

Limit of Detection (LoD) Estimation Using Parametric Curve Fitting to (Hit) Rate Data: The LoD_Est SAS[®] Macro

Jesse A. Canchola, Roche Molecular Diagnostics, Pleasanton, California, USA
Pari Hemyari, Roche Molecular Diagnostics, Pleasanton, California, USA

ABSTRACT

The Limit of Detection (LoD) is defined as the lowest concentration or amount of material, target or analyte that is consistently detectable (for PCR quantitative studies, in at least 95% of the samples tested)¹. In practice, the estimation of the LoD uses a parametric curve fit to a set of panel member (PM1, PM2, PM3, etc.) data where the responses are binary. Typically, the parametric curve fit to the percent detection levels takes on the form of a probit or logistic distribution. For this, the SAS PROBIT procedure can be used to fit a variety of distributions, including both the probit and logistic. We introduce the LOD_EST SAS macro that takes advantage of the SAS PROBIT procedure's strengths and returns an information-rich graphic as well as a percent detection table with associated 95% exact (Clopper-Pearson) confidence intervals for the hit rates at each level.

INTRODUCTION AND BACKGROUND

For analytical sensitivity in a PCR assay, we would like to know the lowest amount of analyte we can reliably detect (typically taken as 95% of samples tested). This is important in the diagnosis and monitoring of target viruses (e.g., HBV, HCV, HIV)* or bacteria (e.g., Listeria, MAP, MRSA)**. For example, the analytical sensitivity of the assay test for the HIV-1 virus in the blood stream can be used to diagnose disease in a patient who is then placed on a drug treatment regimen. Subsequently, the viral load of the patient is monitored regularly to ensure the treatment is efficacious.

For quantitative assays, the analytical sensitivity is measured by the limit of detection (LoD) and sometimes called the lower limit of detection (LLoD). For the remaining presentation, we use the more common LoD acronym and "sensitivity" to mean "analytical sensitivity", as it relates to a quantitative assay, throughout this document.

To begin, an experiment is performed that collects information about the lower end of the quantitative assay. Typically, several (panel) levels are targeted at the lower portion of the assay range (Table 1), where it may not be linear on the log₁₀ response, in order to fit a parametric model curve to obtain an estimate of the LoD where the curve crosses 95% detection or hit rate (Figure 1).

Operationally, the levels or panels in the experiment are chosen to include at least one panel member at 100% detection, another to anchor the parametric curve at the bottom end (not including zero) with the remaining three or more levels targeting the region where one believes the LoD might be located. Table 1 shows an example of six panel levels (not including zero) with the top end anchored by an HCV RNA assay level at 50 IU/mL and one anchor at the bottom at 2.5 IU/mL. There are four levels in between both the top and bottom ends: 25, 15, 10 and 5 IU/mL.

Table 1. LoD in EDTA Plasma from "Empower change in HCV" for COBAS[®] AmpliPrep/COBAS[®] TaqMan[®] HCV Qualitative Test, v2.0²

Input Titer (HCV RNA IU/mL)	No. of Valid Replicates	No. of Positives	Hit Rate in %
50	252	252	100
25	252	251	100
15	251	246	98
10	252	233	92
5	252	183	73
2.5	251	124	49
0	250	0	0
LOD by PROBIT at 95% Hit Rate 12 IU/mL 95% confidence range: 10 – 14 IU/mL			

At the 95%
hit rate...

The LoD is
somewhere
between 10
to 15 IU/mL.

* HBV=Hepatitis B Virus, HCV=Hepatitis C Virus, HIV=Human Immunodeficiency Virus.

