Slide 1

Time to Event Models
Practical Implementation

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

Slide 2

Encoding Single Events
Interval Starts with DV=3

ID	TIME	DV	MDV (NONMEM)	Comment
1	0	.	1	Start observing
1	50	1	0	Exact Time Event
2	0	.	1	Start observing
2	100	0	0	Censored Event
3	0	.	1	Start observing
3	55	3	1	Start Event Interval
3	70	2	0	End Event Interval

A record at time=0 is needed to define when the hazard integration starts.
Exact event has DV=1, Interval Censored event has DV=2
A DV value of 3 is used to signal the start of an interval containing an event.

Slide 3

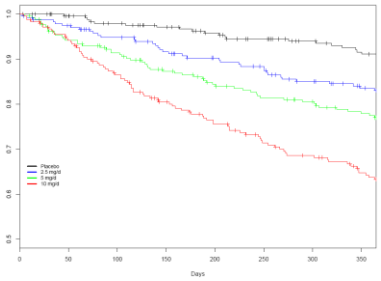
Single Event Time Varying Hazard (CP)
Interval starts with DV=3

```
$ESTIM MAXEVAL=9990 METHOD=COND
NSIG=3 SIGL=9
LAPLACE LIKE

$THETA
10  FIX    ; CL
100 FIX    ; V
(0,0.01)   ; BASE
0.1        ; BETACP

$OMEGA
0.1 FIX ; PPV_CL
0.1 FIX ; PPV_V

$SUBR ADVAN=6 TOL=9
$MODEL
   COMP=(CENTRAL)
   COMP=(CUMHAZ)

$PK
   IF (NEWIND.LE.1) THEN
      ; for interval censoring
      SURLAST=1
   ENDIF
   CL=THETA(1) *EXP(ETA(1))
   V=THETA(2) *EXP(ETA(3))
   BASHAZ=THETA(3)
   BETACP=THETA(4)
```

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

$ERROR
CP=A(1)/V
HAZ= BASHAZ*EXP(BETACP*CP)
CUMHAZ=A(2)
SUR=EXP(-CUMHAZ)

; Estimation Code
IF (DV.EQ.0) THEN ; Censored event
   Y=SUR ; Likelihood is the survivor function
ENDIF

IF (DV.EQ.1) THEN ; Exact Time event
   Y=SUR*HAZ ; likelihood of event in interval
ENDIF

IF (DV.EQ.2) THEN ; Interval Censored event
   Y=SURLAST-SUR ; likelihood of event in interval
ENDIF

IF (DV.EQ.3) THEN ; start of new event interval
   SURLAST=SUR ; survivor function at start of new
interval
ELSE ; save recursive random variable
   SURLAST=SURLAST ; survivor function at end of
previous interval
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. 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. Random effects on hazard model parameters (e.g. BASHAZ and BETACP) are not estimable with single events.

Slide 4	**Joint Model Data** **Interval start uses DV=3** | ID	TIME	TRT	DVID	DV	MDV	Comment | |---|---|---|---|---|---|---| | 1 | 0 | 0 | . | . | 1 | Start observing | | 1 | 20 | 0 | 1 | 67.4 | 0 | Biomarker | | 1 | 30 | 0 | 1 | 43.2 | 0 | Biomarker | | 1 | 50 | 0 | 2 | 1 | 0 | Exact Time Event | | 2 | 0 | 1 | . | . | 1 | Start observing | | 2 | 20 | 1 | 1 | 50.2 | 0 | Biomarker | | 2 | 30 | 1 | 2 | 3 | 1 | Start event interval | | 2 | 30 | 1 | 1 | 13.5 | 0 | Biomarker | | 2 | 50 | 1 | 2 | 2 | 0 | Interval Censored Event | 	The TRT data item indicates if the subject is receiving active treatment (TRT=1) or not (TRT=0). DVID is used to distinguish between continuous value biomarker observations (e.g. DVID=1 for drug concentration) and event observations (e.g. DVID=2). Start of an event interval is marked by DV=3 and DVID=2. Exact event has DV=1, Interval Censored event has DV=2
Slide 5	**Disease Progress and Time Varying Hazard** **Interval start uses DV=3** $INPUT ID TRT DVID TIME DV MDV $ESTIM MAX=9990 NSIG=3 SIGL=9 METHOD=CONDITIONAL LAPLACE $SUBR ADVAN=6 TOL=9 $MODEL COMP=(CUMHAZ) $PK IF (NEWIND.LE.1) THEN ; Initialize SURLAST=1 CUMLAST=0 ENDIF ;-------------------------------- ; Hazard BASHAZ = THETA(1) ; Baseline hazard BETADP = THETA(2) ; Disease progress effect ;-------------------------------- ; Symptomatic treatment effect EFFECT = TRT*THETA(3) ;-------------------------------- ;Disease Progress INTRI = (THETA(4)+ EFFECT)*EXP(ETA(1) SLOPI = THETA(5)* EXP(ETA(2) $DES DPRG = INTRI + SLOPI*T DADT(1) = BASHAZ*EXP(BETADP*DPRG) ; h(t) $ERROR CUMHAZ=A(1) ; Cumulative hazard CUMDIF=CUMHAZ-CUMLAST ; CUMHAZ since last event SUR=EXP(-CUMDIF) DISPRG=INTRI + SLOPI*TIME ;-------------------------------- IF (DVID.EQ.1) THEN ; disease progress F_FLAG = 0 ; Continuous Y = DISPRG + ERR(1); Disease Progress ENDIF ;-------------------------------- IF (DVID.EQ.2.AND.DV.EQ.0) THEN ; right censored F_FLAG = 1 ; Likelihood Y = EXP(-CUMHAZ) ENDIF ;-------------------------------- IF (DVID.EQ.2.AND.DV.EQ.1) THEN ; exact time F_FLAG = 1 ; Likelihood HAZARD = BASHAZ*EXP(BETADP*DISPRG) Y = SUR*HAZARD ENDIF ;-------------------------------- IF (DVID.EQ.2.AND.DV.EQ.2) THEN ; interval censored F_FLAG = 1 ; Likelihood Y = SURLAST - SUR SURLAST=SUR ENDIF IF (DVID.EQ.2.AND.DV.NE.3) THEN ; save recursive random variable SURLAST=SURLAST 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. 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(Likelihood).
Slide 6	**Time to Event Analysis** **Diagnostic Plots** 	

