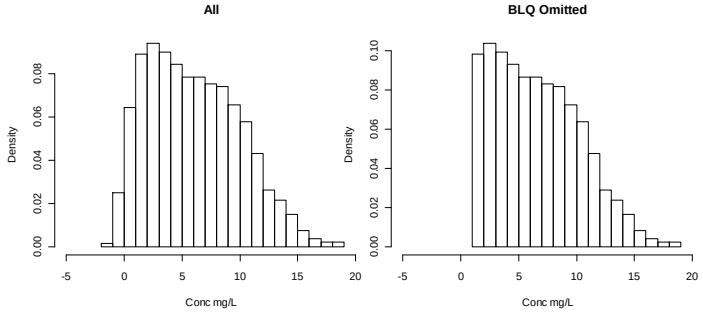
Slide 1	# Missing Data – Left Censoring The BLQ Problem Nick Holford University of Auckland Auckland, New Zealand	
Slide 2	## The Problem • Chemical analyst defines Lower Limit of Quantitation (LLOQ) as part of bioanalyical method validation • If measured conc is less than LLOQ the value is reported as Below Limit of Quantitation (BLQ) • BLQ is not a number! 	
Slide 3	## The Cause of the Problem? Guidance for Industry **Bioanalytical Method Validation** 1. *Lower Limit of Quantification (LLOQ)* – The lowest standard on the calibration curve should be accepted as the limit of quantification if the following conditions are met: – The analyte response at the LLOQ should be at least 5 times the response compared to blank response. – Analyte peak (response) should be identifiable, discrete, and reproducible with a precision of 20% and accuracy of 80-120%. FDA. Bioanalytical Method Validation http://www.fda.gov/cder/guidance/4252fnl.htm. 2001. 	FDA. Bioanalytical Method Validation http://www.fda.gov/cder/guidance/4252fnl.htm. 2001

Slide 4	**Missing Data in the FDA Guidance** • Does not say what values should be reported if BLQ! • This missing data is interpreted to mean that the measured value should not be reported 	
Slide 5	**Left Censoring LLOQ=1 mg/L**  	Missing low concentrations are referred to as 'left censored' because the left hand part of the distribution of concentrations is missing.
Slide 6	**Methods For Dealing with Censored Concentrations** • Ignore – treat as Missing Data Value – MDV Treat BLQ measured conc as missing • Estimate Likelihood – YLO Likelihood that all measured concs > BLQ are greater than LLOQ – M3 Likelihood that BLQ conc is less than LLOQ – M4 Likelihood that BLQ conc is less than LLOQ and greater than zero • Impute Value – Replace BLQ with LLOQ/2 – Replace BLQ with zero 	These three methods of dealing with missing data (Ignore, Estimate Likelihood and Impute) are general methods applied to all kinds of missing data problems.

Slide 7

The Key Ideas

Ways to Fit a PK Model with Some Data Below the Quantification Limit

Stuart L. Beal[1,2]

Received April 24, 2001—Accepted July 13, 2001

- Tough statistics
- Methods have numbers not names
  - i.e. M1 to M7

Beal SL. Ways to fit a PK model with some data below the quantification limit. Journal of Pharmacokinetics & Pharmacodynamics. 2001;28(5):481-504.

This work by Stuart Beal (a close colleague of Lewis Sheiner) was probably stimulated by this PharmPK thread http://gaps.cpb.ouhsc.edu/nm/99may2097.html

>From lewis@c255.ucsf.edu Tue May 20 16:32:53 1997 Subject: Re: Concentration values below assay limits

I have the feeling this has been discussed before... Briefly, BQL values can convey information, especially if they occur rather isolated in time (that is, there are no above QL levels "nearby"), and should then be included in the data analysis. While the very best way to do so is not yet known, here is a way that works pretty well:

- Use an error model that has an additive and proportional component; e.g., Y = F + F*EPS(1) + EPS(2)
- Fix var(eps(2)) to .25*QL**2. Say QL is .5 mcg/ml, then use, e.g., $SIGMA .04 .0625 FIX
- Record BQL values as QL/2; i.e., for the above example as .25.

>From n.holford@auckland.ac.nz Tue May 20 16:57:19 1997 Subject: Re: Concentration values below assay limits

Thanks for the recipe. But how about the rationale?

