Paper SA21-2011

Outlier resistance of multivariate bioequivalence procedures

Dr. Srinand P. Nandakumar, Quintiles Inc., Kansas, USA

Dr. Joseph W. McKean, Western Michigan University, Michigan, USA

ABSTRACT

Bioequivalence tests of two drugs on pharmacokinetic parameters such as AUC (Area under the curve) and Cmax (maximum concentration) are implemented using least squares (LS) procedure on Schuirmann's (Schuirmann, 1987) two one-sided hypotheses. Nandakumar and McKean (Nandakumar, et al., 2011) generalized Schuirmann's hypothesis to a multivariate setting (for LS and its robust analogue). The least squares procedures are extremely susceptible to outliers and require strict adherence to the normality assumption. In this paper, a brief discussion on the merits of robust analysis following the principle of component-wise ranks on small sample data with outliers is investigated. SAS® programs are presented to illustrate through an example, the effects of outliers on LS procedures as applied to bivariate bioequivalence. Furthermore, the resistance to outliers by the robust procedure on this dataset is shown.

KEYWORDS: Component-wise Ranks, Average Bioequivalence, Hodges Lehmann's Estimator, Confidence Ellipse, Robust Statistics.

INTRODUCTION

Bioequivalence between drugs are routinely tested on pharmacokinetic parameters like AUC (Area under the concentration versus time curve) and Cmax (maximum concentration). Univariate LS (Least Squares) procedures are employed to test equivalence. (Ghosh, et al., 2007) and (Berger, et al., 1996) have challenged the normality assumption of these parameters. Also, the bioequivalence studies are often analyzed on small sample sizes (about 20 subjects). Small samples are more susceptible to outliers and yield dramatically different results if the normality assumption is violated. In this paper, we investigate the resistance to outliers of two multivariate tests for bioequivalence. One test is based on traditional least squares (LS) methods and the second is based on Wilcoxon rank-based procedures.

Outliers in bioequivalence analyses are seen often due to the inherent properties of the pharmacokinetic (PK) parameters. The PK parameters AUC and Cmax are estimated using some sort of interpolation/extrapolation techniques. In the case of extended release drugs, these extrapolations can lead to abnormally high estimates of AUC. Similarly, in case of rapid release drugs, the outliers can originate due to under-estimation of AUC (i.e very low values of AUC). Exclusion of such data points as possible outliers in the bioequivalence studies is not an option as they run the risk of bias with safety violations.

COMPARISON OF LEAST SQUARES AND ROBUST PROCEDURES

The LS methods that are employed in small sample studies run the risk of being biased due to outliers. An outlier has a low probability of originating from the same statistical distribution as the other

observations. Let $X_1 \dots X_n$ be a random sample from a distribution which is symmetric about θ. Then we can write the location model as

$$X_i = \theta + e_i\,, \;\; i = 1 \dots n$$

where $e_1 \dots e_n$ are iid random errors with density function f(x), which is symmetric about 0, and distribution function F(x). For the traditional LS analysis, the location for this sample is estimated as

$$\widehat{\theta_{LS}} = Argmin \sum_{i=1}^{n} |X_i - \theta|.$$

Of course, $\widehat{\theta_{LS}} = \bar{X}$, the sample mean, which minimizes the Euclidean (squared) norm. This estimate is a zero breakdown estimate (since the breakdown value is 1/n i.e. even one outlier can dramatically influence the location). Several robust location estimates have been previously suggested which have a higher breakdown (ex. the median, the Hodges Lehmann's estimate). For the rank-based estimates, the norm is defined by

$$\widehat{\theta_W} = Argmin \sum_{i=1}^{n} Rank(|X_i - \theta|)|X_i - \theta|.$$

The inference procedures based on this norm generalize the simple Wilcoxon test statistic. Minimizing the above norm yields the Hodges–Lehmann estimate (Hollander, et al., 1999) which is given by $\widehat{\theta_W} = median_{i \le j}\left\{\frac{X_i + X_j}{2}\right\}$. The properties and breakdown value are discussed in detail in (Hettmansperger, et al., 2011). Similar to location, the usual standard error of the LS estimate of location ($\sigma/\sqrt{n}$) is also very sensitive to outliers. The Hodges–Lehmann estimator is asymptotically normal with mean θ and variance τ^2/n where τ is the scale parameter given by

