DOI: 10.5455/bcp.20120928030834

Examination of Outliers in Bioequivalence Studies

Durdu Karasoy[1], Gulcan Daghan[2]

ÖZET:

Biyoeşdeğerlik çalışmalarında aykırı değerlerin incelenmesi

Amaç: Biyoeşdeğerlik çalışmalarında aykırı değerlerin yer alması, çalışma sonuçları üzerinde önemli farklılıklara yol açan bir problemdir. Aykırı değerlerin varlığı formülasyonlar gerçekte eşdeğer iken eşdeğer değildir ya da gerçekte eşdeğer değil iken eşdeğerdir sonucuna götürebilir. Bu çalışmada, biyoeşdeğerlik çalışmaları ve aykırı değerler hakkında genel bilgi verilerek, biyoeşdeğerlik çalışmalarında kullanılan aykırı değer inceleme yöntemleri incelenmiştir.

Yöntemler: Çalışmaya 23 katılımcı dahil edilmiştir. Aykırı değerleri belirlemek için olabilirlik uzaklıkları, tahmin uzaklıkları, Hotelling T^2, ortalama-kayma testleri ve Liu ve Weng'in artıklar yöntemleri kullanılmıştır.

Bulgular: Test ve referans ilaçlar sağlıklı katılımcılara verilmiştir. Belirlenen zaman noktalarında alınan kan örneklerine göre analiz yapılmıştır. Doğruluğu kanıtlanmış bir yöntemle ölçüm sıvı içindeki madde yoğunluk değerleri belirlenmiş ve istatistiksel çözümleme aşamasına geçilmiştir. Aykırı değer inceleme yöntemleri kullanılarak 6. birimin aykırı değer olduğu görülmüştür.

Sonuç: Aykırı değer olarak bulunan 6. birimin varlığının test ilaç düzeyinde yapılan hesaplama sonuçlarını farklılaştırarak test ve referans ilacın benzerliklerini azaltıcı etki gösterdiği sonucuna varılmıştır.

Anahtar sözcükler: Biyoeşdeğerlik, biyoyararlanım, aykırılıklar

Klinik Psikofarmakoloji Bülteni 2012;22(4):307-12

ABSTRACT:

Examination of outliers in bioequivalence studies

Objective: The existence of outliers in bioequivalence studies is a problem causing important differences in the results of the study. The presence of outliers may lead to the conclusion that formulations are not bioequivalent even if in reality they are or that they are bioequivalent even if in fact they are not. In this study by giving general information about bioequivalence studies and outliers, the methods to examine outliers are analyzed.

Methods: Twenty-three participants were included in this study. The likelihood distance test, estimates distance test, Hotelling T^2 test, mean-shift test and Liu and Weng's residual test were used for examining outliers.

Results: The test and reference drugs were given to healthy participants. The participants were examined by obtaining blood samples at specific time points. The density values of the element in the blood were determined by a reliable method. By the methods such as likelihood distance test, estimates distance test, Hotelling T^2 test, mean-shift test and Liu and Weng's residual test, the sixth observation was detected as an outlier.

Conclusion: It was concluded that the presence of the outlier reduces the similarities between the test and reference drugs.

Key words: Bioequivalence, bioavailability, outliers

Bulletin of Clinical Psychopharmacology 2012;22(4):307-12

[1]Assoc. Prof. Dr., Hacettepe University, Faculty of Science, Department of Statistics, Ankara - Turkey
[2]MSc, Ziraat Bank, İstanbul - Turkey

Yazışma Adresi / Address reprint requests to: Assoc. Prof. Dr. Durdu Karasoy, Hacettepe University, Faculty of Science, Department of Statistics, 06800, Beytepe, Çankaya / Ankara - Turkey

Telefon / Phone: +90-312-297-7900/121

Faks / Fax: +90-312-297-7913

Elektronik posta adresi / E-mail address: durdu@hacettepe.edu.tr

Gönderme tarihi / Date of submission: 7 Mayıs 2012 / May 7, 2012

Kabul tarihi / Date of acceptance: 28 Eylül 2012 / September 28, 2012

Bağıntı beyanı:
D.K., G.D.: Yazarlar bu makale ile ilgili olarak herhangi bir çıkar çatışması bildirmemişlerdir.

Declaration of interest:
D.K., G.D.: The authors reported no conflict of interest related to this article.

