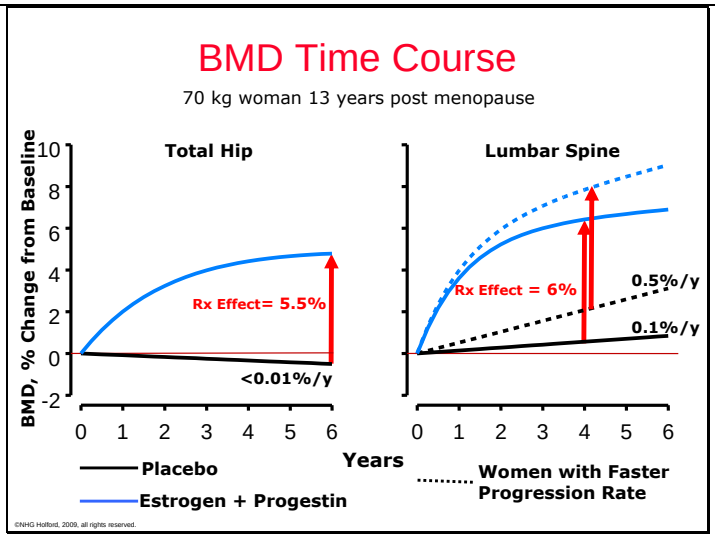
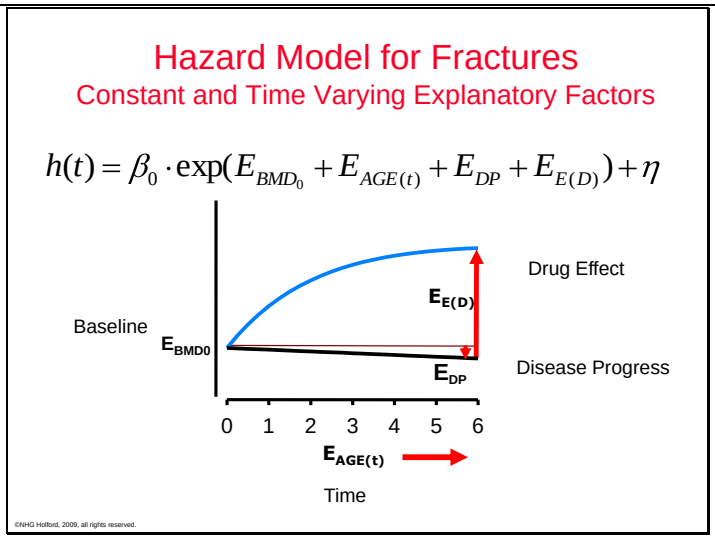
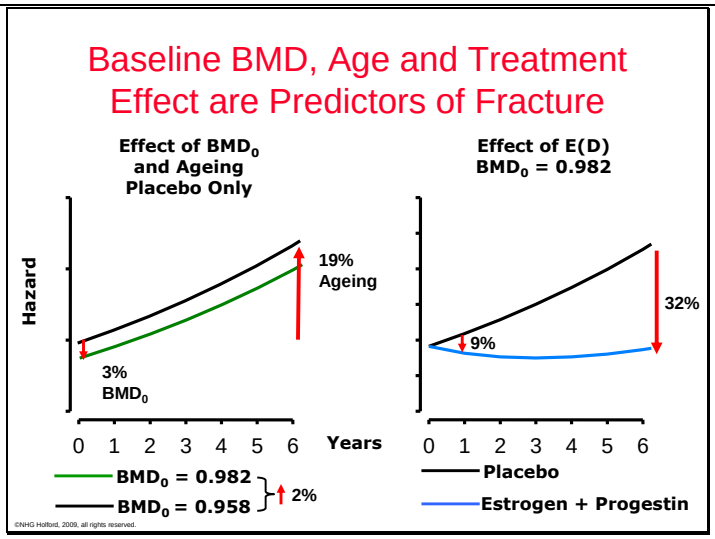
Interval Censored Data (NM7 or NMVI)