Slide 7	**Outline** - The Kaplan-Meier Plot - Simulating time to event - Visual predictive checks 	
Slide 8	**Kaplan-Meier Plot** - Derived directly from events - Non-parametric estimate of survivor function - Incorporates censored events  $\hat{S}(t) = \prod_{j=1}^{k} \left(\frac{n_j - d_j}{n_j} \right)$ n_j=alive in interval, d_j=die in interval 	Computation of KM estimates can be tricky e.g. how to handle tied events (more than one event at the same time). It is best done using a well written procedure in a statistical package. See Collett, D. (2003). Modelling survival data in medical research. Boca Raton, CRC Press for details. The following example shows how to generate the KM plot shown above and how to add predictions of the survivor function to the KM plot.
Slide 9	**R Code for Kaplan-Meier** ``` # Read NONMEM data file input= "warf_sim_p1_TRT" warf=read.csv(paste(input,".csv",sep=""),stringsAsFactors=F) names(warf)[1]="ID" # Make a variable called ID # Select columns to use event = subset(warf, MDV==0, select=c("ID","TIME","DV","TRT")) event[,"DV"] = as.numeric(event[,"DV"]) # Run survfit() with TRT as the covariate fit1 = survfit(Surv(TIME,DV)~TRT, data=event, conf.type="none") ``` 	Factors are a special kind of data type that can cause problems for R functions expecting numeric values. By default R will treat the DV column as a *factor* because the data file contains "." for DV when MDV=1. The stringsAsFactors=F option is used in read.csv() but this means all values in the DV column will be converted to character data type. Then names() function is used to rename the first variable in the warf data frame from "_ID" to "ID". By default it will have the name "_ID" because the CSV file header for the first column is "#ID" and R converts the "#" to "_". After extraction of records with events (MDV=0) using the subset() function it is necessary to convert the character values of DV to numeric values for survfit() to work properly. The automatic conversion to factor is one of the 'features' of R that often requires the kind of workaround shown here. survfit() computes the non-parametric Kaplan-Meier estimates of the survivor function. The formula for the survivor function includes "~TRT" which means TRT is a covariate distinguishing different groups so that a separate set of KM estimates is produced for each treatment.

Slide 10

R Code for T2E Plots

```
# Create plot using output from survfit()

   minProb=0.6; maxTime=365
   xscale = c(0,maxTime); yscale = c(minProb,1)
   kmcolors=c("black","blue","green","red")

   kmplot=plot(fit1,
      col=kmcolors, cex.main=.95,
      ylim=yscale, xlim=xscale,
      xlab="Days", ylab=paste("Prob of No Event",input)
   )
   legend(
       legend = c("Placebo", "2.5 mg/d","5 mg/d","10 mg/d"),
       lty=c(1,1,1,1),lwd=c(2,2,2,2), col=kmcolors, bty='n', cex=.8,
       x=0.01*maxTime,y=minProb+(1-minProb)*0.35
   )

# Add survivor function observation
   pred=read.csv("surv.csv")
   lines(pred[,"time"], pred[,"mean"])
```


The plot() function recognizes that fit1 is an object created by survfit() and knows how to extract the data in order to create a line for each survivor function and to mark censored events.
The legend() function is used to add an informative legend to describe the different treatments associated with each survivor function curve.
The predicted survivor function values from WinBugs are read using read.csv() and overlaid on the KM plot using the lines() function.

