Use of Area under an ROC Curve as a Discriminating Measure for Identifying Data Separation of PK/PD Measurements in BE Studies

Sungwoo Choi[1], Stella Grosser[1] and Jessica Kim[1]

1 Division of Biometrics VIII, Office of Biostatistics, Center for Drug Evaluation and Research, FDA, Silver Spring, Maryland

Abstract

In bioequivalence (BE) studies, data separation observed in the bioanalytical order of pharmacokinetics (PK) or pharmacodynamic (PD) measurements could be one of the issues in data analysis. Recently, there were cases of BE study data showing potential PK measurement separation along with observed values' bioanalytical order and this type of issue becomes a challenge for researchers to provide reasonable explanations of the phenomenon. In this research project, the receiver operating characteristic (ROC) curve was applied to see how likely a separation between two groups could occur as a tool for analyzing a classification performance in binary classification (two groups) problems. Specifically, the area under the ROC curve (AUC_{ROC}) was used to evaluate a classification performance or a discrimination degree of two groups, where it took probabilities of an observation belonging to each group into consideration. In this problem, a classification rule (i.e., ratio of test to reference means ≥ 1) was defined to allocate an individual measurement to one of the two classes and the corresponding AUC_{ROC} was computed accordingly, where a larger AUC_{ROC} value (0.5≤ AUC_{ROC} ≤1) indicated a stronger separation. A simulation study was conducted based on a case example to evaluate how likely its AUC_{ROC} could happen compared to all possible AUC_{ROC} s of the considered case example. The AUC_{ROC} based analysis can help evaluate BE study data with any potential data separation of PK/PD measurements along with observed values' bioanalytical order.

Introduction

- In BE studies, a random selection of study subjects is not expected to result in data separation across the bioanalytical order of PK measurement.
  - Data separation observed in the bioanalytical order of PK measurement could raise concerns on the study data validity.
  - Trend in test to reference (T/R) ratios of PK measurement shows two distinct groups of subjects.
- This may lead to be potentially biased toward favorable BE outcome.

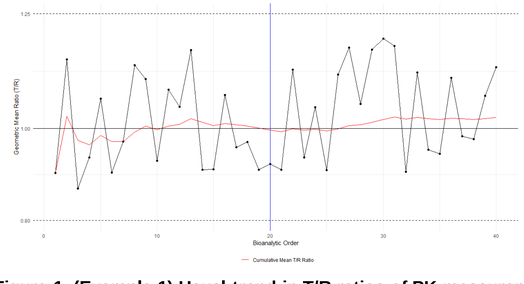

Figure 1. (Example 1) Usual trend in T/R ratios of PK measurement

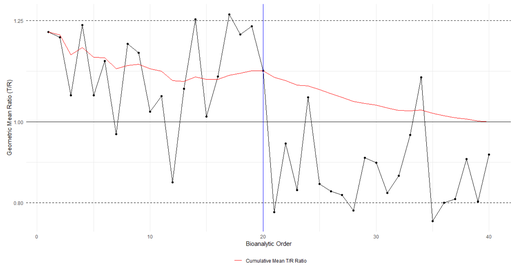

Figure 2. (Example 2) Data separation in the bioanalytical order of PK measurement

- The likelihood of separation between groups, as seen in Figure 2, can be assessed by investigating the classification performance of a binary classification.

Materials and Methods

Receiver Operating Characteristic (ROC) Curve

- Useful tool for analyzing a classification performance in binary classification
- The plot of true positive rate (TPR) on the horizontal axis and false positive rate (FPR) on the vertical axis

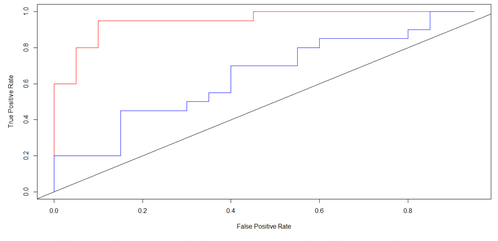

Figure 3. Two ROC curves and a diagonal; Blue (Example 1), Red (Example 2)

- Upper left corner → perfect classification; Diagonal → random allocation
- Classification performance can be evaluated using the area under the ROC curve (AUC_{ROC}).
  - A larger AUC_{ROC} values → a better classification performance (0.5 ≤ AUC_{ROC} ≤ 1)

The Area Under the ROC Curve

The empirical AUC_{ROC} (the Mann–Whitney U statistic) is

$$\widehat{AUC}_{ROC} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} I(X_i > Y_j)$$

where $X_1, \dots X_{n_1}$ and $Y_1, \dots, Y_{n_2}$ are random samples from two classes, respectively and $I(X_i \geq Y_j) = 1$ if $X_i > Y_j$ and $I(X_i < Y_j) = 0$ otherwise.