** MAP=mycobacterium avium subspecies paratuberculosis, MRSA=Methicillin-resistant Staphylococcus aureus.

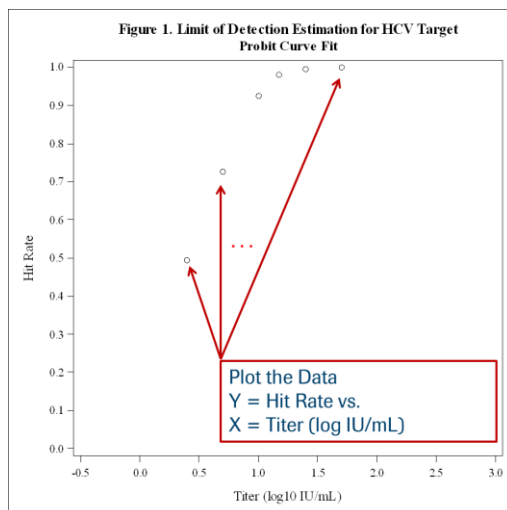
METHODS

The logical stages for visualizing how to find the LoD are as follows:

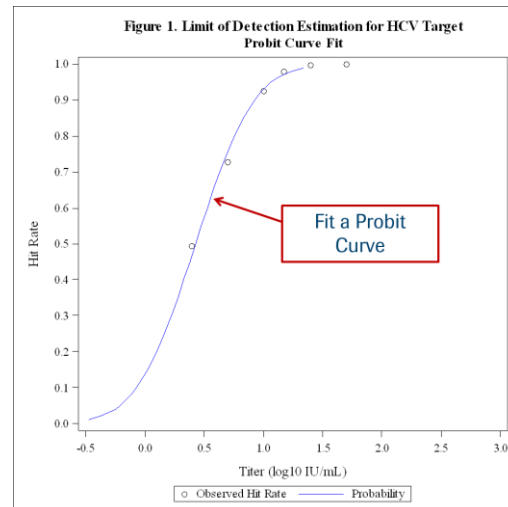
- A. Plot the “Hit Rate” or “Percent Detection” (Y) versus \log_{10} concentration (X);
- B. Fit a parametric distribution (not exhaustive):
 1. Probit Curve
 2. Logistic (Logit) Curve;
- C. Draw a horizontal line at the 95% hit rate (or percent detection) position;
- D. Where the curve in B1/B2 hits the 95% hit rate in C, estimate the corresponding LoD;
- E. Find the corresponding 95% confidence interval for the LoD.

EXPANDED STAGES

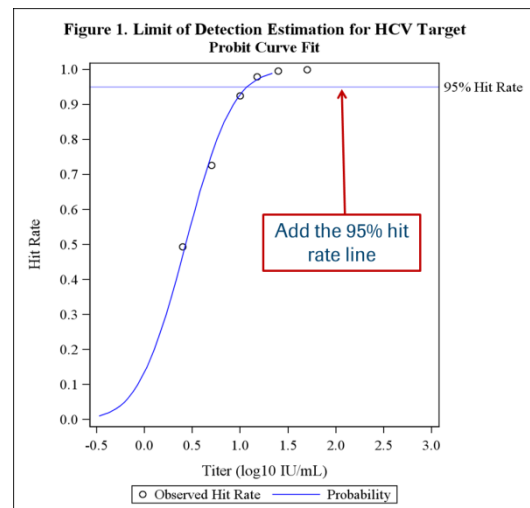
- A. Plot the “Hit Rate” (Y) versus \log_{10} concentration (X):



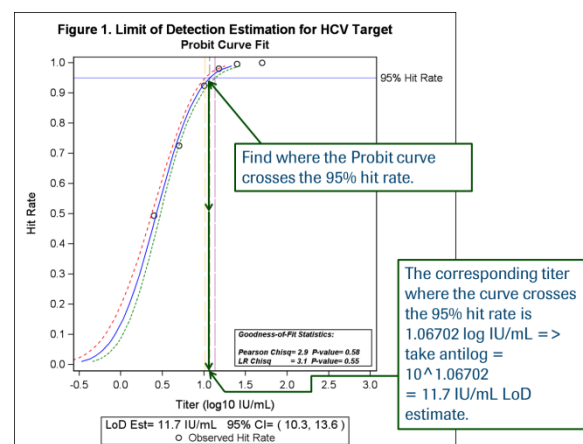
- B. Fit a parametric distribution:
 1. a. Probit Curve or b. Logistic (Logit) Curve:



