

Methods of Outlier Detection in Bioequivalence Process: A Comparative Study

Amal F. Ghania*, Nagwa M. Albehery

Department of Mathematics, Insurance and Applied Statistics, Helwan University, Cairo, Egypt

Abstract Bioequivalence (BE) studies are becoming increasingly important due to the widespread production and use of generic drugs all around the world. These types of studies provide statistical evidence that two drugs or formulations, called the test drug and the reference drug, are therapeutically equivalent in terms of the rate and extent of absorption. Correctly identifying outliers in bioequivalence studies is very important because it may cause important differences in the results of the study. The existence of outliers may give the conclusion that formulations are not bioequivalent but in reality, they are bioequivalent. In the literature, several different methods have been proposed for detecting outliers in bioequivalent studies. In this paper, we compared four outlier detection tests: the estimated distance (ED) test, Hotelling T^2 test, the mean shift test, and principal component analysis (PCA) test. This is the first time that the PCA test has been compared with the Hotelling T^2 test and the mean shift test in the context of BE studies. The results of our simulation study showed that the ED test and PCA test outperformed the other two tests in detecting outliers.

Keywords Bioequivalence, Crossover design, Outliers, Likelihood distance, Estimated distance, Hotelling T^2 test, Mean shift test, Principal Component

1. Introduction

Bioequivalence (BE) studies are performed in the pharmaceutical industry to compare different forms of a drug or different drugs with the same active ingredient in terms of how quickly and completely the active ingredient is absorbed into the bloodstream. The most common pharmacokinetic parameters used to measure the rate and extent of absorption are the area under the blood or plasma concentration-time curve (AUC), the maximum blood or plasma concentration (C_{max}), and the time to the maximum concentration (t_{max}). To conclude that two drugs are bioequivalent, the Food and Drug Administration (FDA) requires statistical evidence that the true ratio of their average pharmacokinetic responses (such as AUC) is between 80% and 125% with 90% confidence. However, a commonly encountered problem in BE studies is that the data set may contain some extreme or outlying values/subjects. These outliers can occur due to laboratory errors, unusual subjects, or simply human transcription errors. Outlier detection is an important step in BE studies, as the presence of outliers can significantly impact the results of the study. Since the 1970s, there has been a lot of research on outlier detection methods.

However, most of this work has focused on outlier detection in linear regression models. For outlier detection in

BE trials, Chow and Tse (1990a) [2] proposed two procedures: the likelihood distance and the estimated distance. Liu and Weng (1991) [9] also proposed a method called two-sample Hotelling T^2 for detection of multiple outliers in bioavailability/bioequivalence data. Wang and Chow (2003) [18] proposed a test procedure called a mean shift model. Enachescu and Enachescu (2009) [7] used principal components to introduce a test for outlier detection in BE studies with crossover design.

In this paper, four methods will be applied for detecting outliers in bioequivalence studies and compared their ability to identify outliers. Specifically, the estimate distance test (ED), Hotelling's T^2 test, mean shift test and principal component analysis squared distance test (PCA) will be used.

It is worth noting that this is the first time that principal component analysis squared distance test (PCA) has been compared to the mean shift test and Hotelling T^2 test. In section 2, we briefly explain each one of the above-mentioned methods. In section 3, the four discussed procedures will be performed on crossover BE data set and a simulation study will be conducted to evaluate the performance of these four procedures in terms of power and level of significance. A concise summary is provided in section 4.

2. Four Outliers Detection Tests in Bioequivalence Studies with Crossover Design

In this section, four outlier detection tests in

* Corresponding author:

dr.amalfouad@commerce.helwan.edu.eg (Amal F. Ghania)

Received: Nov. 1, 2023; Accepted: Nov. 22, 2023; Published: Nov. 23, 2023

Published online at <http://journal.sapub.org/ajms>

bioequivalence studies with crossover design will be discussed. In subsection 2.1, likelihood distance and estimate distance will be explained. Hotelling T^2 test will be demonstrated in subsection 2.2. In subsection 2.3, the mean shift test will be revealed. Principal component analysis squared distance test (PCA) will be clarified in subsection 2.4.

2.1. Likelihood Distance and Estimates Distance

These methods are based on the likelihood function of the data.