$$\tau = \frac{1}{\left[\sqrt{12}\int f^2(x)dx\right]}.$$

We use this estimation in our investigation. In (Koul, et al., 1987), a consistent, robust estimate of τ is developed which is very resistant to outliers. Another consistent estimator of τ is proportional to the length of a distribution free confidence interval for θ (see Chap.1 of (Hettmansperger, et al., 2011)). We use the latter estimate in the example below. The rank-based (1 − α)100% confidence interval for θ is

$$\widehat{\theta_W} \pm t_{\frac{\alpha}{2},n-1}\frac{\hat{\tau}}{\sqrt{n}}$$

where $t_{\frac{\alpha}{2},n-1}$ is the (1 – α/2) quantile of the t-distribution with (n − 1) degrees of freedom. Note that it is analogous to the usual t-interval, except for the location estimate and τ replacing σ. As with LS, the term $\frac{\hat{\tau}}{\sqrt{n}}$ is called the standard error of the estimate. We illustrate the robustness of this inference with the following example.

A one sample normal data is generated with mean=0 and standard deviation=1. Outliers are added to one data point to illustrate the effect outliers have on the location and spread estimates of both the LS and component-wise ranks

```
data data1;
do i=1 to 20;
      x=normal(1);
      output;
end;
run;

*** x1 indicates one 1σ  outlier added, x2 indicates one 2σ  outlier added;
proc univariate data=data1 normal alpha=0.05;
      var x x1 x2 x3 x4 x5 x6;
      QQplot x x1 x2 x3 x4 x5 x6 /normal(mu=est sigma=est color=red l=1);
run;

axis1 order = (-5 to 5 by 1) label = (height = 1.25 'One Sample Study')
      minor = (number = 1) ;
axis2 order = (1 to 2 by 1) label = (height = 1.25 'Outliers Added') ;
PROC BOXPLOT data=data1;
      PLOT var*id/ vaxis = axis1 haxis = axis2;
run;
```

The summary of analyses on the above data using the LS methods is in the below table.

	Raw data	One 1σ outlier	One 3σ outlier	One 6σ outlier
N	20	20	20	20
Mean	0.103	0.071	0.007	-0.089
Std. Deviation	0.6402	0.6964	0.8578	1.1607
Std. Error	0.1432	0.1557	0.1918	0.2595
2-sided 95% CI	(-0.197, 0.403)	(-0.255, 0.397)	(-0.395, 0.408)	(-0.632, 0.454)
% Change from original location estimate	-	31.07%	93.20%	186.41%

Table 1: Summary of least squares analysis

In (Nandakumar, et al., 2011), the SAS code is presented to calculate the robust location estimate (Hodges Lehman) and the robust spread (τ). For the benefit of the readers, the SAS program to estimate τ is also provided in the Appendix. On running these SAS programs on the above data, we see the effect outliers have had on the LS procedures in Table 2.

	Raw data	One 1σ outlier	One 3σ outlier	One 6σ outlier
N	20	20	20	20
Location	0.046	0.033	0.030	0.019
Spread	0.4324	0.5147	0.5774	0.6391
Std. Error	0.0967	0.1151	0.1291	0.1429
2-sided 95% CI	(-0.156, 0.248)	(-0.208, 0.274)	(-0.240, 0.300)	(-0.280, 0.318)
% Change from original location estimate	-	28.26%	34.78%	58.70%

Table 2: Summary of component-wise ranks analysis

There is a 186% change in the original value when one 6σ outlier was added and analyzed by LS methods. The influence of this outlier is mitigated by component-wise ranks where the effect of the same outlier leads to only a marginal change. Also, the impact on the spread, hence the 2-sided 95% CI is also seen. The length of the CI estimated by the LS procedure nearly doubles while that of the robust procedure remains smaller. To further summarize the effect of these outliers, we review the box plots (Figure 1) and the Q-Q plots (Figure 2) which give a graphical description of the data. The box plot shows that the addition of each outlier elongated the whiskers. A Q-Q plot is used to check whether the sample could have come from a specific target population (in our case, the Normal). The Q-Q plot (with one 6σ outlier) shows the influence of one outlier by violating the normality assumption.

