Paper 206-31

The Estimation of Sensitivity and Specificity of Clustered Binary Data

William F. McCarthy**,** Maryland Medical Research Institute, Baltimore, MD
Nan Guo, Maryland Medical Research Institute, Baltimore, MD

ABSTRACT

This paper reviews a methodology for the estimation of sensitivity and specificity of non-clustered binary data (patient level analysis) and presents a methodology for the estimation of sensitivity and specificity that considers the patient as a cluster and the coronary arteries (or coronary artery segments) as the diagnostic units of the study (DUOS) within each cluster. In addition, we present a SAS® Program with a SAS® Macro for the estimation of sensitivity and specificity that considers clustered binary data.

Key Words: Clustered Binary Data; Diagnostic Accuracy; Cardiovascular Imaging; SAS® Program; SAS® Macro

INTRODUCTION

Contrast-enhanced multi-detector row spiral computed tomography (MDCT) has been introduced as a method for non-invasive visualization of coronary artery stenosis. To determine the diagnostic accuracy of MDCT coronary angiography, as compared to the "gold standard" invasive coronary angiography, sensitivity and specificity are estimated (95% CIs). Three separate levels of estimation are computed: at the patient level, at the coronary artery level, and at the coronary artery segment level. We review the methodology for the estimation of sensitivity and specificity of non-clustered binary data (patient level analysis) and present a methodology for the estimation of sensitivity and specificity that considers the patient as a cluster and the coronary arteries (or coronary artery segments) as the diagnostic units of the study (DUOS) within each cluster. Finally, we present a SAS Macro for the estimation of sensitivity and specificity that considers clustered binary data.

The conventional binomial variance estimate [Equations 1.2, 1.3], which assumes that all measurements are independent, will not be valid for calculating the variance of sensitivity and specificity of clustered binary data. As an example, consider the assessment of sensitivity. When considering clustered binary data, the point estimate of sensitivity [Equation 1.7] gives the same estimate that would result if we had ignored the clustering, assumed independence, and used the conventional binomial point estimate [Equation 1.0]. However, if the variance of the sensitivity is estimated using the conventional binomial variance estimate [Equation 1.2], the variance will be underestimated if the correlation between the diagnostic units is positive (i.e., the estimated variance will be smaller than the true variance) or will be overestimated if the correlation between the diagnostic units is negative (i.e., the estimated variance will be larger than the true variance). The most likely scenario for our study is that the correlation between the diagnostic units is positive. A positive correlation indicates that if one diagnostic unit tests as a true positive (true negative) in a particular cluster, the probability of testing other diagnostic units in the same cluster as true positives (true negatives) will increase. To correct for this underestimated (overestimated) variance problem (bias), a ratio estimator for the variance of clustered binary data has been derived [Equation 1.8] (Cochran, 1977; Rao and Scott, 1992).

DEFINITIONS AND TERMS

The sensitivity of a diagnostic test is its ability to detect the condition of interest when it is present in patients.

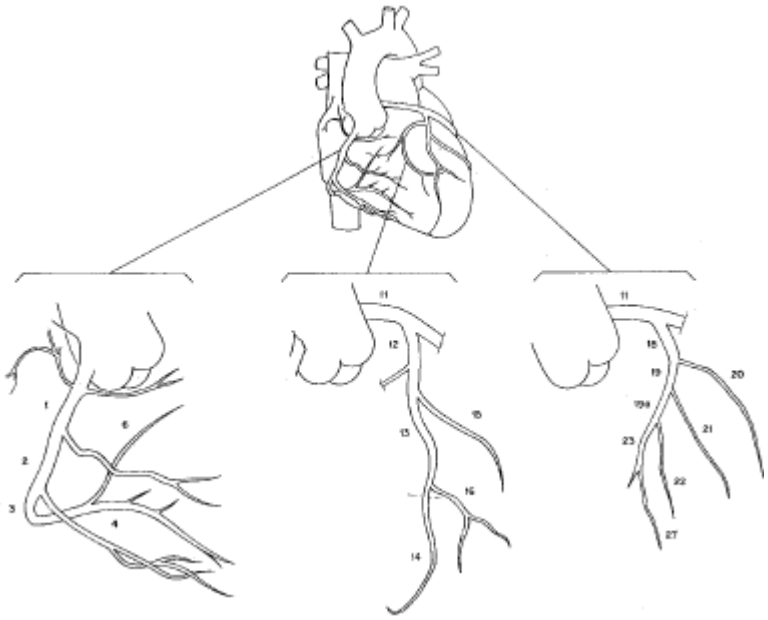
The specificity of a diagnostic test is its ability to exclude the condition of interest in patients without the condition.