The model for a $k \times k$ crossover design comparing f formulations of a drug product under the assumption that there are no period effects, and no formulation effects can be expressed as:

$$Y_{ij} = \mu + S_i + e_{ij}, j = 1, \dots, f; i = 1, \dots, n. \quad (2.1)$$

Where Y_{ij} is the response variable on the i th subject under the j th formulation,

μ is the overall mean, S_i is the random effect of the i th subject, e_{ij} is the error term. S_i and e_{ij} are assumed to be independently and normally distributed with means 0 and variances σ_s^2 and σ_e^2 respectively.

Chow and Tse (1990a) [2] introduced two test procedures, the likelihood distance (LD) and the estimates distance (ED) under the above model. Let $\theta = (\theta_1, \theta_2, \theta_3)'$ where:

$$\theta_1 = \mu, \theta_2 = \sigma_e^2 \text{ and } \theta_3 = \sigma_e^2 + f\sigma_s^2. \quad (2.2)$$

The log-likelihood function is given by:

$$L(\theta) = -\frac{nf}{2} \log 2\pi - \frac{n}{2} \log(\theta_2 \theta_3^{f-1}) - \frac{1}{2\theta_2} \sum_{i=1}^n \sum_{j=1}^f (Y_{ij} - \theta_1)^2 - \frac{f}{2} \left(\frac{1}{\theta_3} - \frac{1}{\theta_2} \right) \sum_{i=1}^n (\bar{Y}_i - \theta_1)^2. \quad (2.3)$$

The maximum likelihood estimator (MLE)

$\hat{\theta}$ of θ is given by

$$\hat{\theta}_1 = \bar{Y} = \frac{1}{nf} \sum_{i=1}^n \sum_{j=1}^f Y_{ij}, \hat{\theta}_2 = m_1, \hat{\theta}_3 = \frac{(n-1)m_2}{n}. \quad (2.4)$$

Where

$$m_1 = \frac{1}{n(f-1)} \sum_{i=1}^n \sum_{j=1}^f (Y_{ij} - \bar{Y}_i)^2, \quad \text{and} \quad m_2 = \frac{f}{(n-1)} \sum_{i=1}^n (\bar{Y}_i - \bar{Y})^2 \quad (2.5)$$

The LD test procedure can be expressed as $LD_i(\hat{\theta}) = 2[L(\hat{\theta}) - L(\hat{\theta}_{(i)})]$,

where $\hat{\theta}_{(i)}$ refers to the MLE of θ with deletion of the i th subject. Chow and Tse (1990 a) [2] demonstrated that as n tends to infinity, $LD_i(\hat{\theta})$ is asymptotically distributed as a chi-square variable with 3 degrees of freedom.

Thus, the i th subject is considered as an outlying subject if $LD_i(\hat{\theta}) > \chi_3^2(\alpha)$, where $\chi_3^2(\alpha)$ is the α th upper percentile point of a central chi-square distribution with 3 degrees of freedom.

They also introduced the estimates distance test, which is defined as:

$$ED_i(\hat{\theta}) = n^2(\hat{\theta} - \hat{\theta}_{(i)})\hat{\Sigma}^{-1}(\hat{\theta} - \hat{\theta}_{(i)}), \quad (2.6)$$

Where $\hat{\Sigma}$ is given by

$$\begin{bmatrix} \frac{\hat{\theta}_3}{f} & 0 & 0 \\ 0 & \frac{2\hat{\theta}_2^2}{f-1} & 0 \\ 0 & 0 & 2\hat{\theta}_3^2 \end{bmatrix} \quad (2.7)$$

Chow and Tse (1990 a) [2] also showed that the ED test follows a chi-squared distribution with 3 degrees of freedom as the sample size (n) increases. Therefore, the i th subject is considered as an outlying subject if $ED_i(\hat{\theta}) > \chi_3^2(\alpha)$.

It is worthwhile to note that the sample size for a bioavailability study is sometimes too small to apply asymptotic distributions of LD and ED of Chow and Tse's procedures, which is considered as a drawback of these procedures. Liu and Weng (1991) [9] applied the bootstrap technique to evaluate the sampling distribution of the LD and ED.

2.2. Hotelling T^2

To detect possible outlying subjects, Liu and Weng (1991) [9] introduced a procedure based on the two-sample Hotelling T^2 . The procedure works by first calculating the two-sample Hotelling T^2 statistics for each subject. The procedure identifies subjects that are likely to be outliers by looking at the order statistics of the T^2 statistics. Subjects that have T^2 statistics that are much larger or smaller than the majority of the T^2 statistics are likely to be outliers.