Example of Two Different Scenarios

Classifier	Group 1	Group 2
Example 1	−+−−+−−++−+++−−+−−−−	−+−+−+++++−+−−+−−++
Example 2	++++++−++++−++++++++	−−−+−−−−−−−−+−−−−−−

- The classification rule is defined as to allocate an individual T/R ratio to one of the two classes: + if a T/R ratio is greater than or equal to 1; – if a T/R ratio is less than 1.
- AUC_{ROC} of Example 1 is 0.6500 and AUC_{ROC} of Example 2 is 0.9525.
- The comparison of classification performance shows that Example 2 discriminates two groups better than Example 1.

Results and Discussion

Case Study 1

- The trend in test to reference (T/R) ratios of PK measurement seemed to show two distinct groups of subjects with most T/R ratios greater than 1 in the first half and the most T/R ratios less than 1 in the second half among 38 subjects.

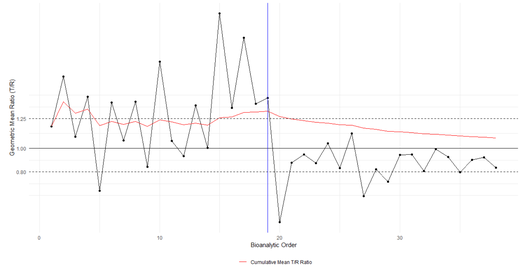

Figure 4. T/R ratios of PK measurement in Case study 1

- The classification rule is defined as to allocate an individual T/R ratio to one of the two classes: + if a T/R ratio is greater than or equal to 1; – if a T/R ratio is less than 1.
- Let the classifier for case study 1 be Classifier CS1. Then Table 1 shows Classifier CS1 of Case study 1 in its bioanalytical order.

	Subjects 1-19	Subjects 20-38
Classifier CS1	++++−+++−++−+++++++	−−−−+−+−−−−−−−−−−−−

Table 1. Classifier CS1

- AUC_{ROC} value of Classifier CS1 is 0.8920
- A simulation study is conducted to assess how likely this AUC_{ROC} value (=0.8920) would be observed out of the AUC_{ROC} values of the 1,000,000 randomly ordered sets obtained from Case Study 1 data.

Simulation

1. Let $(r_1, r_2, \dots, r_{38})$ be the set of T/R ratios of this study (Case Study 1) data
2. Generate 1,000,000 randomly ordered sets of $(r_1, r_2, \dots, r_{38})$
3. Calculate AUC_{ROC} s of 1,000,000 randomly ordered sets of $(r_1, r_2, \dots, r_{38})$

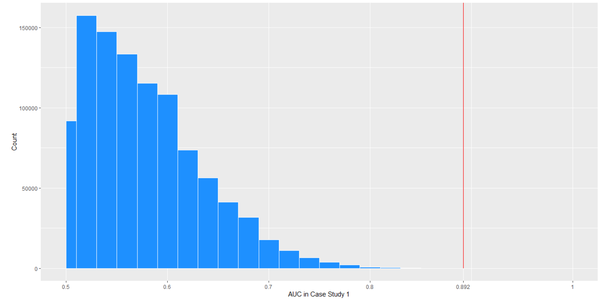

Figure 5. Histogram AUC_{ROC} values in Case study 1

Conclusion

- The comparison of classification performance of Case study 1 to simulated randomly ordered study data sets shows that the ability of any randomly ordered study data to discriminate two groups better than the Case study 1 will be extremely rare; $P(AUC_{ROC} \geq 0.8920) = 0.00068$.
- This research applies a statistical classification technique, ROC curve to look at how likely a separation between groups in bioequivalence study data could occur, which helps evaluate bioequivalence study data with any potential data trend.
- Specifically, the area under the ROC curve (AUC_{ROC}) was used to evaluate a classification performance or a discrimination degree of two groups, where it took probabilities of an observation belonging to each group into consideration.
- The probability of observing as large as or more extreme than the AUC_{ROC} of a study data estimated from a simulation was used as a quantified data separation information; the smaller the value is, it is unlikely that this separation between two groups is just an observed phenomenon or an inherent characteristic of randomization.
- For the future study, we will consider accuracy and misclassification measurements for an alternative method to evaluate the degree of data separation.

Disclaimer

The opinions and information in this poster are those of the authors, and do not represent the views and/or policies of the U.S. Food and Drug Administration.