INTRODUCTION

Two drugs are bioequivalent if they contain the same active substance in the same or similar pharmaceutical form at the same molar dose and if the rate and extent of absorption from the administered dosage form are the same within specific limits.

Data sets usually contain some extraordinary measures. It is one of the most commonly encountered problems in bioequivalence studies. These extraordinary measures are called outliers and may affect bioequivalence results. It might be mistakenly concluded that formulations are not bioequivalent even if in fact they are, or bioequivalent if in reality they are not. Hence outliers should be analyzed carefully to avoid these mistaken conclusions.

Since including or excluding outliers in bioavailability/

bioequivalence studies causes different inferences, many researchers have worked on this topic in the past (1-4).

Chow and Tse (1990) have proposed two procedures for detection of a possible outlier in bioequivalence studies based on Cook's likelihood distance and the estimates distance (3). Liu and Weng (1991) examined procedures based on Hotelling T^2 statistics and residuals for the same purpose (5). Wang and Chow (2003) introduced a general test procedure based on a mean-shift model (6).

In this study, we aimed to introduce methods that may be used to examine outliers and apply them to a real data set of a real bioequivalence analysis.

METHODS

Important Concepts and Outliers in Bioequivalence Studies

Bioequivalence is mainly considered when there is the same active substance within different drugs. Studies investigating the bioavailability of two pharmaceutically bioequivalent drugs whether they are the same or not, after administering the same molar dose, are called bioequivalence studies (7).

"The reference drug" is the compound that was developed first and obtained approval to be marketed to treat a condition after it showed satisfactory efficacy and safety. After the patent/data protection duration of the innovating firm on the reference drug has expired, the products that are introduced to the market by other pharmaceutical companies, are called "equivalent drugs", meaning they are pharmaceutical equivalents or alternatives to the original medicine. In bioequivalence studies, these drugs are called the "test" products.

The most widely used pharmacokinetic parameters in bioequivalence studies are as follows: Area under the concentration curve (AUC), maximum plasma concentration (C_{max}), and time to reach maximum plasma concentration (t_{max}).

For two formulations to be considered as bioequivalent, the FDA regulations require that the limits of a 90% confidence interval for the ratio of the geometric means of the two kinetic responses (AUC and C_{max}) should be within the interval of [80%-125%] (8).

The average, population, and individual bioequivalence methods are used in bioequivalence studies (9-11).

Generally the utilized design is a 2x2 crossover, which takes into account period and formulation effects. The likelihood function for a 2x2 crossover design is

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

(1)

where θ parameter vector is $(\theta_1, \theta_2, \theta_3)^T$. The maximum likelihood estimator (MLE), which makes maximum the likelihood function in Equation (1), is $MLE(\hat{\theta}) = (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3)^T$. Here,

$$\hat{\theta}_1 = \overline{Y} = \frac{1}{nf}\sum_i\sum_j Y_{ij}, \quad \hat{\theta}_2 = \frac{1}{n(f-1)}\sum_i\sum_j (Y_{ij}-\overline{Y}_i)^2 \text{ and } \hat{\theta}_3 = \frac{f}{n}\sum_i(\overline{Y}_i-\overline{Y})^2$$

(2) where, f is the number of formulations, n is the number of subjects, and Y_{ij} expresses the measurement value for i^{th} subject, j^{th} formulation (3).

Outliers can be described as data that are incompatible with other data in a data set. Incompatibility between data is an undesirable condition because of it has an important effect on statistical studies (12-14).

Outliers in bioavailability/bioequivalence studies may occur because of fluctuation in laboratory tests, for example an extraordinary subject having too high or too low bioavailability, or humanistic measuring faults.

The methods, that are used for detecting outliers in bioequivalence studies are explained below.

Likelihood Distance Test

The likelihood distance test (LD) is one of the tests used for determining outliers or influential observations in a bioequivalence study and was developed by Cook and Weisberg (1982) based on likelihood distances.

The LD statistic for the i^{th} subject is twice of the difference between the log likelihood evaluated by using the estimates from all of the subjects and from the estimates obtained after deleting the i^{th} subject. This is expressed as: $LD_i(\hat{\theta}) = 2[L(\hat{\theta}) - L(\hat{\theta}_{-i})]$, where, $\hat{\theta}_{-i}$ is the maximum likelihood estimator of θ obtained by deleting the i^{th} subject from the data.

