U.S. FOOD & DRUG ADMINISTRATION

Generic Drug Evaluation and R-package SABE

Elena Rantou, PhD

GASP, September 23, 2019

Washington, DC

Elena.Rantou@fda.hhs.gov

Disclaimer

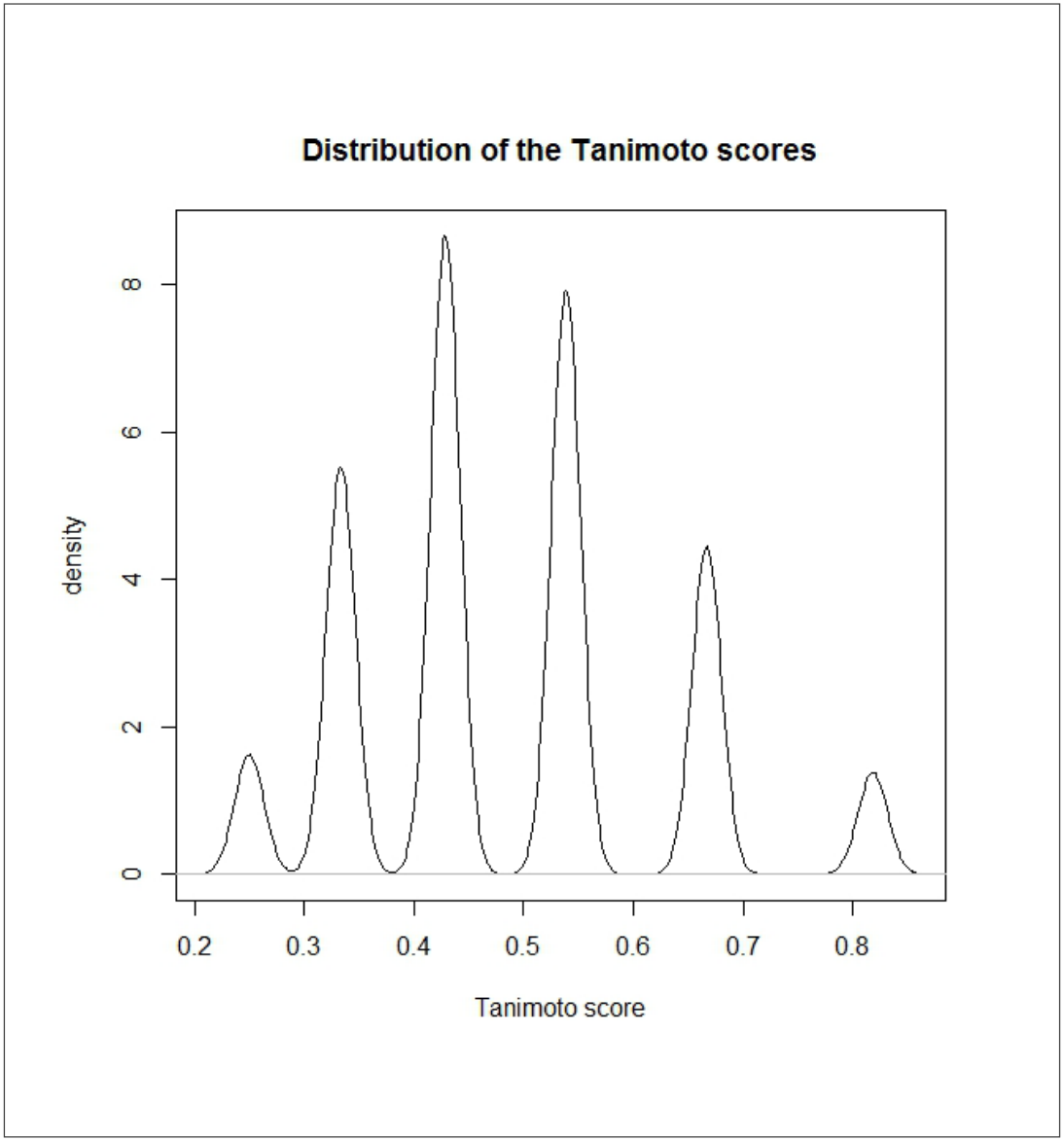
- This presentation reflects the views of the presenter and should not be construed to represent the United States Food and Drug Administration's views or policies
- All data sets shown in this presentation have been previously de-identified