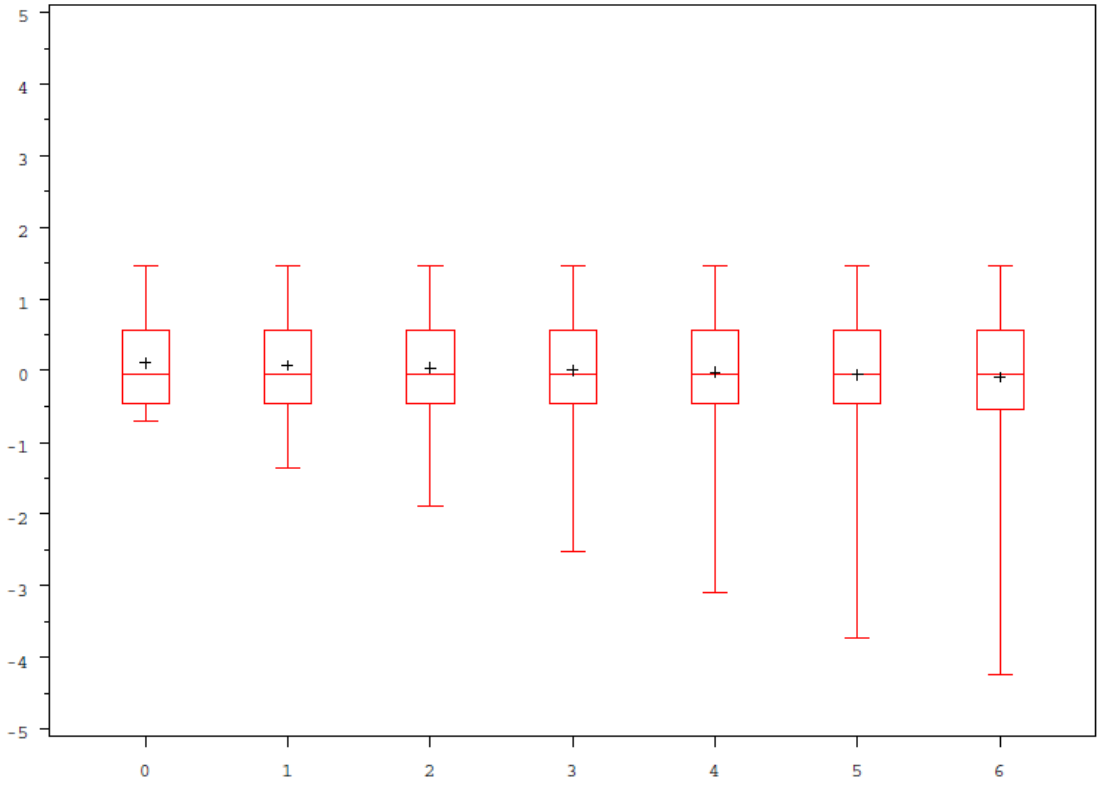

Figure 1: Box plot depicting the original normal sample and effects of one 1σ, 2 σ … 6σ outlier