Slide 11

Simulating Time to Event

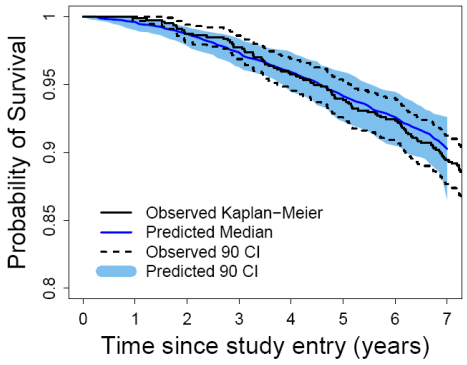
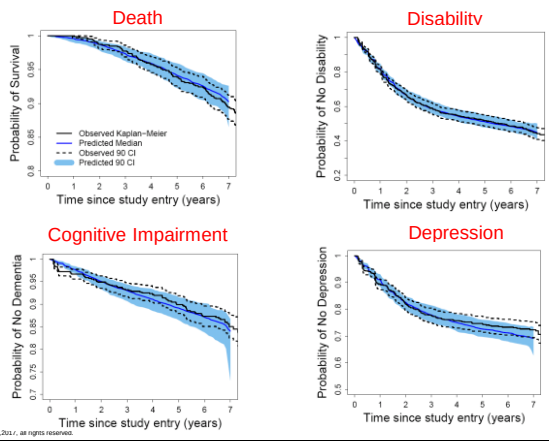
- There is no simple solution to simulate an exact event time
- A general solution for interval censored event times exists
  - » Requires pre-specifying the intervals e.g. every day for 1 year
- Simulation in NONMEM is complex. The following code snippets show some of the key features.

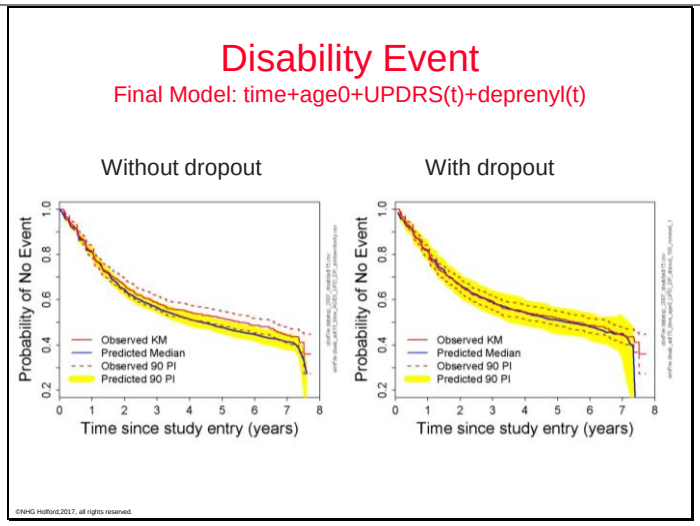
Slide 12

NONMEM Data Template for Simulation of Warfarin

#ID	TIME	TRT	AMT	RATE	ADDL	II	WT	DVID	DV	MDV
1	0	1	5	-2	364	1	70	0	.	1
1	0	1	.	.	.	.	70	1	.	1
1	0	1	.	.	.	.	70	2	.	1
1	0	1	.	.	.	.	70	3	.	1
1	1	1	.	.	.	.	70	1	.	1
1	1	1	.	.	.	.	70	2	.	1
1	1	1	.	.	.	.	70	3	.	1

Data template can be used to simulate warfarin PK (AMT, RATE, ADDL, II) and changes in prothrombin complex activity (PCA; DVID=1) and international normalised ratio (INR, DVID=2) as well as interval censored time to event (DVID=3). Between subject variability in warfarin PK can be simulated in part by simulating different weights.

Slide 13	**Visual Predictive Check of Time to Event** - Simulate many data sets e.g. 100 - Calculate the KM survivor function estimates at each time for each data set - Interpolate the KM survivor function at frequent intervals e.g. every day - Calculate median and prediction intervals from the set of interpolated KM survivor functions 	
Slide 14	**Time to Event VPC**  	
Slide 15	**Evaluation of Hazard Models visual predictive check**  36	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.

Slide 16		Time to event model for disability in Parkinson's disease (from Vu, T. C., J. G. Nutt, et al. (2009). "Disease progress and response to treatment as predictors of survival, disability, cognitive impairment, and depression in Parkinson's disease." In Preparation.) The VPC shows the need to include the dropout model (due to death or non-death random censoring) in order to match the disability time to event model predictions with the observed KM survivor function estimates.