Slide 1

Modelling Likelihoods using NONMEM

Nick Holford
University of Auckland

Slide 2

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 3

-2 Log Likelihood

Continuous Variable

$$ELS = \sum_{i=1}^{NSUB}\sum_{j=1}^{NOBS}\left(\frac{(Y_{OBS,ij} - Y_{PRED,ij})^2}{Var_{ij}} + \ln(Var_{ij})\right)$$

$$-2LL_{ij} = \frac{(Y_{OBS,ij} - Y_{PRED,ij})^2}{Var_{ij}} + \ln(Var_{ij})$$

$$-2LL_{ij} = \left(\frac{Y_{OBS,ij} - Y_{PRED,ij}}{SD_{ij}}\right)^2 + 2\cdot\ln(D_{ij})$$

Likelihood = Probability

$$-2LL_{i,j} = -2\cdot\ln(\Pr_{i,j})$$

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/2*Ln(2*Pi)) but this is not needed in order to minimize -2LL and obtain maximum likelihood estimates.

The likelihood of an observation is also the probability of an observation so if we have a predicted probability it can be easily converted to the -2LL equivalent.

Slide 4

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.

Slide 5

Standard Method

```
$PROB theophylline pharmacodynamics
;standard control stream
$DATA theopd.dat IGNORE #
$INPUT ID TIME THEO AGE WT
SEX RACE DIAG DV
$ESTIM METHOD=COND LAPLACIAN
$COV

$THETA (0,150.,)  ; POP_E0
$THETA (0,200.,)  ; POP_EMAX
$THETA (.001,10,) ; POP_EC50

$OMEGA 0.5  ; PPV_E0
$OMEGA 0.5  ; PPV_EMAX
$OMEGA 0.5  ; PPV_EC50

$SIGMA 100   ; RUV_SD
```

```
$PRED

E0=POP_E0*EXP(PPV_E0)
EMAX=POP_EMAX*EXP(PPV_EMAX)
EC50=POP_EC50*EXP(PPV_EC50)

EFFECT = E0 + EMAX*THEO/(THEO+EC50)

Y=EFFECT + RUV_SD
```


Here is an example of part of an NM-TRAN control stream which shows the standard method that NONMEM uses to compute the likelihood. The prediction is returned in the variable Y and NONMEM will figure out from Y and the DV and the RUV_SD (e.g. EPS(1)) how to compute the ELS contribution to -2LL.

Slide 6

-2LL Example

```
$PROB theophylline pharmacodynamics
;-2LL option using SD
$DATA theopd.dat IGNORE #
$INPUT ID TIME THEO AGE WT
SEX RACE DIAG DV
$ESTIM METHOD=COND LAPLACIAN -2LL
$COV

$THETA (0,150.,)  ; POP_E0
$THETA (0,200.,)  ; POP_EMAX
$THETA (.001,10,) ; POP_EC50

$OMEGA 0.5  ; PPV_E0
$OMEGA 0.5  ; PPV_EMAX
$OMEGA 0.5  ; PPV_EC50

$THETA 10   ; RUV_SD
```

```
$PRED

E0=POP_E0*EXP(PPV_E0)
EMAX=POP_EMAX*EXP(PPV_EMAX)
EC50=POP_EC50*EXP(PPV_EC50)

EFFECT = E0 + EMAX*THEO/(THEO+EC50)

Y=((DV-EFFECT)/RUV_SD)**2 + 2*LOG(RUV_SD)
```


Here is one way to compute -2LL explicitly and return the -2LL value in the variable Y. The $ESTIMATION option -2LL tells NONMEM not to try to compute the -2LL contribution internally.

Slide 7

Alternatives

- -2LL (variance)

```
$ESTIM -2LL
$THETA (1,250) ; RUV_VAR
M2LL=(DV-EFFECT)**2/RUV_VAR + LOG(RUV_VAR)
Y=-0.5*M2LL
```

- LIKE (standard deviation)

```
$ESTIM LIKE
$THETA (1,10) ; RUV_SD
M2LL=((DV-EFFECT)/RUV_SD)**2 + 2*LOG(RUV_SD)
Y=EXP(-0.5*M2LL)
```

- LIKE (variance)

```
$ESTIM LIKE
$THETA (1,250) ; RUV_VAR
M2LL=(DV-EFFECT)**2/RUV_VAR + LOG(RUV_VAR)
Y=EXP(-0.5*M2LL)
```


Other ways to compute -2LL or the LIKELIHOOD are shown here. Note that results of these different methods may be not be the same because of small numerical differences in the way the calculations are performed. There is no right answer.