They adapted the following model:

$$Y_{ij} = B_j + \epsilon_{ij}, i = 1, \dots, n, j = 1, \dots, f \quad (2.8)$$

where $B_j = \mu + F_j$

Let $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{if})'$ be $f \times 1$ vector of the responses observed on subject i to f formulations. Thus, \mathbf{Y}_i are f -dimensional multivariate normal (MVN) random vectors with mean vector \mathbf{B} and covariance matrix \mathbf{A} , where

$\mathbf{B} = (B_1, \dots, B_f)'$ and

$$\mathbf{A} = \text{cov}(Y_{ij}, Y_{i'j'}) = \begin{cases} \sigma_j^2 & \text{if } i = i' \text{ and } j = j', \\ \sigma_{jj'} & \text{if } i = i' \text{ and } j \neq j', \\ 0 & \text{otherwise} \end{cases} \quad (2.9)$$

The formulation of the hypothesis for outlying subjects caused by a location shift can be written in the form of n sub-hypothesis as follows:

$$H_{0(i)}: \mathbf{Y}_i \sim \text{MVN}(\mathbf{B}, \mathbf{A})$$

$$H_{a(i)}: \mathbf{Y}_i \sim \text{MVN}(\mathbf{B} + \mathbf{A}_i, \mathbf{A}), i = 1, \dots, n$$

Hotelling T^2 statistics can be expressed as

$$T_i^2 = \frac{(n-2)D_i^2}{[(n-1)-D_i^2]}, \quad (2.10)$$

Where

$$D_i^2 = (\mathbf{Y}_i - \bar{\mathbf{Y}})' \mathbf{A}^{-1} (\mathbf{Y}_i - \bar{\mathbf{Y}}), i = 1, \dots, n \quad (2.11)$$

And \mathbf{A} is matrix of the sums of squares and cross products

computed from Y_1, \dots, Y_n , respectively.

Hypothesis $H_{0(i)}$ is rejected if

$$P\{\max_{1 \leq j \leq n} W_j^2 > T_{(i)}^2\} < \alpha$$

provided that $H_{0(n)}, H_{0(n-1)}, \dots, H_{0(i+1)}$ are rejected at the α level of significance, where $0 < \alpha < 1$, and (W_1^2, \dots, W_n^2) be a vector of n Hotelling T^2 statistics computed from a sample of size n from an f -dimensional multivariate normal with mean $\mathbf{0}$ and covariance matrix \mathbf{I}_f . The joint distribution of order statistics of T_i^2 is complicated. Instead, Liu and Weng (1991) [9] applied the Monte Carlo or bootstrap simulation to evaluate the sampling distribution of test statistics. They presented tables that give the 5% and 1% upper quantiles of the distribution of ordered T^2 statistics for $f = 2, 3$ and varies sample sizes.

2.3. Mean-Shift Model

Wang and Chow (2003) [18] developed a method for identifying both outlying subjects and outlying observations in bioequivalence studies using the mean-shift model proposed by Srikantan (1961) [17],

which is given as

$$Y_{ij} = \mu + S_i + F_j + \lambda_j \delta_{it} + e_{ij}, \quad j = 1, \dots, f; i = 1, \dots, n. \quad (2.12)$$

Where F_j is the fixed formulation effect $\delta_{it} = 1$ if $t = i$ and 0 otherwise. μ, S_i , and e_{ij} are as defined before. Based on this model, the hypothesis of interest can be expressed as:

$$H_{0(i)}: \lambda = (\lambda_1, \dots, \lambda_f)' = \mathbf{0}$$

Versus $H_{a(i)}: \lambda = (\lambda_1, \dots, \lambda_f)' \neq \mathbf{0}, i = 1, \dots, n$

The score test for the above hypothesis is given as

$D_i = n(f-1)T_{i1n} + nT_{i2n}$, where

$$T_{i1n} = \frac{(e_i - \bar{e}_i \mathbf{1})' (e_i - \bar{e}_i \mathbf{1})}{\mathbf{e}_s' (\mathbf{e}_s - \bar{e}_s \mathbf{1})' (\mathbf{e}_s - \bar{e}_s \mathbf{1})} \quad (2.13)$$

where $\mathbf{e}_i = (e_{i1}, \dots, e_{if})'$ is the residual vector of subject i after fitting model (2.12) and $\mathbf{1}$ is an $f \times 1$ vector of 1s.