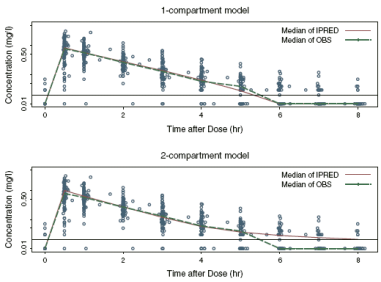
>From lewis@c255.ucsf.edu Tue May 20 17:11:08 1997 Subject: Re: Concentration values below assay limits

Basically, yes, although the formulas are actually: 1. var(eps(2)) = (f*QL)**2, where f can be taklen to be .5, but see below. 2. Record BQL values as QL/2

The rationale is

- the CV at QL is about f*100% (many define QL to be at f=.2; f=.5 gives BQL observations even less weight and is conservative in that sense).
- If all we know is that an observation is BQL, a reasonable ignorance assumption is that the true value is drawn from a uniform distribution with support on 0 - QL; a reasonable imputed value is then the mean of this distribution, viz., QL/2.

-LBS.

Slide 8	**Beal's Methods** • M1 = Ignore missing values 'MDV' • M2 = Likelihood assumes all values are censored at LLOQ 'YLO' • M3 = Estimate likelihood at times measurements are BLQ • M4 = Like M3 but also assume measurements are >=0 • M5 = Replace *all* BLQ with LLOQ/2 • M6 = Replace *first* BLQ with LLOQ/2, ignore others • M7 = Replace *all* BLQ with zero 	Beal wrote "A fourth method (M4) is like M3. However, it includes an adjustment to recognize that a measurement cannot really be negative." However, this is not correct. Concentrations cannot be negative but measurements (with additive error) can be negative.
Slide 9	**Simple Example** **Impact of censoring data below an arbitrary quantification limit on structural model misspecification** Wonkyung Byon · Courtney V. Fletcher · Richard C. Brundage • Used M2 ("YLO") method in NONMEM VI with simulations of one compartment model • "This simulation study has shown that the practice of assigning a LLOQ during analytical methods development, although well intentioned, can lead to incorrect decisions regarding the structure of the pharmacokinetic model. " • "The standard operating procedures in analytical laboratories should be adjusted to provide a quantitative value for all samples assayed in the drug development setting where sophisticated modeling may occur. However, the current level of precision may need to be maintained when laboratory results are to be used for direct patient care in a clinical setting." Byon W, Fletcher CV, Brundage RC. Impact of censoring data below an arbitrary quantification limit on structural model misspecification. J Pharmacokinet Pharmacodyn. 2008;35(1):101-16. 	
Slide 10	**Model Selection Error**  **Fig. 1** Plot of observed concentrations (*open circles*) versus time after dose: median of IPRED (*solid line*) and median of OBS (*dashed line*) were plotted at each nominal time point for the one-compartment model (upper panel) and two-compartment model (lower panel). The horizontal line represents the LLOQ of ZDV (0.02 mg/l). All concentrations below the LLOQ (both censored and simulated IPRED) were assigned a value of one-half LLOQ (0.01 mg/l) for the graphical presentation 1 cpt model used for simulation 1 and 2 cpt model used for estimation Without YLO the 2 cpt model was incorrectly selected frequently but with YLO the Type I error was close to nominal 5% Byon W, Fletcher CV, Brundage RC. Impact of censoring data below an arbitrary quantification limit on structural model misspecification. J Pharmacokinet Pharmacodyn. 2008;35(1):101-16. 	

Slide 11

Fancier Methods

Likelihood based approaches to handling data below the quantification limit using NONMEM VI

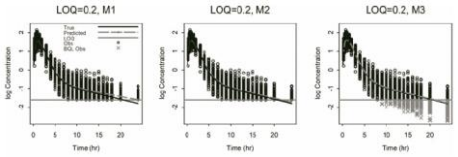
Jae Eun Ahn · Mats O. Karlsson · Adrian Dunne · Thomas M. Ludden

- Used M1 ("MDV"), M2 ("YLO") , M3 and M4 with 2 compartment model and first order absorption.
- "For the data simulated with a proportional error model, the overall performance was best for M3 followed by M2 and M1"
- "M3 and M4 resulted in similar estimates in analyses without log transformation."

Ahn JE, Karlsson MO, Dunne A, Ludden TM. Likelihood based approaches to handling data below the quantification limit using NONMEM VI. J Pharmacokinet Pharmacodyn. 2008;35(4):401-21.