Asymptotically, $LD_i(\hat{\theta})$ is distributed as a chi-square statistic with three degrees of freedom. If $LD_i(\hat{\theta}) > \chi^2_{(3,\alpha)}$, the i^{th} subject is an outlier (3).

Estimates Distance Test

The second method for examining the effect of i^{th} subject in the study is based on the difference in the parameter estimates arising from the deletion of the i^{th} subject.

Estimates distance test (ED) is similar to the LD because of accounting the distances of parameter estimates in the case of the presnce or absence of the i^{th} subject. The ED statistic is $ED_i(\hat{\theta}) = f^2(\hat{\theta} - \hat{\theta}_{-i})^t \hat{\Sigma}^{-1}(\hat{\theta} - \hat{\theta}_{-i})$ (3) where $\hat{\Sigma}^{-1}$ is the MLE of the variance matrix.

Chow and Tse (1990) showed that $ED_i(\hat{\theta})$ is asymptotically distributed as a chi-square with three degrees of freedom. If $ED_i(\hat{\theta}) > \chi^2_{3,\alpha}$, the i^{th} subject is an outlier (3).

Hotelling T^2 Test

Liu ve Weng (1991) suggested a procedure based on the order statistics of the two sample Hotelling T^2 (HT) statistics to identify possible outlying subjects. Let $Y_i = (Y_{i1},..., Y_{if})^t$ be the response variables vector observed on the f^{th} formulation of the i^{th} subject. $\overline{Y}$ and A are the sample mean and the matrix of the sums of squares by cross products of $Y_1, Y_2, \ldots, Y_n$, respectively.
Define

$$D_i^2 = (Y_i - \overline{Y})^t A^{-1}(Y_i - \overline{Y}) \quad (4)$$

In this case, HT statistic for the i^{th} subject is given by

$$T_i^2 = \frac{(n-2)D_i^2}{(\frac{n-1}{n} - D_i^2)} \quad (5)$$

The value, which is obtained from Equation (5), is compared with the critical value to decide whether or not the i^{th} subject is an outlier (5).

Liu and Weng's Residuals Test

Liu and Weng (1991) suggested a test for determining outliers benefiting from the means of the formulation. Let $\overline{Y}_{.j}$ be the j^{th} formulation mean. Studentized residuals are obtained as r_{ij} (i=1, 2, …, n; j=1, 2, .., f). Maximum values of standardized residuals are compared with a critical value to decide whether or not the i^{th} subject is an outlier (5).

Mean Shift Test

Wang and Chow (2003) developed a test procedure based on the mean-shift model for the i^{th} subject's response to the j^{th} formulation. They defined two quantities such as

$$T_{1n} = \frac{(e_t - \bar{e}_t 1)'(e_t - \bar{e}_t 1)}{\sum_{s=1}^{n}(e_s - \bar{e}_s 1)'(e_s - \bar{e}_s 1)} \quad (6)$$

and

$$T_{2n} = \frac{\bar{e}_t^2}{\sum_{s=1}^{n}\bar{e}_s^2} \quad (7)$$

In Equations (6) and (7); e_t is the vector of residuals for the t^{th} subject, $\bar{e}_t$ represents the mean of e_t, 1 is a vector whose members are 1. Then, they showed that the test statistic

$$D_t = nT_{1n} + nT_{2n} \quad (8)$$

can be used to test whether the t^{th} subject is an outlier, or not (6).

To test the practicality of outlier examination methods by using 2x2 crossover method that is applied to two different pharmacokinetic profiles' bioequivalence by Novagenix Bioanalytical Drug R.D. Center we used a real data set taken from 23 volunteers at 21 different time points.

The participants in this study provided voluntary informed consent. This study was approved by a local institutional review board and was conducted in accordance with the Declaration of Helsinki.

Microsoft Excel 2007 was used for statistical analysis.

RESULTS

After administering the test and reference drugs, involving the same active substance, to 23 subjects in the study, measures were taken from the blood at 21 time points. The AUC parameters, which were calculated for test and reference drugs for any subject, are seen in Figure 1.