Slide 8

Logistic Model

$$P = \frac{e^{\beta_0 + \beta_1 \cdot X}}{1 + e^{\beta_0 + \beta_1 \cdot X}}$$

$$Logit = \beta_0 + \beta_1 \cdot X$$

The logistic model is a trick for predicting probabilities which must lie between 0 and 1. There are other transformations that might be used e.g. see http://www.page-meeting.org/page/page2005/PAGE2005O16.pdf (talk by Adrian Dunne at PAGE in Pamplona). We can right a model using any form we like in the logistic domain. The value obtained is called the logit. It is transformed back to a probability (i.e. likelihood) and this value can then be used by NONMEM.

Slide 9

Binary Response

```
$ESTIMATION METHOD=CONDITIONAL
LAPLACIAN LIKELIHOOD
$THETA
(0,0.3,1) ; BASEP
0.1 ; BETA
$OMEGA .5 ; PPV_EVENT

$ERROR

  BASE=LOG(BASEP/(1-BASEP))
  LGST=BASE + BETA*CP + PPV_EVENT

  P1=1/(1+EXP(-LGST))

  IF (DV.EQ.1) Y=P1
  IF (DV.EQ.0) Y=1-P1
```


The simplest kind of non-continuous response is binary. This is a fragment of NM-TRAN code that implements the logistic transformation. BASEP is the baseline probability of the response with a DV value of 1. The probability of a DV of 0 is 1 minus this probability.

Slide 10

Ordered Categorical Response

```
; Effect of drug
  Edrug=Beta*Cp

; Compute cumulative logit
  B0=Lgt0
  B1=B0 + Lgt1
  B2=B1 + Lgt2

; A0 - A2 are the cumulative logits with
subject-specific random effects
  A0=B0 + Edrug + ppv_event
  A1=B1 + Edrug + ppv_event
  A2=B2 + Edrug + ppv_event
```

```
; P0 - P2 are the cumulative probabilities
; p0=P(Y<=0), p1=P(Y<=1) =p0+p1,
; p2=P(Y<=2) =p0+p1+p2

  P0=1/(1+EXP(-A0)
  P1=1/(1+EXP(-A1)
  P2=1/(1+EXP(-A2)

; exit if the probabilities are
; not ordered correctly

   IF (P0.GE.P1) EXIT 1 101
   IF (P1.GE.P2) EXIT 1 102

   IF (DV.EQ.0) Y=P0
   IF (DV.EQ.1) Y=P1-P0
   IF (DV.EQ.2) Y=P2-P1
   IF (DV.EQ.3) Y=1-P2
```


Ordered categorical responses are expressed as the cumulative logit for each level of response. In this case there are 4 categories of response. The trick is to compute the cumulative logit for each category then compute the individual contributions to the cumulative probability at the end in order to predict the probability of the observed DV value.

Slide 11

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 12

Hazard, Risk, Survival

$$hazard = \beta_0 \cdot e^{\beta \cdot X} \qquad Risk_i = \int_0^{t_i} hazard$$

$$Survival_i = e^{-Risk_i}$$

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 risk of the event. As time passes the cumulative hazard predicts the risk of having the event over the interval 0-t.
The hazard model shown here is a proportional hazard models which is widely used for survival analysis. Beta0 is the baseline hazard. Beta is a parameter describing how the hazard depends on some other value, X.
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.

Slide 13

Survival in a Bathtub

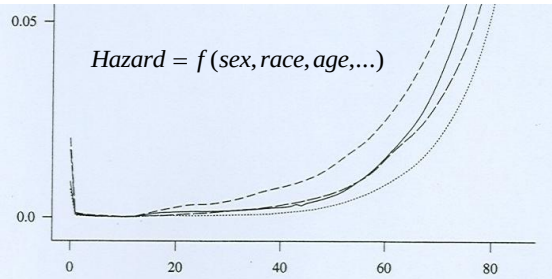


Figure 2.6 *Hazard functions for all cause mortality for the US population in 1989. White males (———); white females (·····); black males (- - - - -); black females (— — —).*

"… 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.

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 14

Right Censored Time to Event

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

$$Like_{NoEvent} = e^{-\int_{t=0}^{t=T} h(t)}$$

ELSE; event at T

$$Like_{Event} = e^{-\int_{t=0}^{t=T} h(t)} \cdot h(T)$$

Using NONMEM nomenclature, DV is a the observed drop-out state. If it is 0 it means the subject has not dropped out. If it is 1 then the subject has dropped out at some time between the last observation when they had not dropped out and the current time when it is known they have dropped out.
The probability of a drop-out event is conditional on whether drop-out has not occurred at the time of the probability prediction.
This model shows the probability of not dropping out depends on the hazard accumulated from time 0 until the time of the current observation given that the subject has not dropped out (DV=0). In this example the hazard uses the predicted disease state value rather than the observed disease state value.
At the time a subject is known to have dropped out (DV=1) the probability of this happening is the product of the probability they the subject had not

dropped out at the time of the last observation (when it was known they had not dropped out) and the probability of dropping out at some time between the last observation and the current time when it known the subject has dropped out.