Slide 12

Which Method?



"If the times of the BQL observations are not available, then adjusting the likelihoods for the remaining data (M2) may improve parameter estimates compared to no adjustment (M1). M2 estimates were not best but quite comparable to those of more refined methods like M3 and/or M4. This method can also be easily implemented using YLO feature in NONMEM VI."

"In most cases, the time when the BQL observations occur are recorded and a better approach for handling BQL data would be treating them as censored observation (M3)."

Ahn JE, Karlsson MO, Dunne A, Ludden TM. Likelihood based approaches to handling data below the quantification limit using NONMEM VI. J Pharmacokinet Pharmacodyn. 2008;35(4):401-21.

Slide 13

Simulation Code

```
$THETA
(0.01,0.15,1) ; POPCL L/H/70KG
(0.01,8,20) ; POPV L/70KG
(0.01,0.5,24) ; POPTABS H
(0.01,0.75,24) ; POPLAG H
(0,0.1,) ; RUVCV
(0,0.5,) ; RUVSD MG/L

$OMEGA
0.1 ; PPVCL
0.02 ; PPVV
0.5 ; PPVTABS
0.5 ; PPVLAG

$SIGMA 1 FIX ; EPS1

$PK
IF (NEWIND.EQ.0) THEN
   LLOQ=1 ; MG/L
ENDIF
```

```
$ERROR

   CP=A(2)/V
   PROP=CP*RUVCV
   ADD=RUVSD
   SD=SQRT(PROP*PROP+ADD*ADD)
   Y=CP + SD*EPS1

   IF (ICALL.EQ.4) THEN
      IF (AMT.GT.0.OR.Y.LT.LLOQ) THEN
         MDVX=1
      ELSE
         MDVX=0
      ENDIF
   ENDIF

$TABLE ID TIME WT AGE SEX DVID DV
MDVX EVID CMT AMT
```


Residual error model is parameterized using THETA and a SIGMA FIXED to 1. This parameterization is often convenient especially because the estimates describe standard deviation like quantities rather than variance. It is also used to implement the M3 and M4 censoring methods.

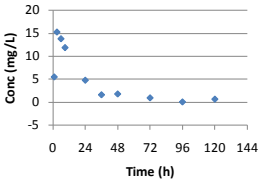
A new variable MDVX is created during the simulation process (when ICALL is 4). This variable is set to 1 for dose records (AMT>0) and when the predicted observation is less than LLOQ. Otherwise MDVX is set to 0. The MDVX variable is saved in the table file in the place of MDV.

Note that attempts to change MDV directly when there is more than one simulation subproblem may lead to undesired behaviour because each sub-problem will get the set of simulated MDV values from the previous problem.

Slide 14

Simulated Data

#ID	TIME	DV	MDVX	AMT
16	0	0	1	85
16	1	5.4486	0	0
16	3	15.252	0	0
16	6	13.783	0	0
16	9	11.852	0	0
16	24	4.7359	0	0
16	36	1.5538	0	0
16	48	1.7341	0	0
16	72	0.88145	1	0
16	96	-0.00852	1	0
16	120	0.60127	1	0



Note: Negative simulated concentration at 96 h because of additive residual error component

Slide 15

Estimation Code - MDV

```
$INPUT ID TIME WT AGE SEX DVID DV MDV EVID CMT AMT

$DATA ..\ka11_blq1_simln.reg\ka11_blq1_simln.fit
IGNORE=@ ; Ignore NONMEM table headers

; or

$DATA ka11_blq1_simln.csv ; CSV file created from NONMEM table


$EST METHOD=COND INTER NOABORT
MAX=9990 NSIG=3 SIGL=9
```

```
$ERROR

   CP=A(2)/V
   PROP=CP*RUVCV
   ADD=RUVSD
   SD=SQRT(PROP*PROP+ADD*ADD)
   Y=CP + SD*EPS1
```


Residual error model is parameterized using THETA and a SIGMA FIXED to 1. This parameterization is often convenient especially because the estimates describe standard deviation like quantities rather than variance. It is also used to implement the M3 and M4 censoring methods.