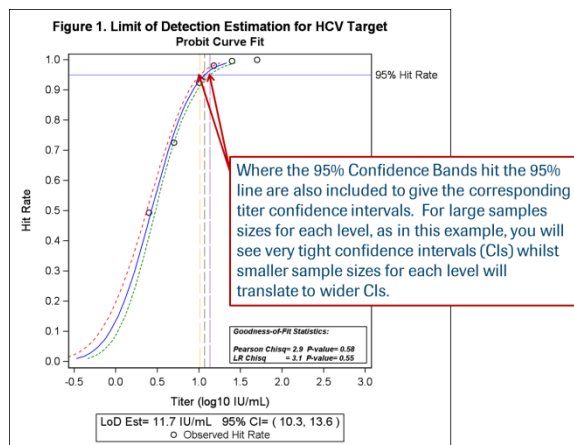
- C. Draw a horizontal line at the 95% hit rate (or percent detection) position:



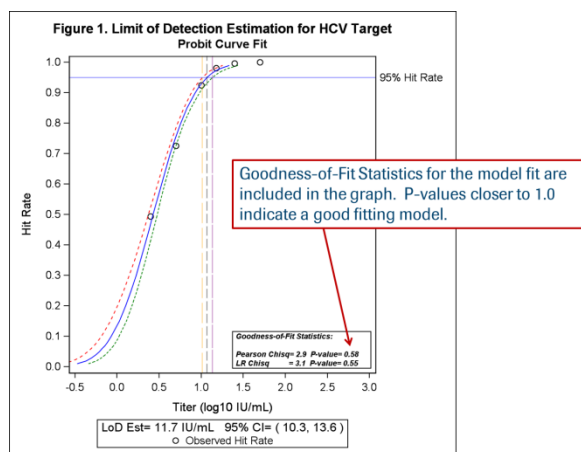
- D. Where the curve in B1 or B2 hits the 95% hit rate in C, estimate the corresponding LoD:



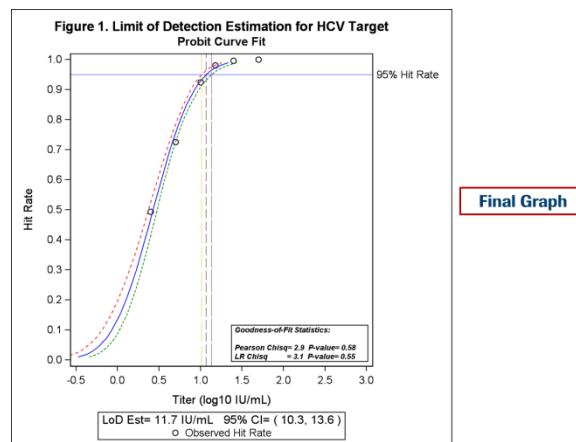
- E. Find the corresponding 95% confidence interval for the LoD:



Moreover, the statistics for how well the curve fits the data, the so-called goodness-of-fit (GOF) are included in the produced figure. The GOF statistics presented are the Pearson chi-square and the Likelihood Ratio (LR) chi-square. Associated p-values closer to 1.0 for these statistics indicate better fitting models.



The final figure is then presented as follows:



INTRODUCING THE LOD_EST SAS MACRO

LoD_Est is a SAS macro for fitting parametric curves to hit-rate (or percent detection) data in order to find an estimate of the Limit of Detection (LoD). The macro uses the SAS PROBIT procedure as the main computation engine.

There are four main steps to successfully run the LoD_Est SAS macro that follow.