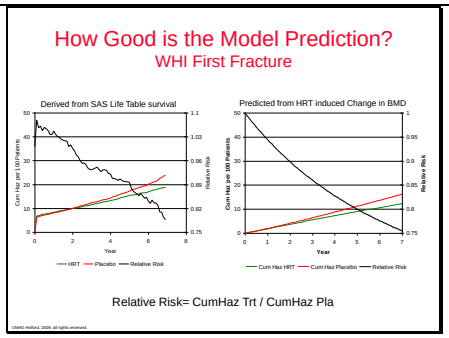
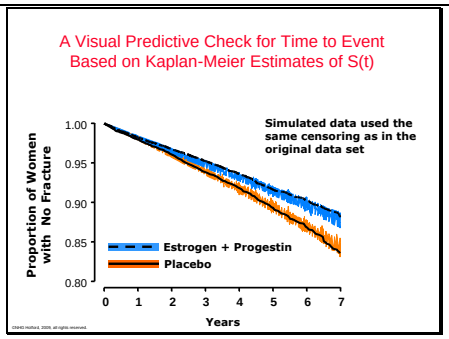
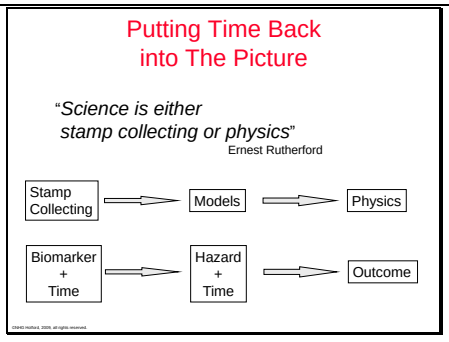
#ID	TRT	TIME	DV	DVID	Comment
1	1	0	-0.6	1	Biomarker Obs
1	1	25	28.1	1	Biomarker Obs
1	1	50	53.2	1	Biomarker Obs
1	1	75	81.8	1	Biomarker Obs
1	1	100	108.7	1	Biomarker Obs
1	**1**	**100**	**0**	**2**	**Censored Event**
2	0	0	0.1	1	Biomarker Obs
2	0	25	28.8	3	Last Non-event Obs
2	**0**	**50**	**1**	**2**	**End Event Interval**

ID 2: Event is between time 25 and 50.

The NONMEM VI dataset format is shown here (can be used with NMV if the hazard does not involve a random variable)
When DVID is 1 this means the DV observation is of the disease state (e.g. viral load).
When DVID is 2 this means the DV observation is event status (0=censored event, 1=had event)
When DVID is 3 this means this is the last observation time before the interval censored event
In this example the first subject did not have an event during the study and the final record has DV=0 to indicate this.
The second subject had an event between time 25 and 50. The DVID is 3 at time 25 to indicate this is the last time the subject was known not to have the event. The final record for this subject indicates that they were known to have had the event by time 50.
The TRT data item is 0 for placebo and 1 for active treatment.
The LOCF data item is used when testing the random missingness model.

Slide 30	# Bone Fractures and Hormone Replacement Therapy Integration of PKPD, Disease Progress and Hazard to Explain the Time Course of Benefit 	
Slide 31	## Effects of Estrogen Plus Progestin on Risk of Fracture and Bone Mineral Density The Women's Health Initiative Randomized Trial Jane A. Cauley, DrPH John Robbins, MD Zhao Chen, PhD Steven R. Cummings, MD Rebecca D. Jackson, MD Andrea Z. LaCroix, PhD, MPH Meryl LeBoff, MD Cora E. Lewis, MD, MSPH Joan McGowan, PhD Joan Neuner, MD, MPH Mary Pettinger, MS Marcia L. Stefanick, PhD Jean Wactawski-Wende, PhD Nelson B. Watts, MD for the Women's Health Initiative Investigators  **JAMA, 2003;290:1729-1738** 	NOTE: Ca/vitamin D at doses administered show increases in BMD without added reduction in fracture (hip and total fracture) Jackson RD, LaCroix AZ, Gass M, Wallace RB, Robbins J, Lewis CE, et al. Calcium plus Vitamin D Supplementation and the Risk of Fractures. N Engl J Med. 2006 February 16, 2006;354(7):669-83.
Slide 32	## Bone Mineral Density – Slow Biomarker Women's Health Initiative Randomized Trial  Simulated Observed Analysis with Christine Garnett (CDDS/FDA) 	The Women's Health Initiative trial observed the time course of changes in bone mineral density in 1000 women who were treated with placebo or with hormone replacement therapy. Both groups were treated with vitamin D and calcium. Half of the placebo patients were given placebo vitamin D and calcium. This plot is a visual predictive check showing the median and 90% interval for the observed (black) and predicted (red) BMD changes. The increase in BMD in the placebo group (and some of the change in the HRT group) is attributable to treatment with vitamin D and calcium.