When Figure 1 is analyzed, it may be seen that a minimum value was measured for the 6^{th} subject and this subject is different from the other subjects. It may be speculated that the 6^{th} subject is an outlier, because outliers in the test drug levels have a converter effect on the results

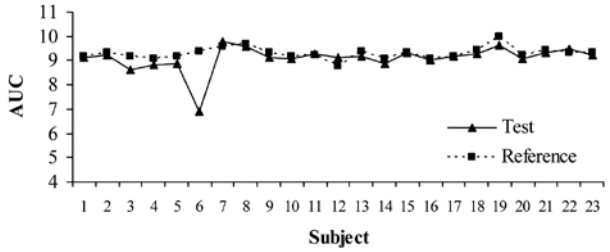

Figure 1: AUC values for test and reference drugs

of the study. Whether or not the 6th subject was an outlier was evaluated by using the outlier examination methods as previously described.

Likelihood Distance Test

Estimators of the maximum likelihood function in Equation (1) are obtained as $\hat{\theta}_1$= 9.192; $\hat{\theta}_2$= 0.1569, and $\hat{\theta}_3$= 0.2056 by using Equation (2). To determine, whether the 6th subject is an outlier or not, the 6th subject was deleted from the data and estimators were calculated as $\hat{\theta}_{1\text{-}6}$= 9.239; $\hat{\theta}_{2\text{-}6}$= 0.0241, and $\hat{\theta}_{3\text{-}6}$= 0.1157. For $\hat{\theta}_1$,$\hat{\theta}_2$, $\hat{\theta}_3$, the likelihood function value in Equation (1) is $L(\hat{\theta})$= -25.786. When the 6th subject was deleted from the data, the likelihood function statistic becomes $L(\hat{\theta}_{-6})$=-84.17715. Then, the LD statistic of the 6th subject as obtained as $LD_6(\hat{\theta}) = 2(-25.786 + 84.17715) = 116.7823$.

As $LD_6(\hat{\theta}) = 116.7823 > \chi^2_{3,0,05}$ =7.815, the 6th subject was an outlier.

Estimates Distance Test

For this test, the ED statistic was calculated using Equation (3). The 6th subject's ED statistic was detected as 32.8991. As $ED_6(\hat{\theta}) = 32.8991 > \chi^2_{3,0,05} = 7.815$, the 6th subject was an outlier.

Hotelling T^2 Test

For testing whether or not the 6th subject was an outlier, the D_6^2 statistic for the 6th subject was calculated as 0.93 by Equation (4). This value was used in Equation (5) and the T_6^2 statistic for the 6th subject was calculated as 751.15. This value was greater than the critical value of 21.88 for T^2 order statistics at α=0.05 significance level with a sample size 23. Therefore, the 6th subject was an outlier.

Liu and Weng's Residuals Test

In this test, outliers were determined by using residuals (e_i). Standardized residuals (r_i) were obtained by dividing calculated residuals by the standard deviation (0.4183) and these values were compared with the critical value of 2.83 at α=0.05, n=23, and p=3.

The maximum residual value (5.223) was obtained at the sixth subject's test drug level. The sixth subject was an outlier because this value was greater than the critical value of 2.83.

Mean Shift Test

In this study, T_{1n}, T_{2n}, and D_t statistics were calculated using Equation (6), Equation (7), and Equation (8) for all subjects, respectively. The maximum D_t value (29.65) was found for the test drug level of the 6th subject. The 6th subject was an outlier because this value was greater than critical value of 10.67 at α= 0.05 significance level with n=23.

Bioequivalence Examination

From all of the tests for the analysis of outliers, it was concluded that the sixth subject was an outlier. The effect of this subject on the result of the bioequivalence study was examined by the average bioequivalence method.

Ninety percent confidence intervals of test/reference proportions of AUC, which are obtained using the data with and without the sixth subject by plasma concentrations in the subjects' blood, are given at Table 1.

When Table 1 is examined, it is seen that decision parameter maximum and minimum limits of AUC's proportions are within the 80%-125% limits. The inclusion of the sixth subject in the study did not change the bioequivalence result but decreased the similarities of test and reference drugs by changing the results at the test drug level, as seen by the means of test/reference drug ratio (p=0.4>0.05). When the sixth subject is deleted, the mean of the test/reference drug ratio became closer to 1.