Slide 15

Right Censored Time to Event

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

$SUBR ADVAN=6 TOL=3
$MODEL
   COMP=(CENTRAL)
   COMP=(RISK)
$PK
   CL=POP_CL
   V=POP_V
   BETA=BETA
   BASHAZ=BASE*EXP(PPV_HAZ)
```

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

$ERROR
  CP=A(1)/V
  CUMHAZ=A(2)
  PSURV=EXP(-CUMHAZ)

  IF (DV.EQ.0) THEN
     Y=PSURV ; censored event
  ELSE
     HAZNOW=BASHAZ*EXP(BETA*CP)
     Y=PSURV*HAZNOW ; Prob 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 (PSURV) i.e. the probability of not having the event at a particular time (DV=0). If the event occurs (DV=1) at the time of the record then the probability of having it is the product of the probability of not having it times the instantaneous risk.

Note that this code cannot be used for simulating time to event. Simulation of time to event is a tricky business with NONMEM.

Slide 16

Right Censored Data Single Event

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

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

Slide 17

Right Censored Data Repeated Events

#ID	TIME	DV	MDV	EVID	CMT	AMT	Comment
1	0	0	1	1	1	100	Dose
1	50	1	0	0	2	0	First Event
1	50	0	1	2	-2	0	Reset Cum Hazard
1	50	0	1	2	2	0	Turn on
1	75	1	0	0	2	0	Second Event
1	75	0	1	2	-2	0	Reset Cum Hazard
1	75	0	1	2	2	0	Turn on
1	100	0	0	0	2	0	Censored Event

Slide 18

Joint Models

- Basic concept
  Compute LIKE (or -2LL) for ANY response
- All types of response can be combined
- F_FLAG (NMVI) or user written CCONTR (NMV) allows combining continuous responses with non-continuous responses

Any kind of response, continuous or non-continuous, can be combined in a NONMEM model by using the joint likelihood computed for each observation. A user written CCONTR is used to tell NONMEM how to compute the likelihood (or -2LL) for continuous responses. It is up to the user to compute the likelihood (or -2LL) for any non-continuous responses.

Slide 19

Joint Model NMVI

```
$DATA ID TIME AMT TYPE DV
$ESTIM MAXEVAL=9990 METHOD=COND LAPLACE
NOABORT

$SUBROUTINE ADVAN=2 TRAN=2

$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
   F_FLAG=0 ; ELS
   Y=CPEST ; continuous
ENDIF
IF (TYPE.EQ.3) THEN
   F_FLAG=1; likelihood
   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 20

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 21

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 22

Interval Censored Time to Event

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

$$Like_{NoEvent} = S(T) = e^{-\int_{0}^{t=T} h(t)}$$

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 23

Hazard Models for Interval Censored Events

Completely Random

$$h(t) = \beta_0$$

Random

$$h(t) = \beta_0 \cdot \exp(\beta_{RD} \cdot Yo_{lastobs}) \text{ from } t_{lastobs} \text{ to } t_{end}$$

Informative

$$h(t) = \beta_0 \cdot \exp(\beta_{ID} \cdot Yu(t)) \text{ from } t_{lastobs} \text{ to } t_{end}$$

T = unobserved event time

Yo = observed outcome at $t_{lastobs}$ [LOCF]

Yu = unobserved outcome at t [Model prediction]

Hu C and Sale M. JPKPD 2003;30(1):83-103

Completely random event is modeled as a constant hazard over the course of the study. For this type of event the probability of the event is not influenced by any other factor such as disease progress status.
For the random event model, if the unobserved event time occurs within the interval ti-1 to ti, then hazard for event is constant only within that interval and hazard rates depends only on the last observed status at ti-1. The disease status during the interval is imputed using the last observation carried forward (LOCF) method.
For the informative event model, within the interval ti-1 to ti, the hazard is NOT constant but depends on the unobserved disease progression during that interval and it may or may not depend on the last observed disease status.