Outline

- Office of Biostatistics/DBVIII
- Office of Generic Drugs/ORS/DQMM
- R-package ‘SABE’

Office of Biostatistics / DBVIII

- 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



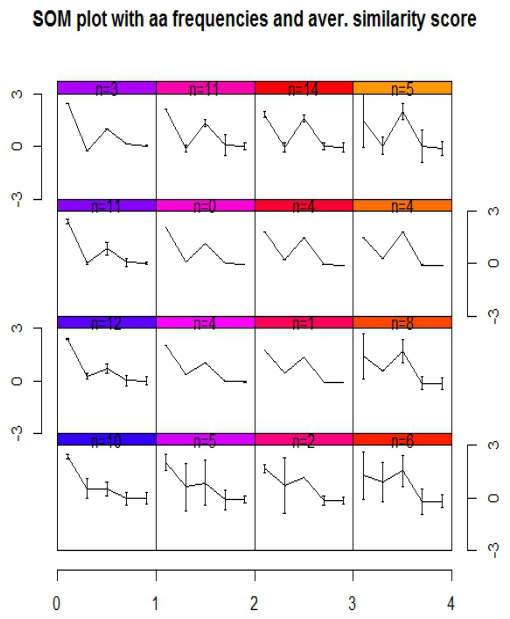
Similarity of amino-acid sequences

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

Tanimoto Distance

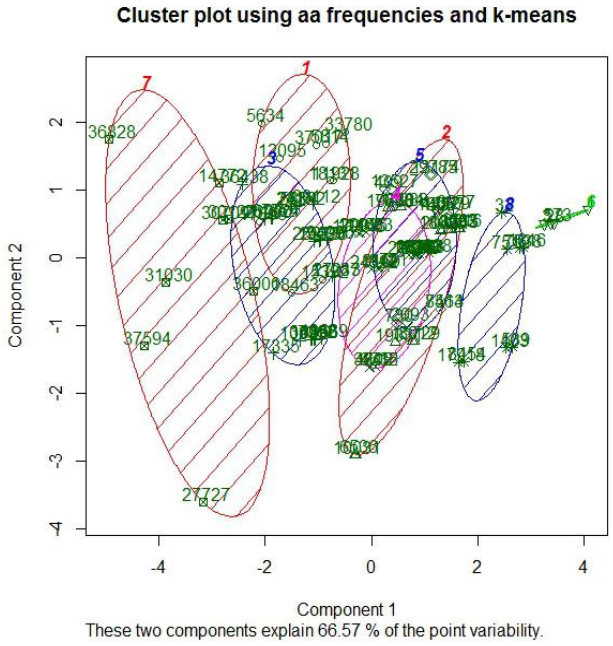
$$T = \frac{N_{A \cap B}}{N_A + N_B - N_{A \cap B}}$$

Self-Organizing Maps (package 'SOM')