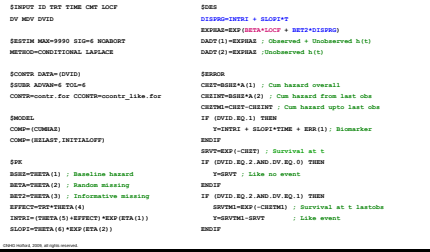
Slide 33		These figures shows some key results for the Hip and Spine models which represent the two different types of bone. **Teq for lumbar spine 0.81 y. Teq for hip 1.53 y.** For the hip bone, there was a trend for bone loss with a progression rate of less than 0.01% per year. Maximum treatment effect was estimated to be 6% of baseline. But by year 6, 94% of treatment effect was observed. For spine, women gained bone mass during the trial. Approximately 52% of the women's progression rate was 0.1% per year and the remaining women gained bone with a rate of 0.5% per year. Maximum treatment effect from hormones was approximately 6% of baseline. Due to the shorter equilibration T1/2, maximum treatment effect was observed by year 4.
Slide 34		Key to our modeling approach was to specify the hazard model for fractures as a function of Bone Mineral Density. Instead of using the predicted BMD as a single time-varying covariate in the hazard, we chose to parameterize the hazard function by including each component of the disease status model as a covariate. This allowed us to assess the relative contributions of each component to the risk of fracture. Some women had more than one fracture which allowed the between subject difference in hazard to be estimated (η). This kind of random effects model is called a 'frailty model' in the statistical survival analysis literature.
Slide 35		This is a deterministic simulation of the final fracture model using total body BMD NOTE: Ca/vitamin D at doses administered show increases in BMD without added reduction in fracture (hip and total fracture. JAMA 2006)

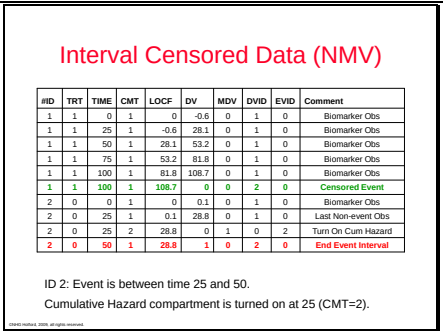
Slide 36			
Slide 37			The slide shows a predictive check using the Kaplan-Meier method to generate the predicted uncertainty in the survivor function (based on 500 replications of the WHI data set). The 90% predictive interval for placebo is shaded in orange and the blue represents the predictive interval for E+P. The black lines represent the observed probability of no fracture. Overall, we concluded that the model describes the observed data well.
Slide 38			

Slide 39	Backup Slides 		

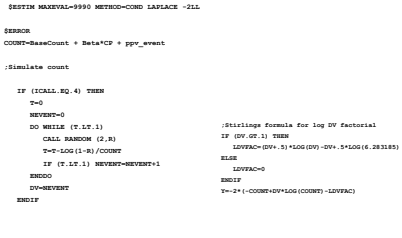
Slide 40	Joint Model NMV
	`$DATA ID TIME AMT TYPE DV`
	`$ESTIM MAXEVAL=9990 METHOD=COND LAPLACE NOABORT`
	`$SUBROUTINE ADVAN=2 TRAN=2`
	`CCONTR=..\ccontr_like.for`
	`CONTR=..\contr.for`
	`$CONTR DATA=(TYPE)`
	`$ERROR`
	`; Continuous response`
	`CP=F`
	`CPEST=CP + ruv_sd`
	`; Non-continuous response`
	`BASE=LOG(BASEP/(1-BASEP))`
	`LGST=BASE + BETA*CP + PPV_EVENT`
	`ODDS=EXP(LGST)`
	`P1=ODDS/(1+ODDS)`
	`IF (DV.EQ.1) ODDEST=P1`
	`IF (DV.EQ.0) ODDEST=1-P1`
	`IF (TYPE.LE.2) THEN`
	`Y=CPEST ; continuous`
	`ENDIF`
	`IF (TYPE.EQ.3) THEN`
	`Y=ODDEST ; non-continuous`
	`ENDIF`
	