Slide 24

Time Varying Hazard with Interval Censoring (NMVI)

```
$INPUT ID TRT TIME LOCF DV DVID

$ESTIM MAX=9990 SIG=6 NOABORT
METHOD=CONDITIONAL LAPLACE

$SUBR ADVAN=6 TOL=6

$MODEL
COMP=(CUMHAZ)

$PK
IF (NEWIND.LE.1) SRVTM1=1 ; initial S(t)
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)
$ERROR
CHZT=BSHZ*A(1) ; Cum hazard overall

IF (DVID.EQ.1.OR.DVID.EQ.3) THEN
   Y=INTRI + SLOPI*TIME + ERR(1); Biomarker
ENDIF
SRVT=EXP(-CHZT) ; 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
   Y=SRVTM1-SRVT ; Like event
ENDIF
IF (DVID.EQ.3) THEN ; last obs before event
   SRVTM1=SRVT ; remember S(t) for next record
ELSE
   SRVTM1=SRVTM1 ; save random variable
ENDIF
```


This illustrates joint modelling for all 3 types of missingness mechanism (completely random, random and informative).
The CRD model is coded by fixing THETA(2) and THETA(3) to zero ($THETA, $OMEGA and $SIGMA records are not shown here)
The RD model is coded by fixing THETA(2) to zero and estimating THETA(1). It uses the LOCF value to predict the hazard during the unobserved period after the last known observation until drop-out is known.
The ID model is coded by fixing THETA(1) to zero. It uses the predicted disease status to predict the hazard.
A differential equation are used to integrate the hazard.
An effect of treatment is assumed to affect the intercept of the disease progress model.

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. The assignment of SRVT to SRVTM1 must to be done only at the time of the last observed non-event time which is signaled by a DVID value of 3.

Note that records with DVID of 1 and DVID of 3 may be valid observations to predict the biomarker.

Slide 25

Interval Censored Data (NMVI)

#ID	TRT	TIME	LOCF	DV	DVID	Comment
1	1	0	0	-0.6	1	Biomarker Obs
1	1	25	-0.6	28.1	1	Biomarker Obs
1	1	50	28.1	53.2	1	Biomarker Obs
1	1	75	53.2	81.8	1	Biomarker Obs
1	1	100	81.8	108.7	1	Biomarker Obs
1	**1**	**100**	**108.7**	**0**	**2**	**Censored Event**
2	0	0	0	0.1	1	Biomarker Obs
2	0	25	0.1	28.8	3	Last Non-event Obs
2	**0**	**50**	**28.8**	**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 26

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=(HZLAST,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
CHZT=BSHZ*A(1) ; Cum hazard overall
CHZINT=BSHZ*A(2) ; Cum hazard from last obs
CHZTM1=CHZT-CHZINT ; Cum hazard upto last obs
IF (DVID.EQ.1) THEN
   Y=INTRI + SLOPI*TIME + ERR(1); Biomarker
ENDIF
SRVT=EXP(-CHZT) ; 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(-CHZTM1) ; Survival at t lastobs
   Y=SRVTM1-SRVT       ; Like event
ENDIF
```


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.
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 27

Interval Censored Data (NMV)

#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**

ID 2: Event is between time 25 and 50.
Cumulative Hazard compartment is turned on at 25 (CMT=2).

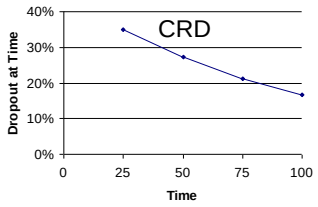
The NONMEM V dataset format is shown here.
When DVID is 1 this means the DV observation is of the disease state (e.g. viral load).
When DVID is 2 this means the DV observation is event status (0=censored event, 1=had event)
The CMT data item is set to 2 at the time of the last observation prior to the 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.

Slide 28

Randomization Test

Simulated Linear Disease Progress Plus CRD

1000 subjects observed at t_i = 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%



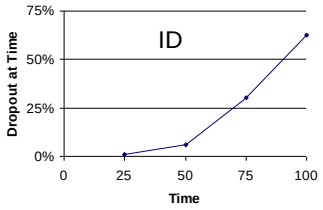
Model Comparison	CritOBJ	Low	High	F success
Null: CRD Alternate: RD	3.75	3.27	4.25	95%
Null: CRD Alternate: ID	3.67	3.33	4.17	95%

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 29

LIKE vs -2LL Methods
ID Bias and Imprecision

1000 subjects observed at t_i = 0, 25, 50, 75 and 100
100 replications
Slope 1 u/time SD 1 u
Baseline 0.0001
ID hazard 0.065
Average Dropout 53% (95 percentile 50-56%)



	LIKE	Success	79%	7 h 4 min	-2LL	Success	44%	7h 53 min
	Bias	loCI	hiCI	RMSE	Bias	loCI	hiCI	RMSE
Slope	0.01%	-0.14%	0.17%	0.7%	0.1%	-0.11%	0.3%	0.67%
PPVslope	-0.15%	-0.77%	0.46%	2.7%	-0.23	-0.77	0.32	2.6%
Baseline	2.3%	-1.3%	5.8%	15.8%	15%	11%	17%	10%
ID hazard	-0.16%	-0.8%	0.5%	2.9%	-2.2%	-2.7%	-1.7%	1.7%

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.