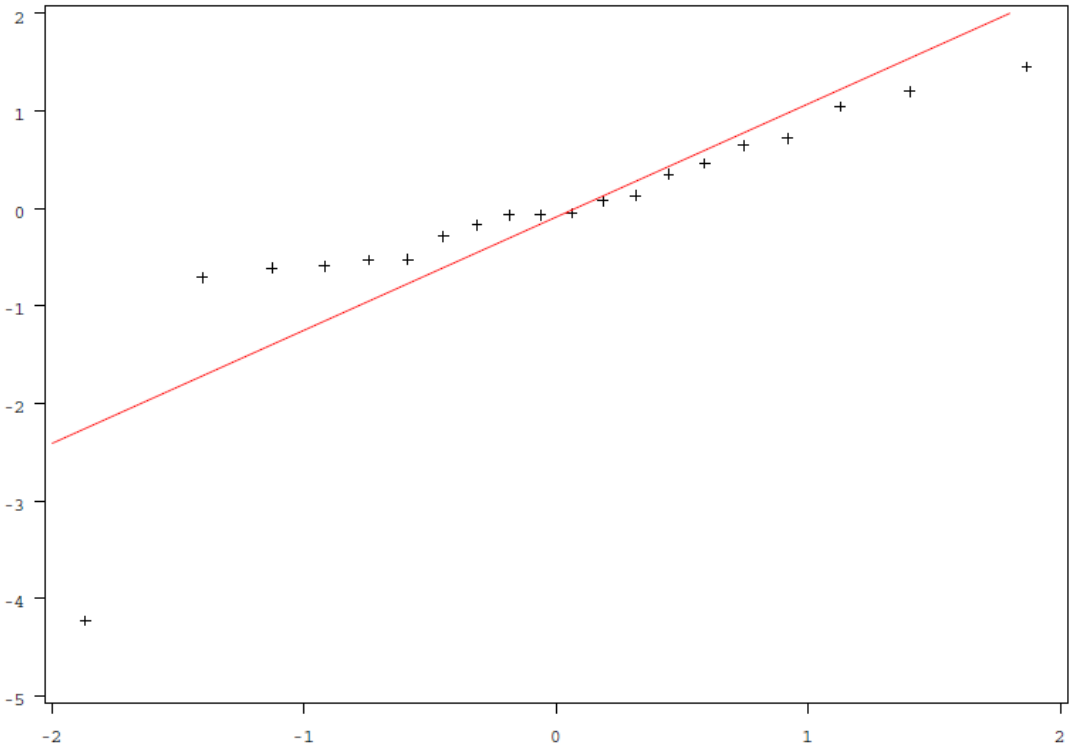

Figure 2: Q-Q plot of the sample quantiles vs the normal quantiles

AVERAGE BIOEQUIVALENCE ANALYSIS

The FDA (FDA, 2001) suggests using Schuirmann's (Schuirmann, 1987) two one-sided hypotheses for the analysis of PK parameters like AUC and Cmax. For this analysis, the PK parameters are log transformed (natural-log i.e. ln) and assumed to be normally distributed. In this paper, though, we are investigating the resistance to outliers of multivariate bioequivalence testing procedures. Hence, our hypotheses of interest are:

$$\text{Ho}: \left|\Delta\mu_{\text{AUC}}\right| > \ln(1.25) \text{ or } \left|\Delta\mu_{\text{Cmax}}\right| > \ln(1.25)$$

$$\text{Ha}: \left|\Delta\mu_{\text{AUC}}\right| \leq \ln(1.25) \text{ and } \left|\Delta\mu_{\text{Cmax}}\right| \leq \ln(1.25)$$

These hypotheses suggest testing procedures based on a confidence region of a multivariate estimator of the vector $\Delta\mu$. Two such tests were proposed in (Nandakumar, et al., 2011) and (Nandakumar, et al., 2009). One is based on the LS estimator of $\Delta\mu$ (componentwise means) and the second is based on componentwise Hodges-Lehmann estimates of $\Delta\mu$. Both the traditional and robust estimates are asymptotically bivariate normal with mean $\Delta\mu$ and different asymptotic covariance matrices. As discussed in (Nandakumar, et al., 2011) and (Nandakumar, et al., 2009), the usual ellipse type confidence region based on the estimates of these covariance matrices can easily be formulated. The only assumption made for the robust procedure is that the differences are symmetric and have finite Fisher's information.

The specifics of the model and the analysis procedure is outlined below

$$Y_{ijk} = \pi_i + \mu_k + s_{j(i)} + e_{ijk}$$

The components of the response Y_{ijk} is the log transformed AUC and log transformed Cmax for treatment k and subject j within sequence i, $s_{j(i)}$ is the random effect and e_{ijk} is the random error. Assume the random subject effect $s_{j(i)}$ to be independently and identically distributed as N (0, Φ1) and the random error e_{ijk}, also independently and identically distributed as N (0, Φo). Random effects $s_{j(i)}$ and e_{ijk} are mutually independent. The difference between the two drug responses eliminates the random subject effect (Stefanescu, et al.) as shown.

$$Y_{1j1} - Y_{1j2} = \mu_1 - \mu_2 + e_{1j1} - e_{1j2}$$
$$Y_{2j1} - Y_{2j2} = \mu_1 - \mu_2 + e_{2j1} - e_{2j2}$$

Based upon the hypothesis stated above, the rectangle bound by the co-ordinates (ln 0.80, ln 1.25) forms the rejection region. A 95% confidence ellipse (Johnson, et al., 1992) is plotted for the difference in the two treatment's ln AUC on the X-axis and ln Cmax on Y-axis. The least squares ellipse is constructed with the ordinary mean differences as estimates of location and least squares covariance structure for spread. Similarly, a robust confidence ellipse is constructed with the Hodges Lehmann's location estimate of differences as estimates of location and Component-wise Rank method for robust spread. The analysis procedure and the SAS program for the above is detailed in (Nandakumar, et al., 2011).

