Generic Drug Evaluation and R-package SABE

Elena Rantou, PhD
GASP, September 23, 2019
Washington, DC
Elena.Rantou@fda.hhs.gov
Disclaimer

- This presentation reflects the views of the presenter and should not be construed to represent the United States Food and Drug Administration’s views or policies.
- All data sets shown in this presentation have been previously de-identified.
Outline

- Office of Biostatistics/DBVIII
- Office of Generic Drugs/ORS/DQMM
- R-package ‘SABE’
- power simulations
- generate the distribution of certain statistics of interest
- assess the similarity of and cluster amino-acid sequences
- determine the validity of data sets categorized for genotoxicity
- characterize outliers in replicated, crossover design PK studies
- compare bioequivalence assessment approaches
- determine important features for identifying clinical sites for inspection
Similarity of amino-acid sequences

Use weighted sampling and select sequences using their frequencies as weights.

**Tanimoto Distance**

\[
T = \frac{N_{A \cap B}}{N_A + N_B - N_{A \cap B}}
\]
Self-Organizing Maps (package ‘SOM’)

Clustering using amino-acid frequencies

- Sample sequences using either random or weighted sampling
- For each sequence define mean similarity score across all other sequences
- For each sequence define the frequency of each amino-acid, i.e., ‘A’, ‘K’, ‘E’ and ‘Y’
Genotoxicity data integrity

- Examining data from the Ames test on different genotoxic impurities. Such data demonstrated suspicious patterns and unusual degree of replication.

- The objective was to analyze the reported positive control data in order to investigate the existence, pattern and likelihood of lack of variation and assess the probability of the occurrence of such outcomes.
Genotoxicity data integrity
Simulation study for likelihood assessment

(R-package ‘compoisson’)

\[ M = \frac{\text{total number of distinct observations}}{\text{total number of observations}} \]

<table>
<thead>
<tr>
<th>Underlying distribution model</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient of Variation CV</td>
</tr>
<tr>
<td>Poisson</td>
<td>0.0000</td>
</tr>
<tr>
<td>COM-Poisson</td>
<td>0.0000</td>
</tr>
<tr>
<td>Data</td>
<td>0.5385</td>
</tr>
<tr>
<td>Historical data 1</td>
<td>0.0000</td>
</tr>
<tr>
<td>Historical data 2</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
Simulation study for likelihood assessment (R-package ‘compoisson’)

The derived sampling distribution of the robust coefficient of variation, $CV_R$ when resampling from the distribution of the historical data 2, shows a marked value on the left tail which is the observed value of $CV_R$ from the reported data. This can be considered as an empirical p-value. If this was the true underlying distribution, the observed value would be extremely rare as it only occurs twice in 10,000 samples.
Outliers in replicated crossover PK studies

When formulations are compared with respect to their PK-characteristics, there may exist

- ‘unusual’ subjects or
- ‘unusual’ observations within a certain formulation

with extremely high or low bioavailability values
Outliers in replicated crossover PK studies

- The $D_t$ statistic (Wang and Chow, 2003) is based on the residuals from a linear model and seems to be a consistent metric for outlier characterization.

- $D_t$ is suitable for replicated crossover designs.
This is for an abbreviated new drug application for a generic topical cream. A traditional approach for establishing BE relies on a clinical endpoint study and uses success proportion (where success = at least 2-grade improvement based on 5-point scale of the condition severity) as a study endpoint.

An applicant proposed a new approach based on AUEC/Emax for establishing BE.

The three graphs above help us comparing the chances of passing 1) equivalence test, 2) superiority test and 3) both tests when using the two approaches, when the test and reference products are indeed equivalent based on simulation.
Objective is to determine if data mining techniques and/or unsupervised statistical monitoring can assist with the process of identifying potential clinical sites for inspection.

Summary of methods used to predict site inspection outcomes.

<table>
<thead>
<tr>
<th></th>
<th>SMART™</th>
<th>CISST</th>
<th>Data mining</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
<td>Detects outliers using distributional assumptions about the data.</td>
<td>Expert opinions used to develop a risk-based model.</td>
<td>Historical data used to train classification models for prediction.</td>
</tr>
<tr>
<td><strong>Predictions</strong></td>
<td>Uses p-value to identify atypical sites.</td>
<td>Assigns risk score to each site.</td>
<td>NAI, VAI, or OAI. NAI or VAI/OAI</td>
</tr>
</tbody>
</table>
Comparison of methods for clinical investigator site inspection selection

Machine learning (ML) methodology to predict Abbreviated New Drug Application (ANDA) submissions

Application of ML for Time-to-Event analysis

Equivalence Testing of Complex Particle Size Distribution Profiles
Predictive analysis of first ANDA submission for new chemical entities based on machine learning methodology

- Random Survival Forest (RSF) ML method is employed to forecast the time to first ANDA submission, referencing a new chemical entities (NCE) drug product
- RSF is superior in predictive performance comparing to conventional time-to-event methodology
- Variable importance of predictors (e.g., drug product, regulatory and pharmacoeconomic information variables) is assessed

Big data toolsets to pharmacometrics: Application of machine learning for time-to-event analysis