Clustering using amino-acid frequencies

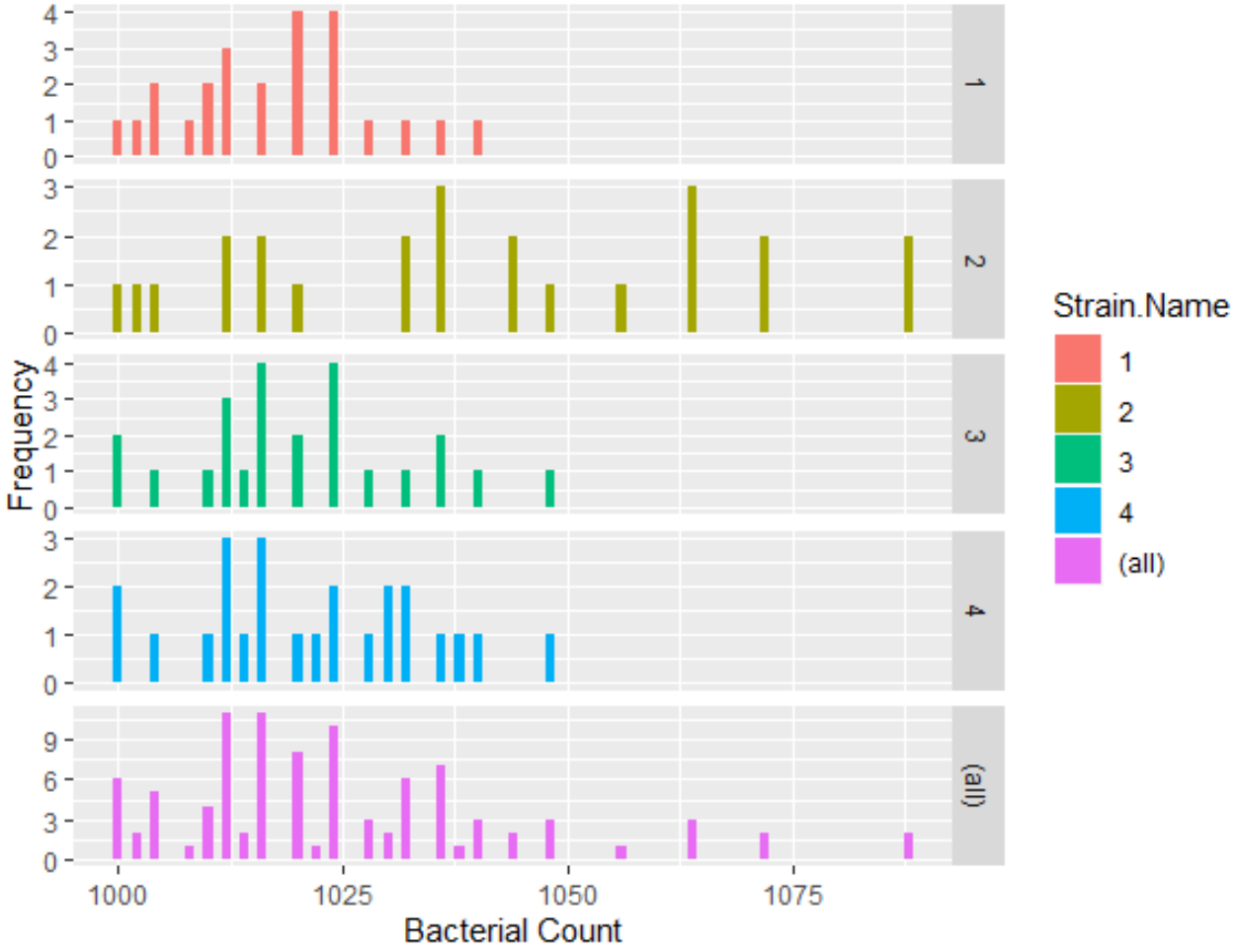
- Sample sequences using either random or weighted sampling
- For each sequence define mean similarity score across all other sequences
- For each sequence define the frequency of each amino-acid, i.e., 'A', 'K', 'E' and 'Y'



Genotoxicity data integrity

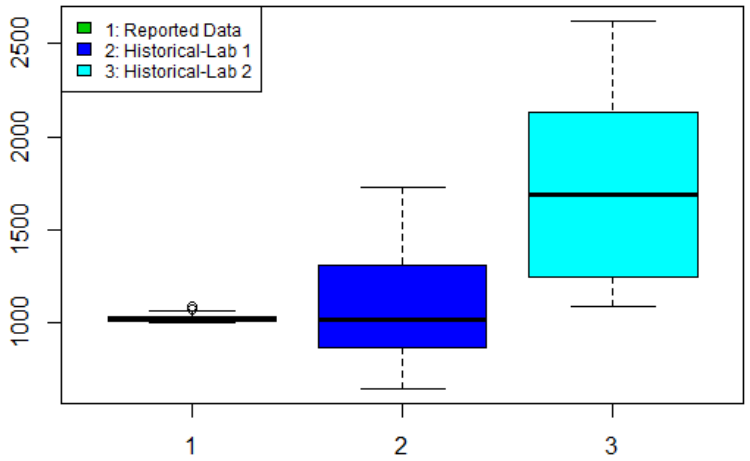
- Examining data from the Ames test on different genotoxic impurities. Such data demonstrated suspicious patterns and unusual degree of replication
- The objective was to analyze the reported positive control data in order to investigate the existence, pattern and likelihood of lack of variation and assess the probability of the occurrence of such outcomes