OUTLIER ANALYSIS IN BIOEQUIVALENCE STUDY

Twenty-six healthy subjects (Sub) were allocated randomly to one of two treatment sequences (Seq) in a two period, two treatment crossover design. The objective of the trial was to determine if the pharmacokinetic characteristics of drug A was bioequivalent to that of drug B. Subjects in the first sequences get dosed with drug A, followed by a washout period and then drug B. Subjects randomized to the second sequence get dosed first with drug B, a washout period and then drug A. This example dataset was sourced from the (Bradstreet, 1994). The snippet of the data is shown below.

		AUC		Cmax	
		(ng*hr/ml)		(ng/ml)	
Sub	Seq	A	B	A	B
4	A/B	1818.15	1425.94	767.01	786.82
5	A/B	961.01	1336.8	324.95	746.19
8	A/B	303.2	318.49	150.5	108.95
10	A/B	622.05	435.05	284.15	129.1
12	A/B	853.29	730.59	265.11	331.3
……					
20	B/A	633.89	367.36	348.44	310.44
21	B/A	1111.07	947.2	445.17	418.9
23	B/A	807.18	299.39	464.48	171.05
26	B/A	486.44	395.26	207.21	145.74

The least squares estimate of location for the treatment differences of ln AUC and ln Cmax are (0.111, 0.044) and the estimate of spread i.e. covariance matrix is $\sum_{LS} = \begin{bmatrix} 0.1499 & 0.1667 \\ 0.1667 & 0.3111 \end{bmatrix}$. The SAS code to obtain this is shown below

```
*** where X= Ln(AUC) for Drug A - Ln(AUC) for Drug B for that subject and
Y= Ln(Cmax) for Drug A - Ln(Cmax) for Drug B for that subject;

PROC CORR DATA=DATA1 COV NOCORR;
VAR X Y;
RUN;
```

Similarly, the Hodges Lehmann's location of the treatment differences for ln AUC and ln Cmax are (0.126, 0.082) and the robust covariance matrix is $\sum_{HL} = \begin{bmatrix} 0.1377 & 0.3026 \\ 0.3026 & 0.1514 \end{bmatrix}$. A 95% confidence ellipse is plotted with difference between drug A and drug B by log transformed AUC as the X-axis and log transformed Cmax as the Y-axis in Figure 3. Two confidence ellipses are plotted (a least squares ellipse and a robust ellipse) along with the rejection region.

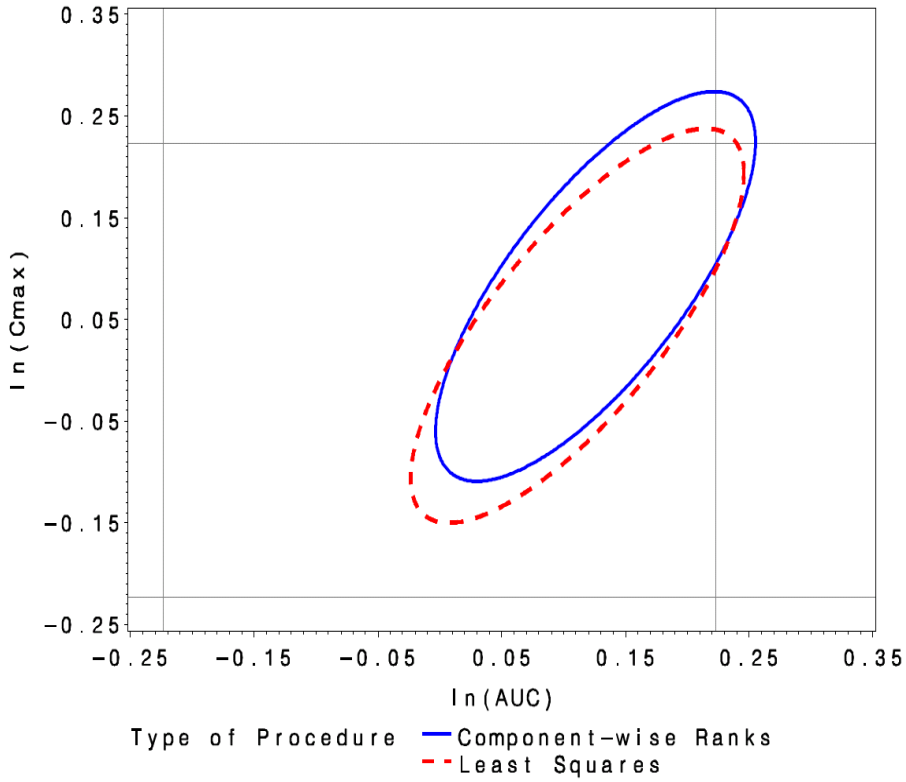

Figure 3: 95% Confidence Ellipses and the Rejection region.