$$T_{i2n} = \frac{\bar{e}_i^2}{\mathbf{e}_s' \bar{e}_s} \text{ and } \bar{e}_i = \frac{\mathbf{1}' \mathbf{e}_i}{f} \quad (2.14)$$

T_{i1n} can be used to detect an outlier observation in a certain formulation for a particular subject. T_{i2n} can be used to detect whether the subject is an outlier.

In addition, T_{i1n} and T_{i2n} are independent beta random variables, i.e.,

$$T_{i1n} \sim \text{beta}\left[\frac{(f-1)}{2}, \frac{(n-1)(f-1)}{2}\right] \\ \text{and } T_{i2n} \sim \text{beta}\left[\frac{1}{2}, (n-1)/2\right] \quad (2.15)$$

It follows that

$$D_i \sim n(f-1)\text{beta}\left[\frac{(f-1)}{2}, (n-1)(f-1)/2\right] + n\text{beta}\left[\frac{1}{2}, (n-1)/2\right] \quad (2.16)$$

Based on statistic D_i , subject i is declared as an outlier at the α significance level if $D_i = D_{(n)} = \max_{1 \leq i \leq n} D_i$ is greater than the α th upper percentile of the distribution of the

maximum of D_i . Wang and Chow produced tables that provide the 10%, 5%, and 1% upper percentiles of the distribution of the maximum of D_i .

A study by Ramsay and Elukum (2005) [14] compared the performance of Likelihood distance, estimates distance, Hotelling T^2 procedure, and the mean-shift model method. The study found that the estimate distance (ED) test was the most powerful method for detecting outliers. However, the study also found that the ED test was more likely to falsely identify subjects as outliers than the other methods.

2.4. Principal Component Analysis (PCA) Test

This method transforms the data into a new set of variables that are uncorrelated. In principal component analysis, we seek to maximize the variance of a linear combination of the variables. Enachescu and Enachescu (2009) [7] used principal components to introduce a test for outlier detection in bioequivalence studies with crossover design. They considered the weighted sum of square distance to zero of the projected data into principal factorial plane and mentioned that the observations with a square distance greater than threshold (m) may be considered as outliers where

$$m = f + 2 \sqrt{2 \sum_{i=1}^f \lambda_i^2} \quad (2.17)$$

λ_i is the Eigen value representing the variance of the i -th principal component, and f is number of formulations. Rasheed A. et al (2013) [16] compared the outliers identification test based on principal components with likelihood distance test and estimates distance test. They also observed the performance of these tests through a simulation study.

3. Application

In this section, a numerical example will be presented in subsection 3.1 and a simulation study will be conducted in subsection 3.2.

3.1. Numerical Example

The above four mentioned tests will be applied to a modified data set of Areas Under the Curve: AUC as given in Table (1). Rasheed, Junaid and Ahmad (2011) [15] employed this modified data set to verify their proposed PCA test. The original data set was taken from a bioequivalence study reported by (Purich, 1980) [13]. In the paper, twelve healthy volunteers were employed to examine the bioequivalence of two test tablet formulations in comparison to a reference solution.

As shown in Table 2, the estimates distance test identified subject 3 as an outlier. The mean shift test was also able to identify subject 3 as an outlier. The PCA test identified subject 3 as an outlier as the value of the squared distance of subject 3 is 19.0678, which is greater than the threshold value m in equation (2.17) of 10.4302. Hotelling T^2 was not able to identify subject 3 as an outlier as subject 3 has the highest Hotelling T^2 value of 27.0577, which is less than

the 5th percentile (upper quantile) of $T^2_{(12)}$ of 30.63. This means that we cannot reject the null hypothesis $H_{0(12)}$ at the 5% significance level, so we cannot conclude that subject 3 is an outlier, and the sequential testing procedure stops.