Step 1: Enter the following (from Table 1):

- Target descriptive ①
- Level concentrations ②
- Number of Positive Hit Counts ③
- Total Test Counts ④

```
data LoD ;
input Target $3. Titer Hits TotalTests @ ;
Log10_Titer = log10(Titer) ;
Obs_HitRate = Hits / TotalTests ;
Obs_HitPcnt = Obs_HitRate * 100 ;
datalines ;
HCV 50 252 252
HCV 25 251 252
HCV 15 246 251
HCV 10 233 252
HCV 5 183 252
HCV 2.5 124 251
;
run ;
```

① → Target
② → Titer
③ → Hits
④ → TotalTests

Step 2: Load the LoD_Est.sas SAS macro into your session.

```
%include "C:\sas\macros\LoD\LoD_Est.sas" ;
```

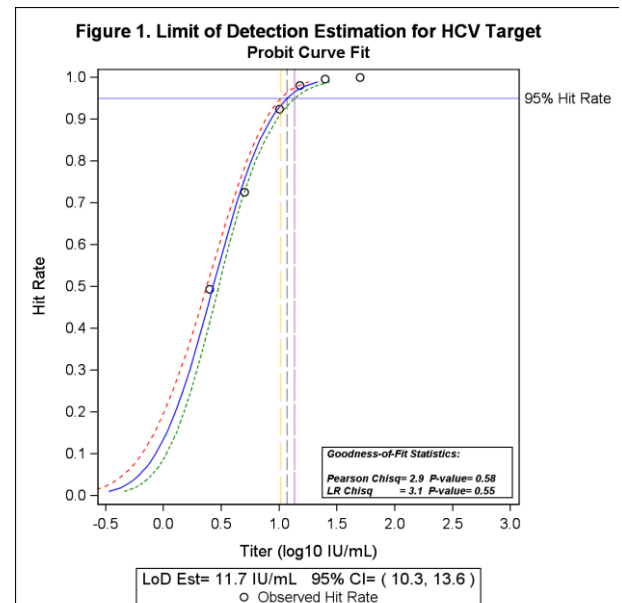
Step 3: Enter LoD_Est SAS macro inputs.

```
%LoD_Est (
datain      = LoD      ,
study       = CTM_1    ,
target      = &target   ,
subset      = target = "HCV" ,
Units       = IU/mL    ,
distn       = Probit   ,
alpha       = 0.05     ,
sided       = 2        ,
TableNo     = 1        ,
FigureNo    = 1        ,
ODStype     = RTF      ,
RTFout      = C:\HIVstudy\RTFout\HIV_LoD.RTF ,
GraphMin    = -0.5     ,
GraphMax    = 2.5      ,
GraphByInc  = 0.5      ,
GraphTitle1 = Figure 1. LoD Estimation for HIV ,
GraphTitle2 = &distn Curve Fit ,
YaxisTitle  = Hit Rate ,
XaxisTitle  = Titer (log10 &Units) ;
```

A Percent Detection table is generated along with associated 95% exact CIs.

Step 4: Run your SAS code and obtain the following results:

Resulting graph for LoD in EDTA Plasma Example



Target	Titer (IU/mL)	Titer (log10 IU/mL)	No. Detected	Total Tests	Percent Detection	95% Exact LL	95% Exact UL
HCV	50	1.6990	252	252	100	98.5	100
	25	1.3979	251	252	99.6	97.8	100
	15	1.1761	246	251	98.0	95.4	99.3
	10	1.0000	233	252	92.4	88.4	95.3
	5.0	0.6990	183	252	72.6	66.6	78.0
	2.5	0.3979	124	251	49.4	43.0	55.7

DISCUSSION

We have seen that the LoD_Est SAS macro is simple to use when a Limit of Detection graph, with information-rich details, is desired. The reader should note that there are some occasions where the distribution does not provide an adequate fit to the percent detection data and the methodology cannot converge to a parameter estimate for the LoD or if it does, it may not be able to compute the 95% confidence interval. Under these instances, the reader can turn to the maximum likelihood methodology that can provide such estimates. For this, see our LoD_MLE SAS macro.

CONCLUSION

The LoD_Est SAS macro can be used to produce a camera-ready graph with an estimate of the LoD and associated 95% confidence interval. In addition, a percent detection (or hit-rate) table is produced with 95% exact (Clopper-Pearson) confidence intervals for the hit rate at each level.