For the benefit of the readers, the SAS program to calculate the robust estimate of spread (τ) from (Nandakumar, et al., 2011) is presented in the Appendix. The SAS program to generate the ellipses (Sinco, 2009) is shown below.

```
DATA PARAMDATA;
N= Total Number of subjects in the balanced Cross-over study;

       R_MSE= estimate of MSE based on robust estimates;
       R_X= Hodges-Lehman estimate of Ln(AUC) differences;
       R_Y= Hodges-Lehman estimate of Ln(Cmax) differences;
       R_S11= Covariance estimate of Ln(AUC) using component-wise ranks ;
       R_S22= Covariance estimate of Ln(Cmax) using component-wise ranks ;
       R_S12= Correlation estimate using component-wise ranks ;

       LS_MSE= LS estimate of MSE;
       LS_X= LS estimate of Ln(AUC) differences;
       LS_Y= LS estimate of Ln(Cmax) differences;
       LS_S11= Covariance estimate of Ln(AUC) using LS procedure;
       LS_S22= Covariance estimate of Ln(Cmax) using LS procedure;
       LS_S12= Correlation estimate using LS procedure;

       R_A=SQRT(R_S11);
       R_B=R_S12/R_A;
       R_C=SQRT(R_S22-(R_B**2));

       LS_A=SQRT(LS_S11);
       LS_B=LS_S12/LS_A;
       LS_C=SQRT(LS_S22-(LS_B**2));
RUN;

/*COMPUTES THE 95% CONFIDENCE ELLIPSE*/
DATA ELLIPSEDATA;
SET PARAMDATA;
       DO T=0 TO 6.285714285714285714285 BY .01; /* Ranges from 0 to 2*PI*/
              SINT=SIN(T);
              COST=COS(T);
              LS_D=SQRT(((N-1)*2*FINV(.95,2,N-2)*LS_MSE)/(N*(N-2)));
              R_D=SQRT(((N-1)*2*FINV(.95,2,N-2)*R_MSE)/(N*(N-2)));

              LS_XCOORD=LS_X+(LS_A*LS_D*COST);
              LS_YCOORD=LS_Y+LS_B*LS_D*COST+(LS_C*LS_D*SINT);

              R_XCOORD=R_X+(R_A*R_D*COST);
              R_YCOORD=R_Y+R_B*R_D*COST+(R_C*R_D*SINT);
              OUTPUT;
       END;
RUN;

PROC SQL;
CREATE TABLE ELLIPSEDATA2 AS SELECT *
       FROM
              (SELECT *, LS_YCOORD AS Y, LS_XCOORD AS X,
                   'Least Squares' AS ID
                     FROM ELLIPSEDATA
              OUTER UNION CORRESPONDING
              SELECT *, R_YCOORD AS Y, R_XCOORD AS X,
                   'Component-wise Ranks' AS ID
                     FROM ELLIPSEDATA);
QUIT;

SYMBOL1 v=none I = JOIN C=BLUE L = 1  width=5;
SYMBOL2 v=none I = JOIN C=RED  L = 20 width=5;
goptions htext=15pt noborder ftext=swisslu;

axis1 order=(-.25 to .35 by 0.1)  label=("ln(AUC)");
axis2 order=(-.25 to .35 by 0.1) label=(a=90 "ln(Cmax)");
TITLE " ";
```

```
PROC GPLOT DATA=ELLIPSEDATA2;
       PLOT Y*X=ID/haxis=axis1 vaxis=axis2
                   HREF=(-.2231 .2231) VREF=(-.2231 .2231);
       label ID="Type of Procedure";
RUN;
QUIT;
```

Since part of both of the ellipses fall outside the rejection region, failure to reject the null hypotheses concludes that the drugs A and B are not bioequivalent. A strong positive correlation between AUC and Cmax is also seen by the shape of the ellipses.