The diagnostic unit of study (DUOS) is the smallest "unit" that is tested (Beam, 1998).

Gold standard: we determine the true condition status by means of a gold standard – "a source of information completely different from the test or tests under evaluation and which tells us the true condition status of the patient" (Zhou et al, 2002). For this paper, we shall say there is concordance between the test being evaluated and the "gold standard" if the same (identical) coronary artery or same (identical) coronary artery segment has the event of interest by both the test being evaluated and the "gold standard". For example, if coronary artery segment 4 is determined to be positive by the test being evaluated as well as by the "gold standard", we have concordance.

Stenosis: Narrowing or constriction of a coronary artery (coronary artery segment).

Coronary Tree with coronary arteries and coronary artery segments (Wang et al, 2004).



CONVENTIONAL BINOMIAL ESTIMATOR METHOD: ESTIMATION IN A SINGLE SAMPLE, BINARY-SCALED DATA

This method is used for the patient level analysis. It is a non-clustered method. As an example, we will use the following binary-scale event definition: Event= Patient has a $\geq$50% stenosis and Non-event= Else. The same (identical) coronary artery or same (identical) coronary artery segment has the event. Only one coronary artery with an event or one coronary artery segment with an event is required. Table 1 shows how the non-clustered data for binary diagnostic test results can be displayed.

Table 1. Data Layout for Non-Clustered Binary-Scale Data

		TEST RESULT		
True Condition Status	Positive (T=1)		Negative (T=0)	Total
Present (D=1)	s_1		s_0	n_1
Absent (D=0)	r_1		r_0	n_0
Total	m_1		m_0	N

The true condition status of the patient is determined by the use of the “gold standard”.

The sensitivity and specificity estimates are:

$$\hat{Se} = \frac{s_1}{n_1}$$ sensitivity [1.0]

and

$$\hat{Sp} = \frac{r_0}{n_0} \qquad \text{specificity} \qquad [1.1]$$

The variance of the sensitivity and specificity is the variance of a proportion:

$$\hat{Var}\left(\hat{Se}\right) = \frac{\hat{Se}(1-\hat{Se})}{n_1} = \frac{s_1 s_0}{n_1^3} \qquad [1.2]$$

and

$$\hat{Var}\left(\hat{Sp}\right) = \frac{\hat{Sp}(1-\hat{Sp})}{n_0} = \frac{r_1 r_0}{n_0^3} \qquad [1.3]$$

The usual approach to constructing a confidence interval for a measure of diagnostic accuracy assumes a large sample size (i.e., follows a normal distribution asymptotically). This confidence interval is referred to as an asymptotic interval. It has the following form:

$$\hat{\Theta} - z_{1-\alpha/2}\sqrt{\hat{Var}\left(\hat{\Theta}\right)},\ \hat{\Theta} + z_{1-\alpha/2}\sqrt{\hat{Var}\left(\hat{\Theta}\right)} \qquad [1.4]$$

where $\hat{\Theta}$ is the estimate of the accuracy measure (either sensitivity or specificity), Θ; $z_{1-\alpha/2}$ is the upper $\alpha/2$ percentile of the standard normal distribution, and $100(1-\alpha)\%$ is the confidence level.

Agresti and Coull (1998) have noted that the "Wald [confidence] interval performs poorly unless n is quite large (e.g., Ghosh, 1979; Blyth and Still, 1983)". Construction of a confidence interval based on Equation 1.4 and using Equations 1.0 and 1.2 and Equations 1.1 and 1.3, is based on the Wald confidence interval.