REFERENCES

1. Clinical and Laboratory Standards Institute (CLSI) document EP17-A2. "Evaluation of Detection Capability for Clinical Laboratory Measurement Procedures; Approved Guideline--Second edition." Wayne, PA. 2012.
2. Roche Molecular Systems, Inc. 2011. "Empower change in HCV" for COBAS® AmpliPrep/COBAS® TaqMan® HCV Qualitative Test, v2.06. Accessed on 05Jun2015. Available at: http://www.roche-diagnostics.ch/content/dam/corporate/roche-dia_ch/documents/broschuere/molecular_diagnostics/virology/06611656001_EN_EA_COBAS-AmpliPrep_COBAS-TaqMan-HCV-Qualitative-Test-v2.0.pdf
3. Pawitan Y. 2013. In All Likelihood: Statistical Modelling and Inference Using Likelihood. Oxford, United Kingdom: Oxford University Press. 528 pp.
4. Tang S, Hemyari P, Canchola J. 2016. "Composite Reference Standard in Diagnostic Research: A New Approach to Reduce Bias in the Presence of Imperfect Reference Tests". *In Review*.
5. Purcell S. 2007. "Maximum Likelihood Estimation". Accessed 05Jun2015. Available at: http://statgen.iop.kcl.ac.uk/bgim/mle/sslike_3.html
6. Canchola JA, Canchola CM, Canchola TL. 2016. "Limit of Detection (LoD) Estimation Using Maximum Likelihood Estimation on (Hit) Rate Data: The *LoD_MLE* SAS Macro." Proceedings of the SAS Global 2016 Conference. Cary, NC: The SAS Institute.

ACKNOWLEDGMENTS

The authors thank Enrique Marino and Alison J. Canchola for their valuable comments and suggestions.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the first author at:

Name: Jesse A. Canchola, MS, PStat®
Company: Roche Molecular Systems, Inc
Address: 4300 Hacienda Drive
City, State ZIP: Pleasanton, California, 94588
E-mail: Jesse.Canchola@Roche.Com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.