Genotoxicity data integrity



Simulation study for likelihood assessment

(R-package '***compoisson***')



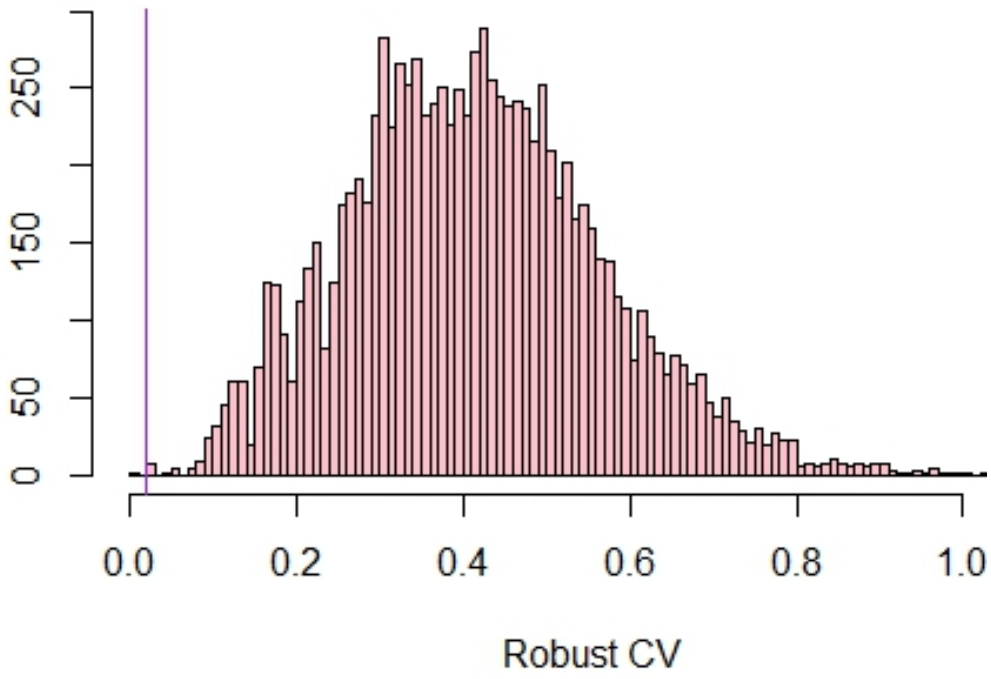
$$M = \frac{total\ number\ of\ distinct\ observations}{total\ number\ of\ observations}$$

Underlying distribution model	p-values		
	Coefficient of Variation CV	Robust Coefficient of Variation CV_R	M
Poisson	0.0000	0.0000	0.0000
COM-Poisson	0.0000	0.0000	0.0000
Data	0.5385	0.5531	0.0001
Historical data 1	0.0000	0.0001	0.0000
Historical data 2	0.0000	0.0002	0.0000