If the sample size is small, then the confidence limits for the sensitivity are estimated with the following equation (Agresti and Coull, 1998):

$$\frac{\hat{Se} + z^2_{1-\alpha/2}/(2n_1) \pm z_{1-\alpha/2}\sqrt{[\hat{Se}(1-\hat{Se}) + z^2_{1-\alpha/2}/(4n_1)]/n_1}}{1 + z^2_{1-\alpha/2}/n_1} \qquad [1.5]$$

If the sample size is small, then the confidence limits for the specificity are estimated with the following equation (Agresti and Coull, 1998):

$$\frac{\hat{Sp} + z^2_{1-\alpha/2}/(2n_0) \pm z_{1-\alpha/2}\sqrt{[\hat{Sp}(1-\hat{Sp}) + z^2_{1-\alpha/2}/(4n_0)]/n_0}}{1 + z^2_{1-\alpha/2}/n_0} \qquad [1.6]$$

ESTIMATION IN A SINGLE SAMPLE: CLUSTERED BINARY-SCALED DATA

As an example, we will use the following binary-scale event definitions: Event= same (identical) coronary artery segment vessel with $\geq$50% stenosis and Non-event= Else. All coronary artery segments with an event are considered.
If one assesses sensitivity and specificity of a diagnostic test when clustered binary-scaled data is used, the data layout presented in Table 1 is modified as follows in Tables 2a and 2b:

Table 2a: Data Layout for the Assessment of Sensitivity

Patient (cluster)	No. of TP_i	No .of Segments (N_i) with Event	$\hat{Se}_i$	$N_i / \bar{N}$	$(N_i / \bar{N})^2 (\hat{Se}_i - \hat{Se})^2$
1					
2					
Etc					

Each patient is considered a cluster (i=1,2,..,I clusters); within the cluster are a number of segments.

No. of TP_i refers to the number of segments that are test positives in cluster i, as determined by the test being evaluated (MDCT angiography).

No. of Segments (N_i) with Event refers to the number of segments that are true positives in cluster i, as determined by the "gold standard" (invasive coronary angiography).

$\hat{Se}_i$ = No. of TP_i / No. of segments with event (N_i).

$N_i / \bar{N}$ = the cluster size for patient$_i$ / mean cluster size.

$$\hat{Se} = \frac{\sum_{i=1}^{I} N_i \hat{Se}_i}{\sum_{i=1}^{I} N_i} \qquad [1.7]$$

$$\hat{Var}(\hat{Se}) = \frac{1}{I(I-1)} \sum_{i=1}^{I} \left[\left(N_i / \bar{N} \right)^2 (\hat{Se}_i - \hat{Se})^2 \right] \qquad [1.8]$$

where $\bar{N} = \Sigma \frac{N_i}{I}$ is the mean cluster size.

Table 2b: Data Layout for the Assessment of Specificity

Patient (cluster)	No. of TN_i	No. of Segments (N_i) without Event	$\hat{Sp}_i$	$N_i / \bar{N}$	$(N_i / \bar{N})^2 (\hat{Sp}_i - \hat{Sp})^2$
1					
2					
Etc					

Each patient is considered a cluster (i=1,2,..,I clusters); within the cluster are a number of segments.

No. of TN_i refers to the number of segments that are test negatives in cluster i, as determined by the test being evaluated (MDCT angiography).

No. of Segments (N_i) without Event refers to the number of segments that are true negatives in cluster i, as determined by the "gold standard" (invasive coronary angiography).

$\hat{Sp}_i$ = No. of TN_i / No. of segments without event (N_i).

$N_i / \bar{N}$ = the cluster size for patient$_i$ / mean cluster size.

$$\hat{Sp} = \frac{\sum_{i=1}^{I} N_i \hat{Sp}_i}{\sum_{i=1}^{I} N_i} \qquad [1.9]$$

$$\hat{Var}(\hat{Sp}) = \frac{1}{I(I-1)} \sum_{i=1}^{I} \left[\left(N_i / \bar{N} \right)^2 (\hat{Sp}_i - \hat{Sp})^2 \right] \qquad [1.10]$$

where $\bar{N} = \Sigma \frac{N_i}{I}$ is the mean cluster size.