			This is a simple example of a continuous response (PK model defined with ADVAN2) and a non-continuous binary response (defined by a logistic model).
			The TYPE data item is used to signal in $ERROR which kind of response to return in the variable Y.
			The $CONTR record tells the CCONTR subroutine what meaning it should attach the TYPE data item.

Slide 41	CCONTR
	`SUBROUTINE CCONTR (ICALL,CNT,P1,P2,IER1,IER2`
	`SAVE`
	`C LVR and NO should match values in NSIZES`
	`PARAMETER(LVR=30,NO=50)`
	`COMMON /ROCM4/ Y(NO),DATA(NO,3)`
	`DOUBLE PRECISION CNT,P1,P2,Y`
	`DIMENSION P1(*),P2(LVR,*)`
	`TYPE=DATA(1,1)`
	`C Value of TYPE is provided as a user defined data item`
	`IF (TYPE.EQ.1)THEN`
	`C CELS is used for continuous type data`
	`CALL CELS(CNT,P1,P2,IER1,IER2)`
	`ELSE`
	`C CLIK is used for LIKE or -2LL`
	`C first argument is 1 for LIKE and 2 for -2LL`
	`CALL CLIK(1,CNT,P1,P2,IER1,IER2)`
	`ENDIF`
	`RETURN`
	`END`
	
			The format of the user supplied CCONTR is shown here. It can be used quite generally. The main user specific feature is to use the value of TYPE (which is determined by a value in the data set) to choose which method should be used t compute the contribution to the likelihood.

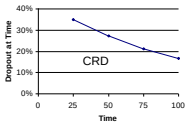
Slide 42	**Time Varying Hazard with Interval Censoring (NMV)**
	 
	$INPUT ID TRT TIME CMT LOCF
	DV MDV DVID
	
	$ESTIM MAX=9990 SIG=6 NOABORT
	METHOD=CONDITIONAL LAPLACE
	
	$CONTR DATA=(DVID)
	$SUBR ADVAN=6 TOL=6
	CONTR=contr.for CCONTR=ccontr_like.for
	
	$MODEL
	COMP=(CUMHAZ)
	COMP=(RELAST,INITIALOFF)
	
	$PK
	BSHZ=THETA(1) ; Baseline hazard
	BETA=THETA(2) ; Random missing
	BET2=THETA(3) ; Informative missing
	EFFECT=TRT*THETA(4)
	INTRI=(THETA(5)+EFFECT)*EXP(ETA(1))
	SLOPI=THETA(6)*EXP(ETA(2))
	
	$DES
	DISPRG=INTRI + SLOPI*T
	EXPHAZ=EXP(BETA*LOCF + BET2*DISPRG)
	DADT(1)=EXPHAZ ; Observed + Unobserved h(t)
	DADT(2)=EXPHAZ ;Unobserved h(t)
	
	$ERROR
	CHST=BSHZ*A(1) ; Cum hazard overall
	CHSINT=BSHZ*A(2) ; Cum hazard from last obs
	CHSTM1=CHST-CHSINT ; Cum hazard upto last obs
	IF (DVID.EQ.1) THEN
	Y=INTRI + SLOPI*TIME + ERR(1) ; Biomarker
	ENDIF
	SRVT=EXP(-CHST) ; Survival at t
	IF (DVID.EQ.2.AND.DV.EQ.0) THEN
	Y=SRVT ; Like no event
	ENDIF
	IF (DVID.EQ.2.AND.DV.EQ.1) THEN
	SRVTM1=EXP(-CHSTM1) ; Survival at t lastobs
	Y=SRVTM1-SRVT ; Like event
	ENDIF
	
		NMV requires a more complex data structure and two differential equations if there is a random effect used to compute the hazard. The data table has to turn on the second integration compartment at the start of the interval during which the event occurs. The cumulative hazard up to the start of this interval is then computed by subtracting the hazard cumulated in the interval from the cumulative hazard at the end of the interval.