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

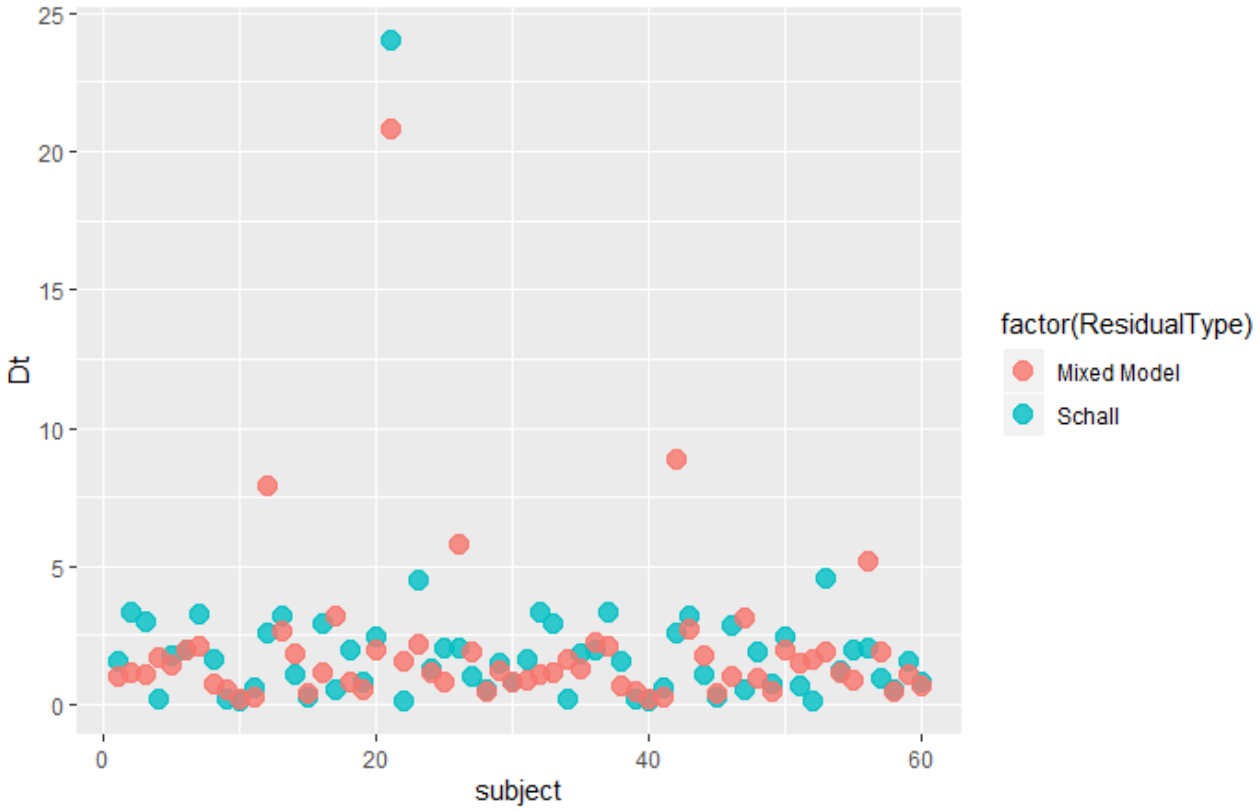
When formulations are compared with respect to their PK-characteristics, there may exist