The usual approach to constructing a confidence interval for a measure of diagnostic accuracy assumes a large sample size (i.e., follows a normal distribution asymptotically). This confidence interval is referred to as an asymptotic interval. It has the following form [Equation 1.4]:

$$\hat{\Theta} - z_{1-\alpha/2} \sqrt{\hat{Var}\left(\hat{\Theta}\right)} \;,\; \hat{\Theta} + z_{1-\alpha/2} \sqrt{\hat{Var}\left(\hat{\Theta}\right)}$$

where $\hat{\Theta}$ is the estimate of the accuracy measure, Θ; $z_{1-\alpha/2}$ is the upper $\alpha/2$ percentile of the standard normal distribution, and $100(1-\alpha)\%$ is the confidence level.

ILLUSTRATION OF CLUSTERING EFFECT ON BINARY-SCALED DATA

Using the data layout outlined in Table 2a and the equations [1.7], [1.8], we illustrate the effect of clustering on binary-scaled data (Refer to Table 3 and 4). We will estimate the 95% CI of sensitivity at the coronary artery segment level.

Table 3. Data Layout of Example
(Based on an adaptation of data and example presented by Zhou et al, 2002)

Patient	No. of TP_i (MDCT)	No. of segments (N_i) with Event (invas. c. angio.)	$\hat{Se}_i$	$N_i / \bar{N}$	$(N_i / \bar{N})^2 (\hat{Se}_i - \hat{Se})^2$
1	0	3	0.000	1.235	0.9391
2	2	3	0.667	1.235	0.0211
3	3	3	1.000	1.235	0.0710
4	1	1	1.000	0.412	0.0079
5	2	3	0.667	1.235	0.0211
6	4	4	1.000	1.647	0.1263
7	3	3	1.000	1.235	0.0710
8	2	2	1.000	0.824	0.0316
9	2	2	1.000	0.824	0.0316
10	1	1	1.000	0.412	0.0079
11	2	3	0.667	1.235	0.0211
12	2	2	1.000	0.824	0.0316
13	3	3	1.000	1.235	0.0710
14	2	2	1.000	0.824	0.0316
15	0	2	0.000	0.824	0.4174
16	2	3	0.667	1.235	0.0211
17	2	3	0.667	1.235	0.0211
18	2	3	0.667	1.235	0.0211
19	2	2	1.000	0.824	0.0316
20	1	1	1.000	0.412	0.0079
21	2	2	1.000	0.824	0.0316
Sum	40	51			2.0357

Each patient is considered a cluster (i=1,2,..,I clusters); within the cluster are a number of segments. We have 21 patients (clusters) with an event based on invasive coronary angiography; we have 51 coronary artery segments (DUOS) with an event based on invasive coronary angiography.

No. of TP_i refers to the number of segments that are test positives in cluster i, as determined by the test being evaluated (MDCT angiography).

No. of Segments (N_i) with Event refers to the number of segments that are true positives in cluster i, as determined by the "gold standard" (invasive coronary angiography).

We shall say there is concordance between the test being evaluated and the "gold standard" if the same (identical) coronary artery segment has the event of interest. That is, if coronary artery segment 4 is determined to be positive by the test being evaluated as well as by the "gold standard", we have concordance. There must be a one-to-one correspondence between the coronary artery segment determined to be positive by the test being evaluated and the "gold standard" in order for the diagnostic accuracy assessment to be meaningful!

$$\hat{Se} = \frac{\sum_{i=1}^{I} N_i \hat{Se}_i}{\sum_{i=1}^{I} N_i} = 40/51 = 0.7843.$$

$N_i / \bar{N}$ = the cluster size for patient$_i$ / 2.43.

$$\hat{Var}(\hat{Se}) = \frac{1}{I(I-1)} \sum_{i=1}^{I} \left[\left(N_i / \bar{N} \right)^2 (\hat{Se}_i - \hat{Se})^2 \right] = \quad 2.0357/[(21)(20)] = 0.0048$$

where $\bar{N} = \Sigma \frac{N_i}{I}$ is the mean cluster size.