APPENDIX A: LOD_EST SAS MACRO CODE

```
/* ***** ;
* Macro LoD Est
* Calculates the LoD based on hit-rate data inputs
* Includes Percent Detection Table
*
* Developer:  Jesse A. Canchola, Roche Molecular Systems, Inc.
*
* Macro Inputs (with examples):
*
* datain      = LoD      ; * enter data set name ;
* study       = YourStudy; * enter study name or investigator name ;
* target      = CMV      ; * enter target/virus ;
* subset      = target="CMV" ; * enter "where" statement language ;
* Units       = IU/mL    ; * enter units of measurement: IU/mL, copies/mL, cp/mL, etc. ;
* distn       = Probit   ; * enter EXTREME, EXTREMEVALUE, GOMPERTZ, GOMPERZ, LOGISTIC, PROBIT ;
* alpha       = 0.05     ; * significance level ;
* sided       = 2        ; * 1 or 2 sided interval ;
* TableNo     = 1        ; * Table Number ;
* FigureNo    = 2        ; * Figure Number ;
* ODStype     = RTF      ; * enter ODS output type: RTF, PDF, HTML, etc. ;
* RTFout      = B:\SAS\Out ; * output &ODStype file ;
* GraphMin    = -1.0     ; * enter on log10 scale ;
* GraphMax    = 3.0      ; * enter on log10 scale ;
* GraphByInc  = 0.5      ; * enter scale increment ;
* GraphTitle1 = Figure &FigureNo.. Limit of Detection Estimation for &target. Target ; * main
*              title at top of graph ;
* GraphTitle2 = &distn Curve Fit ; * smaller title below main title at top of graph ;
* YaxisTitle  = Hit Rate ; * Y-axis title: Hit Rate, Detection Rate, Probability, etc.
* XaxisTitle  = Titer (log10 &Units) ; * X-axis title: Log10 Titer, Log Titer, Titer
*
/* ***** ;
%MACRO LoD_Est (
datain      = ,
study       = ,
target      = ,
subset      = ,
Units       = ,
distn       = ,
alpha       = ,
sided       = ,
TableNo     = ,
FigureNo    = ,
ODStype     = ,
RTFout      = ,
GraphMin    = ,
GraphMax    = ,
GraphByInc  = ,
GraphTitle1 = ,
GraphTitle2 = ,
YaxisTitle  = ,
XaxisTitle  = ,
) ;

proc format ;
  picture mltdecaf (multilabel)
    0-9      = '9.9'
    10-99    = '99'
    100-999  = '999'
    1000-9999 = '9999' ;
  picture mltdecbf (multilabel)
    0-9.9    = '9.99'
    10-99.9  = '99.9'
    100-999.9 = '999'
    1000-9999.9 = '9999' ;
run ;

proc sort data = &DataIn ; by Target ; run ;

data DataIn ;
```

```

Set &DataIn ;
where &subset ;
run ;

* Use the MODSTYLE macro supplied by SAS to change the default colors and markers ;
%modstyle(name = markers ,
          parent = listing ,
          type = CLM ,
          markers = circle triangle square) ;

ods graphics on / outputfmt=EMF ;
ods _ALL_ close ;
ods _ODS_ file = "&RTFout" gtitle style = markers ; * style=analysis ;

proc sort data=DataIn ; by Target ; run ;

*ods listing sge=on style=Analysis ;
ods noproctitle ;
ods output ProbitAnalysis=PredProbs GoodnessOfFit=GOF ;

%IF %upcase(&distn) = PROBIT %THEN %DO ;
  title1 "LoD Estimation with Probit Analysis" ;
  title2 " " ;
  proc probit data = DataIn lackfit ;
    model Hits / TotalTests = log10_titer
      / inversecl d = Normal ;
    predpplot var = log10_titer vref = 0.95 ;
    by Target ;
  run ;
%END ;
%ELSE %DO ;
  title1 "LoD Estimation with &distn Analysis" ;
  title2 " " ;
  proc probit data = DataIn lackfit ;
    model Hits / TotalTests = log10_titer
      / inversecl d = &distn ;
    predpplot var = log10_titer vref = 0.95 ;
    by Target ;
  run ;
%END ;

* combine/concatenate original LoD with PredProbs to overlay in SGPLOT ;
data ComboLoD_Target1 ;
set DataIn PredProbs ( rename=(TARGET=TARGET1) ) ;
run ;

* extracting LoD at 95% and associated 95% CI ;
data Estimates ;
set PredProbs ( rename = (Variable=LogLoDEst LowerCL=LogLCL UpperCL=LogUCL) ) ;

* Get anti-log for actual value in normal units ;
LoD_Est = 10**LogLoDEst ;
LCL95 = 10**LogLCL ;
UCL95 = 10**LogUCL ;

* convert to character with 4.2 format since numeric is out to more decimals and will not work
directly ;
ProbabilityChar = put(Probability,4.2) ;

* Keep only one line at 95% hit rate ;
if ProbabilityChar = "0.95" ;

* output values as macro vars for use in Proc SGPLOT ;
call symput("LogLoDEst",LogLoDEst) ;
call symput("LogLCL",LogLCL) ;
call symput("LogUCL",LogUCL) ;

call symput("LoD_Est" , compress(round(LoD_Est,0.1))) ;
call symput("LCL95" , compress(round(LCL95 ,0.1))) ;
call symput("UCL95" , compress(round(UCL95 ,0.1))) ;
run;