To illustrate the effect of outliers on the analysis, one subject's data was corrupted by including a 3σ outlier to each of ln(AUC) and ln(Cmax). The entire procedure above was rerun and the location and spread were re-estimated. The bivariate least squares estimate of location for the treatment differences of ln(AUC) and ln(Cmax) are (0.189, -0.025) and the estimate of spread i.e. covariance matrix is $\sum_{LS} = \begin{bmatrix} 0.1813 & -0.0348 \\ -0.0348 & 0.6416 \end{bmatrix}$. Similarly, the new Hodges Lehmann's estimate of location for the treatment differences for ln(AUC) and ln(Cmax) are (0.165, 0.082) and the robust covariance matrix is $\sum_{HL} = \begin{bmatrix} 0.1423 & 0.3002 \\ 0.3002 & 0.1033 \end{bmatrix}$. Two confidence ellipses are again plotted (a least squares ellipse and a robust ellipse) along with the rejection region.

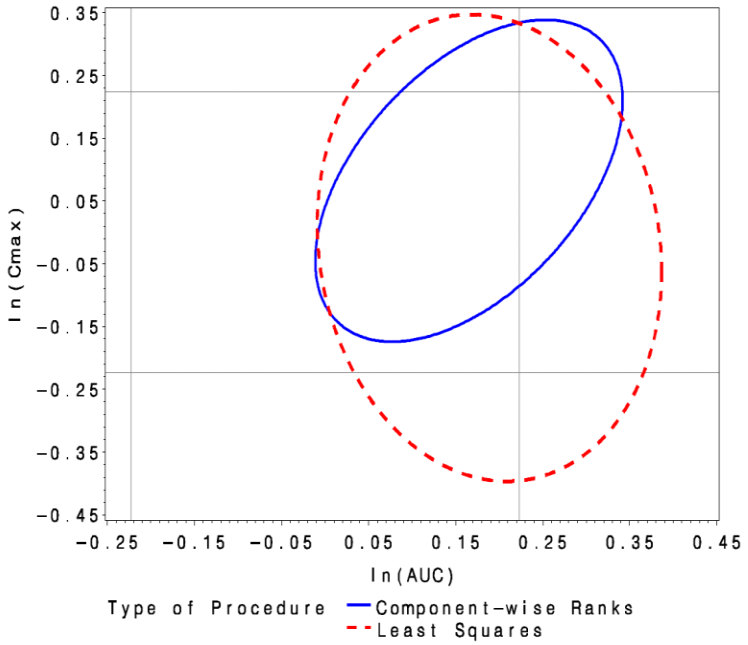

Figure 4: 95% Confidence Ellipses and Rejection region with outliers.

The effect of the outliers is clearly seen on the LS method in Figure 4. The shape of the LS ellipse is severely distorted and the positive correlation is now converted to a mild negative correlation. The influence of one data-point on the entire analysis seems to change the structure of the confidence ellipse. The robust ellipse still maintains the original structure and is not severely distorted by this outlier.

CONCLUSION

As in other datasets from clinical investigations, outliers frequently occur in bioequivalence investigations. These outliers have a profound influence on the LS procedures in analyzing small sample studies as shown above. Exclusion of such outliers is not an option as that could bias the analysis and lead to various safety violations. In this paper, we studied the effects of outliers on LS and robust confidence ellipse in analyzing bivariate bioequivalence. The shape and size of the ellipse is a good indicator of presence of outliers. The LS confidence ellipse has no resistance to outliers and a single outlying point can influence the entire result. However, the robust ellipse maintains its structural integrity and is more resistant to the outliers. The robust confidence ellipse is much smaller than the LS ellipse when analyzed on the data with outliers and, hence, its associated test for bioequivalence is more powerful than the LS test. Thus, in such small sample studies, we suggest that the robust methods also be considered along with the LS methods while testing for bioequivalence.