If the clustering effect was ignored, $\hat{Var}(\hat{Se})$ = 0.7843(0.2157)/51=0.0033 (based on Equation 1.2).

Table 4. 95% Confidence Limits (using Equation 1.4)

Clustering Effect not Considered	Clustering Effect Considered
0.7843 ± 0.1126 (0.6717, 0.8969)	0.7843 ± 0.1358 (0.6485, 0.9201)

One should note that if the clustering effect is not taken into account in this example, the calculated variance (thus the calculated confidence limits) is inappropriately small because the calculation ignores the positive correlation among the coronary artery segments nested within the patient. Even if the amount of correlation among the coronary artery segments nested within the patient is small, the variance (confidence limits) will be biased if the clustering effect is not taken into account.

REFERENCES

Agresti A and Coull B (1998). Approximate is better than "exact" for interval estimation of binomial proportions. Am. Stat. 52: 119-126.

Beam CA (1998). Analysis of clustered data in receiver operating characteristic studies, Stat. Methods Med. Res. 7: 324-336.

Blyth CR and Still HA (1983). Binomial confidence intervals, Journal of the American Statistical Association, 78: 108-116.

Cochran W. (1977). Sampling Techniques, 3rd ed., John Wiley and Sons, New York.

Ghosh BK (1979). A comparison of some approximate confidence intervals for the binomial parameter, Journal of the American Statistical Association, 74: 894-900.

Rao NK and Scott AJ (1992). A simple method for the analysis of clustered binary data, Biometrics 48: 577-585.

Wang JC, Normand S-L T, Mauri L and Kuntz RE (2004). Coronary artery spatial distribution of acute myocardial infarction occlusions, Circulation 110: 278-284.

Zhou X-H, Obuchowski NA, McClish DK (2002). Statistical Methods in Diagnostic Medicine. Wiley, New York.

SAS® PROGRAM WITH SAS® MACRO

Example Data Set.