```

```

proc print data = Estimates ; run ;

* Extract Goodness-of-Fit statistics ;
* Name = "Pearson Chi-Square"
      "L.R.      Chi-Square"
Value
DF
ProbChiSq * ;
data GOFstats ; set GOF ; run ;
Proc IML ;
use GOFstats ;
read ALL VAR{Value DF ProbChiSq} into X ;
close GOFstats ;

PearsonStat = round(X[1,1],0.1) ;
PearsonDF   = round(X[1,2],0.1) ;
PearsonPval = round(X[1,3],0.01) ;

LRStat = round(X[2,1],0.1) ;
LRDF   = round(X[2,2],0.1) ;
LRPval = round(X[2,3],0.01) ;

print PearsonStat LRStat ;

call symputx("PearsonStat",PearsonStat) ;
call symputx("PearsonDF",PearsonDF) ;
call symputx("PearsonPval",PearsonPval) ;

call symputx("LRStat",LRStat) ;
call symputx("LRDF",LRDF) ;
call symputx("LRPval",LRPval) ;
run ;
proc print data=GOFstats ; run ;

ods graphics on / width = 550px height = 550px ;
proc sgplot data = ComboLoD_Target1 ;
  scatter x=Log10_Titer y=Obs_HitRate / LEGENDLABEL="Observed Hit Rate" MARKERATTRS=(SIZE=7px)
name="Assay1" ; * group=Target grouporder=ascending COLOR=blue SYMBOL=triangle) ;

  series x = variable y = Probability / MARKERATTRS=(SIZE=7px) LINEATTRS=(THICKNESS=0.2
COLOR=blue PATTERN=1) name="HitRate" ; */ CurveLabel = "Probit" ;
  series x = lowercl y = Probability / MARKERATTRS=(SIZE=7px) LINEATTRS=(THICKNESS=0.2 COLOR=red
PATTERN=2) name="LowerCL" ; */ CurveLabel = "95% LCL" ;
  series x = uppercl y = Probability / MARKERATTRS=(SIZE=7px) LINEATTRS=(THICKNESS=0.2
COLOR=green PATTERN=3) name="UPPERCL" ; */ CurveLabel = "95% UCL" ;

  * Reference line at 0.95 probability ;
  REFLINE 0.95 / AXIS=Y TRANSPARENCY = 0.5 LINEATTRS = (THICKNESS=0.2 COLOR=blue PATTERN=1)
LABEL = "95% &YaxisTitle" ; * Add reference line to 95% ;

  * Reference line at LoD (log) ;
  REFLINE &LogLoDEst / AXIS=X TRANSPARENCY = 0.5 LINEATTRS = (THICKNESS=0.2 COLOR=black
PATTERN=4) LABEL = "" ;

  * Reference line at lower 95% CI for LoD (log) ;
  REFLINE &LogLCL / AXIS=X TRANSPARENCY = 0.5 LINEATTRS = (THICKNESS=0.2 COLOR=orange
PATTERN=5) LABEL = "" ;

  * Reference line at upper 95% CI for LoD (log) ;
  REFLINE &LogUCL / AXIS=X TRANSPARENCY = 0.5 LINEATTRS = (THICKNESS=0.2 COLOR=purple
PATTERN=6) LABEL = "" ;

  YAXIS LABEL = "&YaxisTitle" VALUES = (0 TO 1 BY 0.1) ;
  XAXIS LABEL = "&XaxisTitle" VALUES = (&GraphMin TO &GraphMax BY &GraphByInc) ;

  KEYLEGEND "Assay1" / LOCATION=outside POSITION=bottom title=" LoD Est= &LoD_Est &Units 95%
CI= ( &LCL95, &UCL95 ) " ;

  INSET "Goodness-of-Fit Statistics: "
      "
      "
      "Pearson Chisq= &PearsonStat P-value= &PearsonPval"

```



```

"LR Chisq          = &LRStat P-value= &LRPval          " / BORDER TEXTATTRS=(Color=black
Family=Arial Size=6px Style=Italic Weight=Bold) ; *POSITION = TOPLEFT ;