Table 1. AUC data for each subject

Subject	R	A	B
1	5.68	6.83	4.12
2	3.6	5.78	5.01
3	15.2	13.2	12.56
4	7.42	7.31	7.87
5	7.26	7.72	6.59
6	9.04	8.91	9.68
7	5.06	4.63	7.23
8	4.82	8.75	7.59
9	7.88	9.02	7.25
10	7.84	7.79	5
11	4.63	6.77	5.72
12	3.87	7.62	6.74

Table 2. Results of ED test, mean shift test, Hotelling T^2 test and PCA-squared distance test

Subject	ED	Mean shift test D_t	Hotelling T^2	PCA-squared distance
1	0.767	2.1492	3.8062	2.0180
2	1.5492	1.8653	1.2481	2.8927
3	45.0929	12.1977	27.0577	19.0678
4	1.1296	0.7183	1.0721	0.2056
5	0.9805	0.3175	0.3563	0.0750
6	1.3764	1.6258	1.9113	1.9922
7	0.7541	4.5830	18.4520	2.6620
8	1.8648	3.6263	7.3520	0.6348
9	0.7493	0.5673	0.6720	0.4083
10	0.8029	3.6216	6.1253	0.9686
11	0.7673	0.9866	0.5715	1.1423
12	1.661	3.7412	4.1868	0.9330

Table 3. Summary of the results of outlier detection from all four methods for AUC data set

Test	Result
ED test	Subj. 3 is an outlier
Hotelling T^2 test	No outlier
Mean shift test	Subj. 3 is an outlier
PCA test	Subj. 3 is an outlier

3.2. Simulation Study

In this subsection, a simulation study will be conducted to compare the power and Level α of the four tests that we introduced in Section 2. We generate random samples under a 2 x 2 crossover design using the procedure described by Wang and Chow (2003) [18]. First, random samples z_{i0} , and z_{ij} have been generated from a bivariate standard

normal distribution. Then, we obtain random samples x_{ij} and y_{ij} as follows:

$$x_{ij} = \sqrt{0.5}(z_{i0} + z_{ij}),$$

$$y_{ij} = cx_{ij} + \mu_j$$

Note that z_{i0} , and z_{ij} are used to capture the between-subject and within-subject variability, respectively. We set the mean of both the reference drug and the test drug to 100. The constant c is set to three different values, $\sqrt{200}$, $\sqrt{800}$ and $\sqrt{1800}$ to ensure that the intra-subject coefficient of variation (CV%) is 10%, 20%, and 30%, respectively.

The first observation is intentionally made to be an outlier by adding qs to the response of test formulation, where s is the standard deviation of the test formulation of a given sample and q is a constant that can be chosen to represent a mild, moderate, or extreme outlier. All simulations were carried out with 3000 simulation samples. Table 4 shows the percentage of detecting designated outlying subjects of the four procedures for sample size 16 and 20.

Table 4. Power of ED test, mean shift test, Hotelling T^2 test and PCA-squared distance test

CV	n	q	ED	Hotelling T^2	Mean shift test D_t	PCA-squared distance
10%	16	3	73.27%	18.83%	27.5%	49.47%
		5	99.23%	77.93%	85.53%	98.83%
		10	100%	100%	100%	100%
	20	3	74.9%	20.43%	26.47%	54.27%
		5	99.4%	83.77%	86.13%	99.53%
		10	100%	100%	100%	100%
20%	16	3	72.83%	18.07%	26.37%	47.87%
		5	99.37%	77.53%	85.23%	98.63%
		10	100%	100%	100%	100%
	20	3	73.6%	19.77%	25.53%	54.33%
		5	99.6%	84.17%	85.4%	99.47%
		10	100%	100%	100%	100%
30%	16	3	73.47%	17.57%	25.23%	46.87%
		5	99.33%	77.43%	85.53%	98.93%
		10	100%	100%	100%	100%
	20	3	75.73%	19.6%	27.57%	57.77%
		5	99.77%	84.47%	86.3%	99.7%
		10	100%	100%	100%	100%

Based on the results of our simulation study as shown in Table 4, we conclude that the estimates distance (ED) test is the most powerful method for detecting outliers among the four methods. PCA-squared distance test is the second powerful method followed by the mean shift test and finally the Hotelling T^2 test.

Figure 1 through Figure 6 portrayed the power of the four tests for sample size $n = 16, 20$ with coefficient of variation 10%, 20%, and 30% and $q=3,5,10$.