ptid	artery	segment	invasive	mdct
1	1	1	0	1
1	1	2	0	0
1	1	3	1	1
1	1	4	0	0
1	1	5	1	0
1	2	11	0	1
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

```
*******************************************************************************
*  Creates a Look-Up Table Z value for N(0,1) distribution for selected err levels *
*******************************************************************************;
data dslkup;
 format err 5.3 tail1 tail2 5.3;
 input err tail1 tail2 @@;
cards;
0.5  0      0.674 0.4  0.253  0.842 0.3  0.524  1.036 0.2  0.842  1.282
0.1  1.282  1.645 0.05  1.645  1.96 0.025  1.96   2.248 0.01  2.326  2.576
0.005  2.576  2.813
;run;

*****************************************
* Segment-Level. Sensitivity & Specificity *
*****************************************;
data segm_tbl2a (keep=ptid TPi    segmEvnts   )  /* Sensitivity data set */
     segm_tbl2b (keep=ptid NotTPi NotsegmEvnts)  /* Specificity data set */;
  retain TPi segmEvnts NotTPi NotsegmEvnts ;
  set sorted;  /* We assume that data set has been sorted already */
  by ptid;
  if first.ptid then
     do;
       TPi         =0; /* No of TPi                          */
       segmEvnts   =0; /* No. of segment (Ni) with Event   */
       NotTPi      =0; /* No of TNi                          */
       NotsegmEvnts=0; /* No. of segment (Ni) without Event*/
     end;
  segmEvnts   =sum(segmEvnts,(invasive=1)         );
  TPi         =sum(TPi,      (invasive=1 & MDct=1));
  NotsegmEvnts=sum(NotsegmEvnts,      (invasive=0));
  NotTPi      =sum(NotTPi,   (invasive=0 & MDct=0));
  if last.ptid then
   do;
     if TPi >0 & segmEvnts >0 then output segm_tbl2a;            * Sentitivity *;
        if NotTPi > 0 & NotsegmEvnts >0 then output segm_tbl2b;  * Specificity *;
   end;
  label
    ptid          = "Patient ID"
    TPi           = "No of TPi"
    segmEvnts     = "No. of segment (Ni) with Event"
    NotTPi        = "No of TNi"
    NotsegmEvnts  = "No. of segment (Ni) without Event";
run;
```

```
%macro Calc_sum(ds,Dxac);
%global sumT sumN I &Dxac meanN;
data _null_;
 retain sum&Dxac.T sum&Dxac.N;
 set &ds end=EOF;
 %if _N_ eq 1 %then
                  %do;sum&Dxac.T=0;sum&Dxac.N=0;%end;
 %if %upcase(&Dxac)=SE %then
      %do;sum&Dxac.T =sum(sum&Dxac.T,TPi);sum&Dxac.N =sum(sum&Dxac.N,SegmEvnts);%end;
  %else %if %upcase(&Dxac)=SP %then
   %do;sum&Dxac.T=sum(sum&Dxac.T,NotTPi);sum&Dxac.N=sum(sum&Dxac.N,NotSegmEvnts);%end;
 if EOF then
    do;
      call symputx('sumT',sum&Dxac.T);
      call symputx('sumN',sum&Dxac.N);
      call symputx('I',_N_);
      call symputx("&Dxac",sum&Dxac.T/sum&Dxac.N);
      call symputx('meanN',sum&Dxac.N/_N_);
    end;
run;
%mend calc_sum;

* --- Sensitivity  --- *;
%calc_sum(segm_tbl2a,se)
data segm_se;
  retain totpart3;
  set segm_tbl2a end=EOF;
  if _N_ eq 1 then totpart3=0;
  Sei     =TPi/SegmEvnts;
  NidivN  =SegmEvnts/&meanN;
  part3   =(NiDivN)**2*(Sei-&se)**2;
  totpart3=sum(totpart3,part3);
  if EOF then call symputx('Varse',put(1/(&I*(&I-1))*totpart3,7.4));
  if _N_ le 5 then put _all_;
  label ptid            = "Patient ID"
        TPi             = "No. of TPi"
        SegmEvnts       = "No. of Segments (Ni) with Event"
        Sei             = "Sei"
        NidivN          = "Ni/Mean N"
        part3           = "(Ni/mean N)**2 * (Sei-Se)**2";
run;

* --- Specificity --- *;
%calc_sum(segm_tbl2b,sp)
data segm_sp;
  retain totpart3;
  set segm_tbl2b end=EOF;
  if _N_ eq 1 then totpart3=0;
  Spi     =NotTPi/NotSegmEvnts;
  NidivN  =NotSegmEvnts/&meanN;
  part3   =(NiDivN)**2*(Spi-&sp)**2;
  totpart3=sum(totpart3,part3);
  if EOF then call symputx('Varsp',put(1/(&I*(&I-1))*totpart3,7.4));
  label ptid            = "Patient ID"
        NotTPi          = " No of TNi"
        NotsegmEvnts    = "No. of Segments (Ni) without Event"
        Spi             = "Spi"
        NidivN          = "Ni/Mean N"
        part3           = "(Ni/mean N)**2 * (Spi-Sp)**2";
run;

%macro cau_Z(conf=95);
```

```
 %global ci Za sen sep spn spp;
 %let ci=&conf; %let Za=; %let sen=;%let sep=;%let spn=;%let spp=;
  data _null_;
    set dslkup (where=(err=%sysevalf(1-&conf/100)));  /* Look-up Table  */
    call symputx('Za',tail2);
  run;

   data temp;
    sqrtsen=sqrt(input("&varse",8.));
    sen=%sysevalf(&se)-%sysevalf(&Za)*sqrtsen;
    sep=%sysevalf(&se)+%sysevalf(&Za)*sqrtsen;
    call symputx('sen',put(sen,7.4));
    call symputx('sep',put(sep,7.4));

    sqrtsep=sqrt(input("&varsp",8.));
    spn=%sysevalf(&sp)-%sysevalf(&Za)*sqrtsep;
    spp=%sysevalf(&sp)+%sysevalf(&Za)*sqrtsep;
    call symputx('spn',put(spn,7.4));
    call symputx('spp',put(spp,7.4));
  run;
  %mend cau_Z;
%cau_Z; /* Default macro variable is 95 */

options nodate nonumber orientation=landscape;
proc template;   /* Creating a template */
  define style myrtf;
  parent=Styles.rtf;
  replace Body from Document
      "Controls the Body file." /
      bottommargin = 0.25in
      topmargin = 0.10in
      rightmargin = 0.10in
      leftmargin = 0.25in;
  end;
run;

ods listing close;  /* ODS starts */
ods rtf file='C:\SUGI 2006 paper\rtf files\segement_level_output.rtf' style=myrtf;
proc print data=segm_se noobs label data=all
 style(header)={font_face=Arial font_weight=bold font_size=8pt
 just=c background=white} style(data)  ={just=center font_face=Arial font_size=8pt};;
var ptid Tpi SegmEvnts Sei  NidivN part3 ;run;
ods escapechar='\';
ods rtf text='\R/RTF"\ul" Segment-Level ';
ods rtf text="\S={just=c font_size=10pt } Sensitivity &se";
ods rtf text="\S={just=c font_size=10pt } Variance of Sensitivity &varse";
ods escapechar='^';
ods rtf text="^S={just=c font_size=10pt } &CI% confidence interval for sensitivity
[&sen, &sep]";
proc print data=segm_sp noobs label data=all
 style(header)={font_face=Arial font_weight=bold font_size=8pt
 just=c background=white} style(data)  ={just=center font_face=Arial font_size=8pt};;
var ptid NotTPi NotsegmEvnts Spi NidivN part3 ;run;
ods escapechar='\';
ods rtf text='\R/RTF"\ul" Segment-Level ';
ods rtf text="\S={just=c font_size=10pt } Specificity &sp";
ods rtf text="\S={just=c font_size=10pt } Variance of Specificity &varsp";
ods escapechar='^';
ods rtf text="^S={just=c font_size=10pt } &CI% confidence interval for specificity
[&spn, &spp]";
ods rtf close;
ods listing; /* ODS Ends */
```