APPENDIX

For the benefit of the readers, the below program is sourced from (Nandakumar, et al., 2011).

```
%macro tau(data=, var=, qt=);
proc sql noprint;
create table data2 as select abs(a.&var.-b.&var.) as d&var.
      from &data. a, &data. b
      order by d&var.;
quit;

proc univariate data=data2 noprint;
      var d&var.;
      output out=pt pctlpre=p_ pctlpts=&qt.;
run;

%global &var.;
proc sql noprint;
select p_&qt./sqrt(&n.) into :tn
      from pt;

create table data3 as select a.*, b.p_&qt.,
      case when d&var. le &tn. then 1
             else 0
      end as gtn
      from data2 a, pt b;

select sum(gtn)/(&n.**2) into :gtn
      from data3;

select distinct (1/((sqrt(3)/&tn.)*&gtn.*sqrt((&n.-2)/&n.)))**2 into :&var.
      from data3;
quit;
%mend tau;

*** The variables in "Resid" data are:
Subject ID,
ResidX  = X-Hodges Lehmann estimate of X,
```

```
ResidY  = Y-Hodges Lehmann estimate of Y,
AResidX = Abs(ResidX),
AResidY = Abs(ResidY);

proc rank data=resid out=a_resid ties=mean;
      var aresidx aresidy;
      ranks aresx aresy;
run;

proc sql noprint;
select 3*sqrt(&residx.)*sqrt(&residy.)*sum(rankx*ranky*sign(residx)*
       sign(residy))/(&n.-1) into :cov
      from
      (select residx, aresidx, aresx/(&n.+1) as rankx, residy,
       aresidy, aresy/(&n.+1) as ranky
            from a_resid);
quit;

%tau (data=resid, var=residx, qt=90);
%tau (data=resid, var=residy, qt=90);
*** Tau is proportional to the length of a distribution free CI for location.
The quantile is associated with this CI (usually 0.90);
%put &residx.;
%put &residy.;
%put &cov.;
```

REFERENCES

Berger R and Hsu J Bioequivalence trials, intersection-union tests and equivalence confidence sets [Journal] // Statistical Science. - 1996. - pp. 283-302. - 11.

Bradstreet T E Favorite Datasets from Early Phases of Drug Research- Part 3 [Conference] // Proceedings of the Section on Statistical Education of the American Statistical Association. - 1994. - http://www.csc.villanova.edu/~short/Bradstreet/part3/part3-table5.html.

FDA Statistical approaches to establishing bioequivalence [Conference]. - [s.l.] : U.S. Department of Health and Human Services Food and Drug Administration-CDER, 2001.

Ghosh P and Gonen M Bayesian modeling of multivariate average bioequivalence [Journal] // Statistics in Medicine. - 2007.

Hettmansperger T. P and McKean Joseph W Robust Nonparametric Statistical Methods [Book]. - New York : Chapman & Hall, 2011.

Hollander Myles and Wolfe Douglas Nonparametric Statistical Methods [Book]. - 1999.

Johnson RA and Wichern DW Applied Multivariate Statistical Analysis [Book]. - [s.l.] : Prentice Hall, 1992.

Koul Hira L, Sievers Gerald L and McKean Joseph W An Estimator of the Scale Parameter for the Rank Analysis of Linear Models under General Score Functions [Journal] // Scandinavian Journal of Statistics. - 1987. - 2 : Vol. 14. - pp. 131-141 .

Nandakumar Srinand P and McKean Joseph W Application of Component-wise Rank Method to Bivariate Bioequivalence Case [Journal] // Pharmaceutical Programming. - [s.l.] : Maney Publishing, Dec 2011. - Accepted pending publication.

Nandakumar Srinand P and Mckean Joseph W Statistical procedures in Bioequivalence analysis [Report] = Statistical Procedures in Bioequivalence Analysis : Technical report / Dept. of Statistics.. - Kalamazoo : Western Michgan Univ., 2009.

Schuirmann D A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability [Journal] // J. Pharmacokinetic. Biopharm. - 1987. - pp. 657-80.

Sinco Brandy [Online] // SAS community. - 2009. - http://www.sascommunity.org/wiki/SAS_Code_for_Ellipse_Confidence_Region.

Stefanescu Catalina and V Devan Mehrotra A more powerful average bioequivalence analysis for the 2 x 2 crossover [Journal] // Communications in Statistics - Simulation and Computation. - Vol. 37. - pp. 212-221.

CONTACT INFORMATION

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

Dr. Srinand P. Nandakumar
Quintiles Inc.
6700 W. 115 St,
Overland Park, Kansas 66211
Work Phone: 913-708-6780
Email: srinand.pn@quintiles.com

Dr. Joseph W. McKean
Department of Statistics,
Western Michigan University,
Kalamazoo, Michigan 49008
Work Phone: 269-387-4541
Email: joseph.mckean@wmich.edu