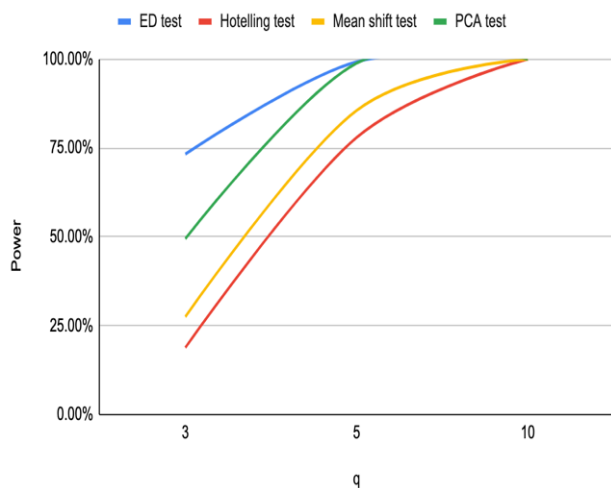


Figure 1. Power of ED test, Hotelling test, mean shift test and PCA test for $n=16$, $CV=10\%$

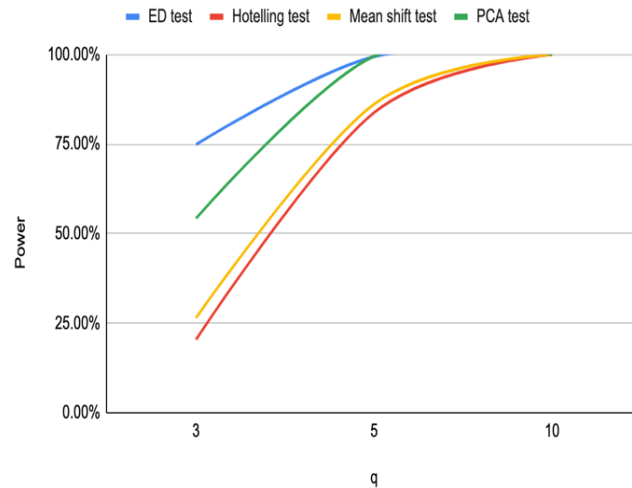


Figure 4. Power of ED test, Hotelling test, mean shift test and PCA test for $n=20$, $CV=10\%$

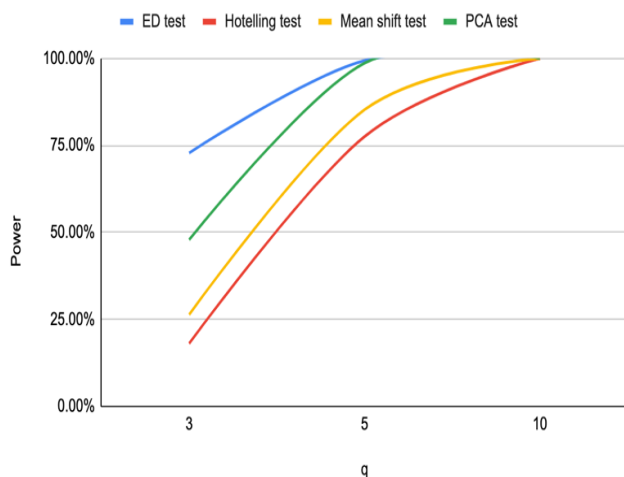


Figure 2. Power of ED test, Hotelling test, Mean shift test and PCA test for $n=16$ and $CV=20\%$

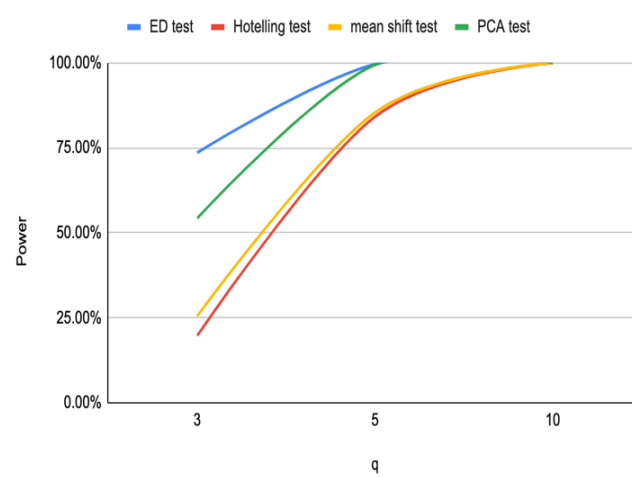


Figure 5. Power of ED test, Hotelling test, mean shift test and PCA test for $n=20$, $CV=20\%$