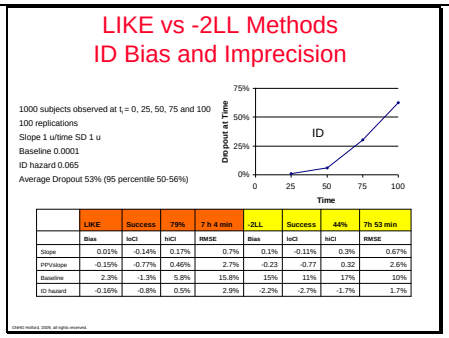
Slide 43	**Interval Censored Data (NMV)**  ID 2: Event is between time 25 and 50. Cumulative Hazard compartment is turned on at 25 (CMT=2). 	The NONMEM V dataset format is shown here. When DVID is 1 this means the DV observation is of the disease state (e.g. viral load). When DVID is 2 this means the DV observation is event status (0=censored event, 1=had event) The CMT data item is set to 2 at the time of the last observation prior to the event. This turns on this compartment so that it accumulates the hazard of the event. In this example the first subject did not have an event during the study and the final record has DV=0 to indicate this. The second subject had an event between time 25 and 50. A special event record (EVID=2) is used to set the CMT variable to 2 and turn on the second compartment. The final record for this subject indicates that they were known to have had the event by time 50. The TRT data item is 0 for placebo and 1 for active treatment. The LOCF data item is used when testing the random missingness model.

#ID	TRT	TIME	CMT	LOCF	DV	MDV	DVID	EVID	Comment
1	1	0	1	0	-0.6	0	1	0	Biomarker Obs
1	1	25	1	-0.6	28.1	0	1	0	Biomarker Obs
1	1	50	1	28.1	53.2	0	1	0	Biomarker Obs
1	1	75	1	53.2	81.8	0	1	0	Biomarker Obs
1	1	100	1	81.8	108.7	0	1	0	Biomarker Obs
1	**1**	**100**	**1**	**108.7**	**0**	**0**	**2**	**0**	**Censored Event**
2	0	0	1	0	0.1	0	1	0	Biomarker Obs
2	0	25	1	0.1	28.8	0	1	0	Last Non-event Obs
2	0	25	2	28.8	0	1	0	2	Turn On Cum Hazard
2	**0**	**50**	**1**	**28.8**	**1**	**0**	**2**	**0**	**End Event Interval**

Slide 44	**Likelihood for Count Response**
	 
	$ESTIM MAXEVAL=9990 METHOD=COND LAPLACE -2LL
	$ERROR
	COUNT=BaseCount + Beta*CP + ppv_event
	;Simulate count
	IF (ICALL.EQ.4) THEN
	T=0
	NEVENT=0
	DO WHILE (T.LT.1)
	CALL RANDOM (2,R)
	T=T-LOG(1-R)/COUNT
	IF (T.LT.1) NEVENT=NEVENT+1
	ENDDO
	DV=NEVENT
	ENDIF
	;Stirlings formula for log DV factorial
	IF (DV.GT.1) THEN
	LDVFAC=(DV+.5)*LOG(DV)-DV+.5*LOG(6.283185)
	ELSE
	LDVFAC=0
	ENDIF
	Y=-2*(-COUNT+DV*LOG(COUNT)-LDVFAC)
	
		Counts are a special kind of categorical response. The probability of a count can be predicted using the Poisson distribution.
		
		This NM-TRAN fragment shows how Stirling's formula is used to calculate the natural log of the factorial of the observed count. This is used in the last line to predict the probability of the observed count (DV) and the predicted count (COUNT).
		
		The ICALL.EQ.4 block shows how to simulate a count under the Poisson distribution (Frame et al 2003). The LOG(1-R) is required rather than the simpler LOG(R) because it is possible that R is 0 but it cannot be 1. Thus the LOG(1-R) code avoids an error caused by LOG(0).
		