Slide 16

Estimation Code - YLO

```
$INPUT ID TIME WT AGE SEX DVID DV MDV EVID CMT AMT
$DATA ka11_blq1_simln.csv ; CSV file created from NONMEM table


$EST METHOD=COND INTER NOABORT
LAPLACIAN
MAX=9990 NSIG=3 SIGL=9

$PK
IF (NEWIND.EQ.0) THEN
   LLOQ=1 ; MG/L
ENDIF
```

```
$ERROR
   YLO=LLOQ
   CP=A(2)/V
   PROP=CP*RUVCV
   ADD=RUVSD
   SD=SQRT(PROP*PROP+ADD*ADD)
   Y=CP + SD*EPS1
```


When NEWIND is 0 this is the first record of the data set. It is most efficient to assign constants a value at this point e.g. here the lower limit of quantitation (LLOQ) is assigned a value of 1.

The special variable YLO is assigned the value of LLOQ in $ERROR so that NONMEM can adjust the likelihood under the assumption that all observed values are above the LLOQ. NONMEM pays no attention to any MDV values although it is important that the DV value when MDV is 1 is no greater than LLOQ otherwise an error is generated. The usual value for DV when MDV is 1 will be 0 (often coded in the data file as ".").

Slide 17	**Estimation Code – M3**
	
	`$INPUT ID TIME WT AGE SEX DVID DV MDVX=DROP EVID CMT AMT`
	`$DATA ka11_blq1_simln.csv ; CSV file created from NONMEM table`
	
	`$EST METHOD=COND INTER NOABORT`
	`LAPLACIAN`
	`MAX=9990 NSIG=3 SIGL=9`
	
	`$PK`
	`IF (NEWIND.EQ.0) THEN`
	`LLOQ=1 ; MG/L`
	`ENDIF`
	
	`$ERROR`
	`CP=A(2)/V`
	`PROP=CP*RUVCV`
	`ADD=RUVSD`
	`SD=SQRT(PROP*PROP+ADD*ADD)`
	`IF (DV.GE.LLOQ) THEN`
	`F_FLAG=0 ; ELS`
	`Y=CP + SD*EPS1`
	`ELSE`
	`F_FLAG=1 ; LIKELIHOOD`
	`Y=PHI((LLOQ-CP)/SD))`
	`ENDIF`
	
	©NHG Holford, 2015, all rights reserved.
		The F_FLAG variable tells NONMEM how to interpret the prediction variable 'Y'.
		The default F_FLAG=0 is used for the usual case of a continuous variable like concentration which has a residual error associated with it. The value returned is the prediction of the observation.
		When F_FLAG=1 it means the likelihood of the observation is returned.
		When F_FLAG=2 it means -2 x log likelihood is returned.
		
		The PHI function computes the likelihood of the predicted concentration being less than the LLOQ.
		
		The MDVX data item is dropped. NM-TRAN will generate MDV=0 for all observations (EVID=0) including those which are below the LLOQ.

Slide 18	**Estimation Code – M4**
	
	`$INPUT ID TIME WT AGE SEX DVID DV MDVX=DROP EVID CMT AMT`
	`$DATA ka11_blq1_simln.csv ; CSV file created from NONMEM table`
	
	`$EST METHOD=COND INTER NOABORT`
	`LAPLACIAN`
	`MAX=9990 NSIG=3 SIGL=9`
	
	`$PK`
	`IF (NEWIND.EQ.0) THEN`
	`LLOQ=1 ; MG/L`
	`ENDIF`
	
	`$ERROR`
	`CP=A(2)/V`
	`PROP=CP*RUVCV`
	`ADD=RUVSD`
	`SD=SQRT(PROP*PROP+ADD*ADD)`
	`IF (DV.GE.LLOQ) THEN`
	`F_FLAG=0 ; ELS`
	`Y=CP + SD*EPS1`
	`ELSE`
	`F_FLAG=1 ; LIKELIHOOD`
	`CUMD=PHI((LLOQ-CP)/SD)`
	`CUMDZ=PHI(-CP/SD)`
	`Y=(CUMD-CUMDZ)/(1-CUMDZ)`
	`ENDIF`
	
	©NHG Holford, 2015, all rights reserved.
		The PHI function is used to compute the likelihood of the predicted concentration being both less than LLOQ (CUMD) and also being non-negative (CUMDZ).
		The MDVX data item is dropped. NM-TRAN will generate MDV=0 for all observations (EVID=0) including those which are below the LLOQ.

Slide 19		http://www.seriouswheels.com/2007/2007-IMSA-Lamborghini-Gallardo-Spyder-Rear-Angle-Top-Down-3-1024x768.htm