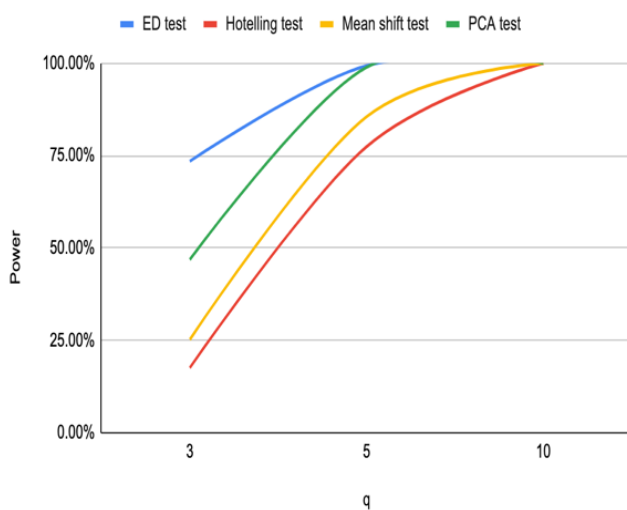


Figure 3. Power of ED test, Hotelling test, mean shift test and PCA test for $n=16$, $CV=30\%$

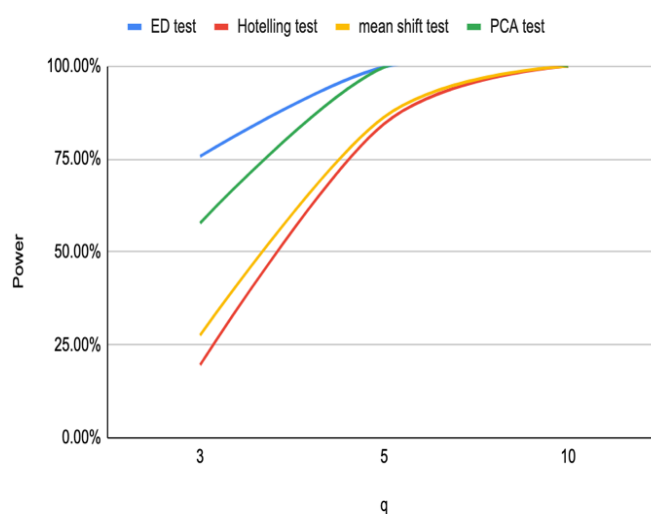


Figure 6. Power of ED test, Hotelling test, mean shift test and PCA test for $n=20$, $CV=30\%$

Level α of ED, mean shift test, Hotelling T^2 and PCA-squared distance test are presented in Table 5.

Table 5. Level α of ED, mean shift test, Hotelling T^2 and PCA-squared distance test

CV	n	q	ED	Hotelling T^2	Mean shift test D_i	PCA-squared distance
10%	16	3	0.0647	0.0047	0.0023	0.0207
		5	0.0287	0.0043	0.001	0.014
		10	0.002	0.0027	0.0003	0.011
	20	3	0.0723	0.0047	0.002	0.028
		5	0.0397	0.0037	0.0007	0.0197
		10	0.0023	0.0027	0	0.0133
20%	16	3	0.0687	0.003	0.0017	0.0213
		5	0.032	0.0027	0.0003	0.0133
		10	0.002	0.003	0	0.0097
	20	3	0.0743	0.0033	0.0023	0.0247
		5	0.0387	0.0023	0.0007	0.017
		10	0.003	0.0023	0	0.0123
30%	16	3	0.0643	0.0047	0.0027	0.02
		5	0.032	0.0033	0.0007	0.014
		10	0.0023	0.0027	0	0.011
	20	3	0.0687	0.0037	0.0017	0.023
		5	0.0327	0.0027	0.0003	0.016
		10	0.0023	0.0027	0	0.0113

Based on Table 5, the ED test is the most likely to flag subjects wrongfully as outliers, followed by the PCA squared distance test. The Hotelling T^2 test and the mean shift test are the least likely to flag subjects wrongfully as outliers. In fact, the mean shift test is much less likely to flag subjects incorrectly as outliers than the Hotelling T^2 test for all sample sizes and all levels of variability used in the simulation study.