- Big Data tools (machine learning, ML) are applied to address pharmacometric problems
- The predictive performance of ML methods is superior compared to the Cox regression model under various simulated scenarios
- ML methods demonstrate less sensitivity to data sizes and censoring rates

Equivalence testing of complex particle size distribution profiles based on Earth Mover’s Distance

- EMD approach is employed to compare complex PSD profiles for equivalence assessment
- The developed approach is both effective and sensitive to pass equivalent products and reject inequivalent products in cases of multimodal PSD

Bioequivalence assessment for topical dermatological products and the In-Vitro Permeation Test (IVPT)

Package ‘SABE’*

*Scaled Average BioEquivalence
IVPT Study Design

Donor 1  Donor 2  Donor 3  Donor 4  Donor 5  Donor n

Test
Reference

Source: Bronaugh and Franz (1986)
IVPT Study Design

The response considered is the log-transformed
- total penetration ($AUC$)
- max flux rate ($J_{max}$)

We consider a sample of

$n$: donors (per treatment),

$r$: replicate skin sections from each one of the $n$ donors are collected for each formulation (replicates from each donor are randomly assigned to each product)

2 treatment formulations: test (generic: T) and reference (R)
Mixed CDER criterion uses the intra (within)-reference variability as a cutoff point.

For $S_{WR} \leq 0.294$, the test and reference formulations are declared bioequivalent if the $(1-2\alpha) * 100\%$ confidence interval:

$$
\bar{I} \pm t_{(n-1),\alpha} * \sqrt{\frac{S_I^2}{n}}
$$

is contained within the limits $\left[ \frac{1}{m}, m \right]$
The scaled BE methodology used in the case that $S_{WR} > 0.294$, adopts the FDA/CDER approach for the analysis of highly variable drugs, modified for the particular design

The hypotheses to be tested are:

$$H_0: \frac{(\mu_T - \mu_R)^2}{\sigma_{WR}^2} > \theta$$

$$H_a: \frac{(\mu_T - \mu_R)^2}{\sigma_{WR}^2} \leq \theta$$

Where $\theta = \frac{(\ln(m))^2}{(0.25)^2}$
Based on the this criterion, the two products are declared equivalent if

1. The point estimate (GMR) is contained within the limits $\left[\frac{1}{m}, m\right]$

2. The upper 95% bound of the scaled confidence interval is $\leq 0$
R-package ‘SABE’

- Tests for BE using the mixed scaled criterion
- Estimates statistical power as a function of the sample size
- Compares statistical power using the mixed scaled criterion (SABE) vs. that of using regular average BE (ABE)
- Estimates statistical power for different levels of the BE margin
- Estimates the size of the test (alpha-level)
R-package ‘SABE’

- Conducts sensitivity analysis with varying the number of replicates per donor, as well as, the inter-donor and within-reference variability levels
- Balances an unbalanced data set using different criteria
- Produces graphical displays that demonstrate the variability levels and potential extreme replicate values (outliers)
Bioequivalence assessment

IVPT.outcome(DataSet)

<table>
<thead>
<tr>
<th>pk_metric</th>
<th>T/R Ratio</th>
<th>Unscaled 90% CI LL</th>
<th>Unscaled 90% CI UL</th>
<th>Swr</th>
<th>Scaled Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>1.00860</td>
<td>0.6416316</td>
<td>1.755730</td>
<td>1.650961</td>
<td>-1.328058</td>
</tr>
<tr>
<td>Cmax</td>
<td>1.11192</td>
<td>0.7576997</td>
<td>1.611803</td>
<td>1.573147</td>
<td>-1.419273</td>
</tr>
</tbody>
</table>

R-package ‘SABE’
Power analysis

R-package ‘SABE’

Power with respect to PK-metric

Power with respect to BE assessment method
alphaTest(PE, matrixT, matrixR, n, r, trialn)

<table>
<thead>
<tr>
<th>SABE</th>
<th>ABE</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.03128</td>
<td>0.005038</td>
<td>4</td>
</tr>
<tr>
<td>0.03054</td>
<td>0.00245</td>
<td>6</td>
</tr>
<tr>
<td>0.02752</td>
<td>0.001334</td>
<td>8</td>
</tr>
<tr>
<td>0.02387</td>
<td>0.000756</td>
<td>10</td>
</tr>
<tr>
<td>0.02037</td>
<td>0.000432</td>
<td>12</td>
</tr>
<tr>
<td>0.01721</td>
<td>0.00024</td>
<td>14</td>
</tr>
<tr>
<td>0.01346</td>
<td>0.000128</td>
<td>16</td>
</tr>
<tr>
<td>0.01083</td>
<td>9.8e-05</td>
<td>18</td>
</tr>
</tbody>
</table>


[https://cran.r-project.org/web/packages/compoisson/compoisson.pdf](https://cran.r-project.org/web/packages/compoisson/compoisson.pdf)

Acknowledgements

Office of Biostatistics/OTS
Stella Grosser
Fairouz Makhlouf
Nam Hee Choi
Sungwoo Choi
Paul Schuette

Office of Generic Drugs
Meng Hu
Liang Zhao