		The theory why this algorithm works is mathematically complex. Mats Karlsson (nmusers 2009) described a simple view: "The code is simulating one event after another on a time interval standardized by lambda. You sample a survival probability, translate that into a time, check if it is beyond the standardized interval, if not increase N and add the event time to the elapsed time in the interval. "
		
		Frame B, Miller R, Lalonde RL. Evaluation of Mixture Modeling with Count Data using NONMEM. Journal of Pharmacokinetics and Pharmacodynamics. 2003;30(3):167-83.

Slide 45	**"Standard Survival Analysis" SAS, Splus etc** ● Two Part Hazard » Baseline Hazard function » Explanatory Variable function ● Parametric Baseline – Exponential – Weibull – Etc ● Non-parametric Baseline – Cox Proportional hazard 	Standard approaches to survival analysis implicitly divide the hazard into the baseline hazard and an explanatory variable hazard function. The Cox proportional hazard method is used to test differences of survival time distributions. It makes the assumption that the baseline hazard is the same in the two groups that are being tested but make no explicit assumption about the baseline hazard shape. The baseline hazard is therefore considered to be defined by a non-parametric assumption. In combination with a parametric explanatory variable function the overall procedure is called semi-parametric.

Slide 46	**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 $\ln(T_i) = \frac{1}{\gamma}\left[\ln(\lambda) - \beta_1 \cdot x_{1i} - \beta_2 \cdot x_{2i} - ... - \beta_p \cdot x_{pi} + \varepsilon_i\right]$ Or more generally $\ln(T_i) = \mu + \alpha_1 x_{1i} + \alpha_2 x_{2i} + ... + \alpha_p x_{pi} + \sigma \cdot \varepsilon_i$ Note that covariates (x1…xp) are usually assumed to be time invariant ***Standard survival analysis is equivalent to non-compartmental PK***	When covariates change with time then the hazard must be integrated in a piecewise fashion. This is exactly analogous to PK problems. If clearance changes from one time period to the next then the concentration prediction must be done piecewise (NONMEM describes this as 'advancing the solution')
Slide 47	**Combined Response Models** ● Continuous Response » Standard PKPD – Bone Mineral Density ● Non-continuous Response » Time to Event – Treatment Discontinuation – Bone fracture ● Combined 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.
Slide 48	**Randomization Test** Simulated Linear Disease Progress Plus CRD 1000 subjects observed at t = 0, 25, 50, 75 and 100 1000 replications; RD and ID: One extra parameter Bootstrap mean and 95% confidence interval Slope 1 u/time SD 1 u Baseline 0.01 ID hazard 0.0 Average Dropout 50% (95 percentile 47-53%) Type I error rate 5%  **Model Comparison** / **CritOBJ** / **Low** / **High** / **F success** Null: CRD Alternate: RD — 3.75 — 3.27 — 4.25 — 95% Null: CRD Alternate: ID — 3.67 — 3.33 — 4.17 — 95%	Hu and Sale investigated the three models for missingness using real data sets. They distinguished between the models by assuming the difference in -2 times the log likelihood was distributed according to the chi-square distribution. This assumption was tested by simulating data under the completely random dropout model and then fitting it with the same model and with the random and informative dropout models. The probability of being observed to have dropped out is shown using the CRD model in the graph. Overall about 50% of subjects are expected to drop-out during the time from 0 to 100. The addition of one extra parameter with either the RD or ID models required a large change in objective function to reject the null with a 5% Type I error rate. Over 90% of runs minimized successfully with both the null model (CRD) and the alternate model (RD or ID).

Slide 49		The influence of using the two different methods (CCONTR and -2LL) for computing the contribution to the likelihood was investigated by Monte Carlo simulation. The probability of being observed to have dropped out is shown using an informative dropout model in the graph. Overall about 50% of subjects are expected to drop-out during the time from 0 to 100. A disease progress observation was made on up to 4 occasions after entry to the study. During simulation the ID model was used to predict if drop-out occurred between the last visit and the current visit. A linear disease progress model with fixed intercept of 0 was used to describe the disease progress state. Simulated data were then fitted to the ID model using NONMEM. The CCONTR method was faster and slightly more NONMEM runs minimized successfully. The bias and imprecision of the parameter estimates was similar for both methods. There were major biases in the estimated of the baseline and informative dropout hazard parameters.