4. Summary

There is no single best way to detect outliers in crossover design in bioequivalence studies. The best method to use will depend on the size of the study, the distribution of the data, and how sensitive you want the test to be. In general, it is best to use multiple outlier detection methods and consider the results of all methods before deciding whether to exclude a subject from the study.

The ED test and the PCA-squared distance test are two generally recommended methods that are both powerful and conservative. This means that they are likely to detect real outliers, but they are also unlikely to flag good data points as outliers.

ACKNOWLEDGEMENTS

We would like to thank Dr. Ibrahim Hassan Ibrahim, professor of applied statistics at the department of mathematics,

applied statistics and insurance, Faculty of Commerce and Business Administration, Helwan University, for referring us to the research topic and for his helpful comments on the initial draft of this research.

REFERENCES

- [1] Anderson, S., and Hauck, W.W., (1984), "A new statistical procedure for testing equivalence in two-group comparative bioavailability trials," *Journal of Pharmacokinetics Biopharmaceutics*, 12, 83 – 91.
- [2] Chow, S.C. and Tse, S.K. (1990a). Outlier detection in bioavailability/bioequivalence studies. *Statistics in Medicine*, 9, 549–558.
- [3] Chow, S.C., and Liu, J. P., (1992), *Design and Analysis of Bioavailability and Bioequivalence Studies*, New York: Marcel Dekker, Inc.
- [4] Chow, S.C., Shao, J., and Wang, H., (2001a), "Statistical Tests for Population Bioequivalence," Technical report, Dept. of Stat., University of Wisconsin, Madison.
- [5] Chow S-C and Liu J-P (2009). *Design and analysis of bioavailability and bioequivalence studies* (3rd ed.). New York: Dekker.
- [6] Davison, A., and Hinkley, D., (1997), *Bootstrap Methods and their Applications*, UK: Cambridge University Press.
- [7] Enachescu D, Enachescu C. (2009). A new approach for outlying records in bioequivalence trials. Paper presented at the The XIII International Conference "Applied Stochastic Models and Data Analysis".
- [8] FDA, (2001), "Guidance for industry on statistical approaches to establishing bioequivalence. Center for Drug Evaluation and Research, Food and Drug Administration, Rockville, Maryland 2001.
- [9] Liu, J. P. and Weng, C. S. (1991). Detection of outlying data in bioavailability/bioequivalence studies. *Statistics in Medicine*, 10, 1375-1389.
- [10] Liu, and Jen-Pei (1998), "Statistical Evaluation of Individual Bioequivalence," *Communications in Statistics, Theory and Methods*, V27, 6, PP 1433-1451.
- [11] Mogos, B (2013). Exploratory data analysis for outlier detection in bioequivalence studies. *Biocybernetics and Biomedical Engineering*, Volume 33, Issue 3, P. 164-170.
- [12] O'Brien (1984). Procedures for comparing samples with multiple endpoints. *Biometrics* 40, 1079-1087.
- [13] Purich E. (1980). *Bioavailability/Bioequivalence Regulations: An FDA Perspective in Drug Absorption and Disposition: Statistical Consideration*, K.S Albert, Edn. American Pharmaceutical Association, Academy of Pharmaceutical Sciences. Washington, DC, 15-137.
- [14] Ramsay T, Elcum N. (2005) A comparison of four different methods for outlier detection in bioequivalence studies. *J. Biopharm Stat*; 15(1): 43-52.
- [15] Rasheed, Junaid and Ahmad (2011). Detection of Outliers in

- Bioequivalence Studies Data Analysis with Williams Design. Proc. 8th International Conference on Recent Advances in Statistics. Lahore, Pakistan – February 8-9, 2011, 9-18.
- [16] Rasheed A. et al. (2013). Comparison of Outlier Detection Methods in Crossover Design Bioequivalence Studies, Journal of Pharmacy and Nutrition Sciences, 3, 163-170.
- [17] Srikantan, K.S. (1961). Testing a single outlier in a regression model. Sankhya. Ser. A vol. 23, No. 3: p251-260.
- [18] Wang, W. and Chow, S.C. (2003). Examining outlying subjects and outlying records in bioequivalence trials. Journal of Biopharmaceutical Statistics, 13, 43–56.