```

```

TITLE1 "&GraphTitle1" ;
TITLE2 "&GraphTitle2" ;
run ;

* ***** * ;
* Hit Rate Table * ;
* ***** * ;
data DataIn1 ;
set DataIn ;

* 2-sided 95% Clopper-Pearson CI for hit rates ;
if &sided = 1 then do ;
    zalpha = probit( 1 - ( &alpha ) ) ;
end ;
else if &sided = 2 then do ;
    zalpha = probit( 1 - ( &alpha / 2 ) ) ;
end ;

n = TotalTests ;
p = Hits / TotalTests ;

** Exact Clopper Pearson ;
x = round( n * p, 0.1 ) ;

* Calculate the lower limit. ;
v1 = 2 * ( n - x + 1 ) ;
v2 = 2 * x ;

if &sided = 1 then do ;
    a = 1 - ( &alpha ) ;
end ;
else if &sided = 2 then do ;
    a = 1 - ( &alpha / 2 ) ;
end ;

coef = ( n - x + 1 ) / x ;
fscore = finv( a, v1, v2 ) ;
exact_lcl = 1 / ( 1 + coef * fscore ) ;

* Calculate the upper limit. ;
v11 = 2 * ( x + 1 ) ;
v22 = 2 * ( n - x ) ;
fscore2 = finv( a, v11, v22 ) ;
coef2 = ( x + 1 ) / (n-x) ;
numer = coef2 * fscore2 ;
exact_ucl = numer / ( 1 + numer ) ;
if exact_ucl = . & p = 1 then exact_ucl = 1.0 ; * check this ;
if exact_lcl = . then exact_lcl = 9999 ;

* converting ci to percent AND HANDLING MISSING DATA IF ANY ;
if exact_lcl ne 9999 then exact_lclp = exact_lcl * 100 ;
if exact_ucl ne 9999 then exact_uclp = exact_ucl * 100 ;
* ***** end Clopper-Pearson CI calculation ***** ;

drop N p x V1 V2 A Coef Fscore V11 V22 Fscore2 Coef2 Numer ;
run ;

* Proc report ;
options orientation = landscape nodate nonumber ;

ods listing close ;
ods escapechar='~' ;
proc report data = DataIn1 split='*' missing headline headskip spanrows
    style(report)=(just=center outputwidth=10 in)
    style(lines) =header{font_size=9pt font_face="Arial" font_weight=medium
background=transparent just=left}

```

```

        style(header)=header{font_size=9pt font_face="Arial" font_weight=bold
background=transparent just=center}
        style(column)=header{font_size=9pt font_face="Arial" font_weight=medium
background=transparent just=center};

columns Target Titer Log10_Titer Hits TotalTests Obs_HitPcnt Exact_LCLP Exact_UCLP ;

define Target / display "Target" group order=freq
    style(header)={just=center cellwidth=.6 in}
    style(column)={just=center cellwidth=.6 in vjust=center} ;

define Titer / display "Titer (&units)" f=mltdecaf.
    style(header)={just=center cellwidth=.55 in}
    style(column)={protectspecialchars=off just=center /*pretext="\qj\tqdec\tx600 */ cellwidth=.55
in} ;

define Log10_Titer / display "Titer*(log10 &units)" f=7.4
    style(header)={just=center cellwidth=.55 in}
    style(column)={protectspecialchars=off just=center /*pretext="\qj\tqdec\tx600 */ cellwidth=.55
in} ;

define Hits / display "No. Detected" f=6.0
    style(header)={just=center cellwidth=.55 in}
    style(column)={protectspecialchars=off just=center /*pretext="\qj\tqdec\tx600 */ cellwidth=.55
in} ;

define TotalTests / display "Total Tests" f=5.0
    style(header)={just=center cellwidth=.5 in}
    style(column)={protectspecialchars=off just=center /*pretext="\qj\tqdec\tx500 */ cellwidth=.5
in} ;

define Obs_HitPcnt / display "Percent*Detection" f=mltdecbf.
    style(header)={just=center cellwidth=.5 in}
    style(column)={protectspecialchars=off just=center /*pretext="\qj\tqdec\tx500 */ cellwidth=.5
in} ;

define Exact_LCLP / display "95% Exact LL" f=mltdecbf.
    style(header)={just=center cellwidth=.5 in}
    style(column)={protectspecialchars=off just=center /*pretext="\qj\tqdec\tx500 */ cellwidth=.5
in} ;

define Exact_UCLP / display "95% Exact UL" f=mltdecbf.
    style(header)={just=center cellwidth=.5 in}
    style(column)={protectspecialchars=off just=center /*pretext="\qj\tqdec\tx500 */ cellwidth=.5
in} ;

title1 bold j=center height=12pt f='Times' "Table &TableNo.. Percent Detection" ;
run ;

ods &ODStype close ;
%MEND LoD_Est ;

```

□