Example SAS® Program with SAS® MACRO for computing Segment-Level Sensitivity and Specificity.

SAS® OUTPUT

Patient ID	No. of TPi	No. of Segments (Ni) with Event	Sei	Ni/Mean N	(Ni/mean N)**2 * (Sei-Se)**2
1	5	10	0.50000	1.0	.000400
2	4	9	0.44444	0.9	.004624
3	6	11	0.54545	1.1	.000784
4	6	11	0.54545	1.1	.000784
5	5	9	0.55556	0.9	.001024

Segment-Level
Sensitivity 0.52
Variance of Sensitivity 0.0004
95% confidence interval for sensitivity [0.4808, 0.5592]

Patient ID	No of TNi	No. of Segments (Ni) without Event	Spi	Ni/Mean N	(Ni/mean N)**2 * (Spi-Sp)**2
1	4	12	0.33333	1.00000	0.013611
2	9	13	0.69231	1.08333	0.068906
3	3	11	0.27273	0.91667	0.026406
4	6	11	0.54545	0.91667	0.007656
5	5	13	0.38462	1.08333	0.005017

Segment-Level
Specificity 0.45
Variance of Specificity 0.0061
95% confidence interval for specificity [0.2969, 0.6031]

CONTACT INFORMATION

William F. McCarthy
Principal Statistician
Maryland Medical Research Institute
600 Wyndhurst Avenue
Baltimore, Maryland 21210-2425
(410) 435-4200
wmccarthy@mmri.org

Nan Guo
Certified Advanced SAS Programmer
Maryland Medical Research Institute
600 Wyndhurst Avenue
Baltimore, Maryland 21210-2425
(410) 435-4200
nguo@mmri.org