Table 1: Bioequivalence study

Pharmacokinetic variable	with 6th subject		without 6th subject	
	90% lower limit	90% upper limit	90% lower limit	90% upper limit
AUC_T / AUC_R	0.9581	0.9989	0.9815	0.9975
Mean (AUC_T / AUC_R)	0.979		0.990	

DISCUSSION

Metzler and Huang (1983) have demonstrated that the existence of outliers has no significant effect on bioequivalence studies. Chow and Tse (1990) and Bolton (1991) exemplified that the presence and absence of outliers have different results. Chow and Tse (1990) proposed two methods for determining probable outliers based on Cook's likelihood distance and estimates distance. Liu and Weng (1991) examined two methods based on Hotelling's T^2 statistic and residuals for the same reason. Wang and Chow (2003) handled a more general method based on the mean-shift model. Ki et al. (1995) examined the effect of outliers to bioequivalence. Ramsay and Elkum (2005) compared different methods which are used to examine outliers in bioequivalence studies.

In this study, likelihood distance, estimates distance, Hotelling T^2, mean-shift test, and Liu and Weng's residuals method have been applied on a real data set and it is concluded that the sixth subject is an outlier.

For examining the effect of the sixth subject on bioequivalence results, the average bioequivalence method proposed 2x2 crossover designs are used in both the presence and absence of the sixth subject in the data. It was concluded that the sixth subject did not change the bioequivalence results, but the similarities of the drugs were less in the study when the sixth subject was involved.

From the results of the average bioequivalence method, it should not be thought that the sixth subject had no effect on the results, because the sixth subject had a masking effect on the mean of the test drug and this effect should be taken into the consideration in the comparison.

Acknowledgements

The authors are grateful to Onursal Sağlam from Novagenix Bioanalytical Drug R.D. Center, who provided application data and also thanks to the editor and referees for their contributions.

References:

1. Rodda BE. Bioequivalence of solid oral dosage forms. A Presentation to U.S. Food and Drug Administration Hearing on Bioequivalence of Solid Oral Dosage Forms, Pharmaceutical Manufacturers Association. September 29 - October 1, Section III, 1986, p. 12-5.
2. Metzler CM, Huang DC. Statistical methods for bioavailability and bioequivalence. Clin Res Pr Drug Regul Aff 1983;1(2):109-32.
3. Chow SC, Tse SK. Outlier detection in bioavailability/bioequivalence studies. Stat Med 1990;9(5):549-58.
4. Bolton S. Outliers–examples and opinions. Presented at bioavailability/bioequivalence: Pharmacokinetic and Statistical Considerations sponsored by Drug Information Association, Bethesda, Maryland, August, 1991.
5. Liu JP, Weng CS. Detection of outlying data in bioavailability/bioequivalence studies. Stat Med 1991;10(9):1375-89.
6. Wang W, Chow SC. Examining outlying subjects and outlying records in bioequivalence trials. J Biopharm Stat 2003;13(1):43-56.
7. Kayaalp SO. Klinik Farmakolojinin Esasları ve Temel Düzenlemeler. 2. Baskı, Ankara: Hacettepe Taş Kitapçılık; 2001. p. 305-47.
8. Food and Drug Administration (FDA). Guidance for industry: Bioavailability and bioequivalence studies for orally administered drug products-general considerations. U.S. Department of Health and Human Services, Center for Drug Evaluation and Research, Rockville, Maryland, March, 2003.
9. Karasoy D, Gündüz İ. Statistical methods in bioequivalence studies and an application. Turkiye Klinikleri J Med Sci 2008;28(5):653-62. (Turkish).
10. Anderson S, Hauck WW. Consideration of individual bioequivalence. J Pharmacokinet Biopharm 1990;18(3):259-73.
11. Chen ML, Patnaik R, Hauck WW, Schuirmann DJ, Hyslop T, Williams R. An individual bioequivalence criterion: Regulatory considerations. Stat Med 2000;19(20):2821-42.
12. Barnett V, Lewis T (editors). Outliers in statistical data. Third ed., John Wiley & Sons; 1994. p. 31-52.

13. Ki FYC, Liu JP, Wang W, Chou SC. The impact of outlying subjects on decision of bioequivalence. J Biopharm Stat 1995;5(1):71-94.

14. Beckman RJ, Cook RD. Outlier...s (with discussion). Technometrics 1983;25(2):119-49.

15. Cook RD, Weisberg S (editors). Residuals and influence in regression. New York and London: Chapman and Hall, 1982, p. 183-205.

16. Ramsay T, Elkum N. A comparison of four different methods for outlier detection in bioequivalence studies. J Biopharm Stat 2005;15(1):43-52.