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

with extremely high or low bioavailability values

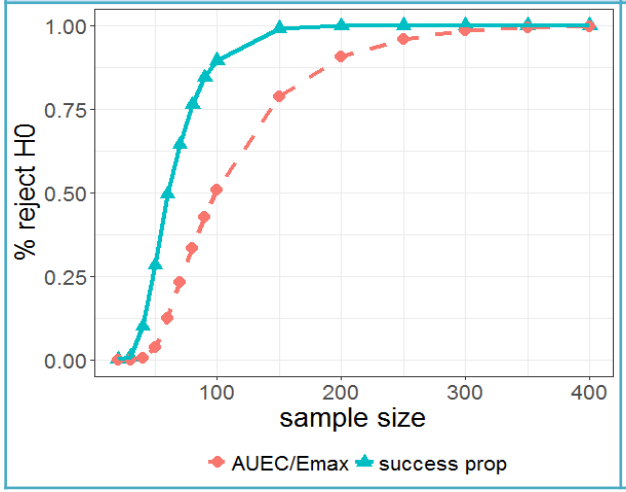
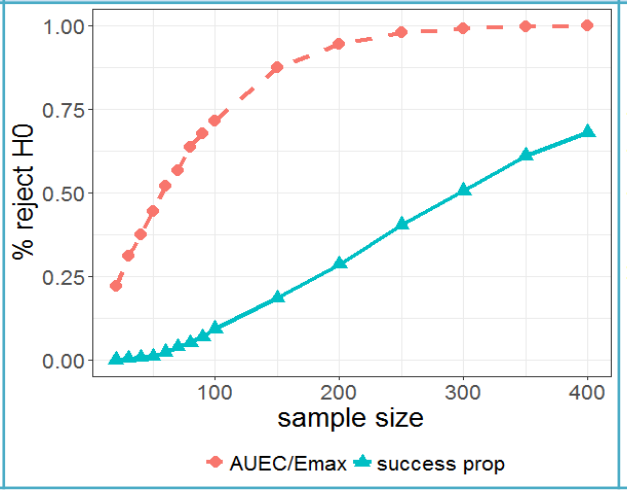
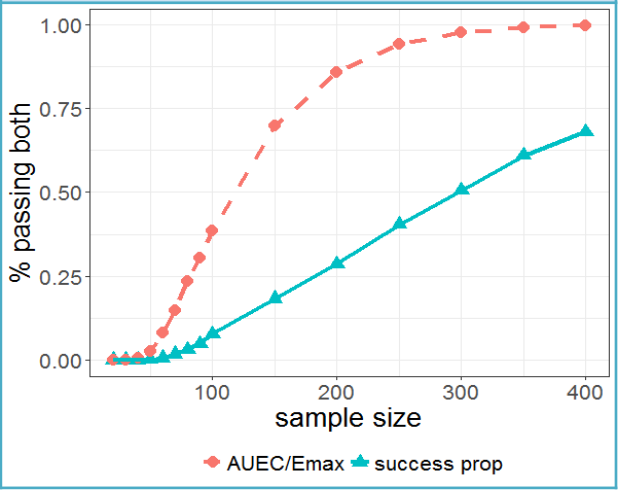
Outliers in replicated crossover PK studies

- 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



Comparison of two BE-assessment approaches

Passing Rates

Equivalence	Superiority	Both
		

- 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.

Comparison of methods for clinical investigator site inspection selection

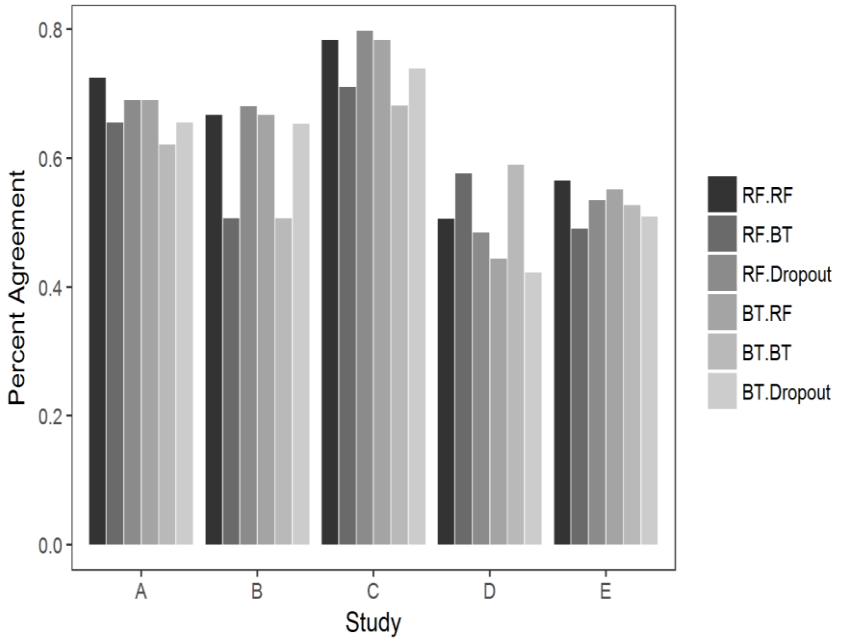
Objective is to determine if data mining techniques and / or unsupervised statistical monitoring can assist with the process of identifying potential clinical sites for inspection

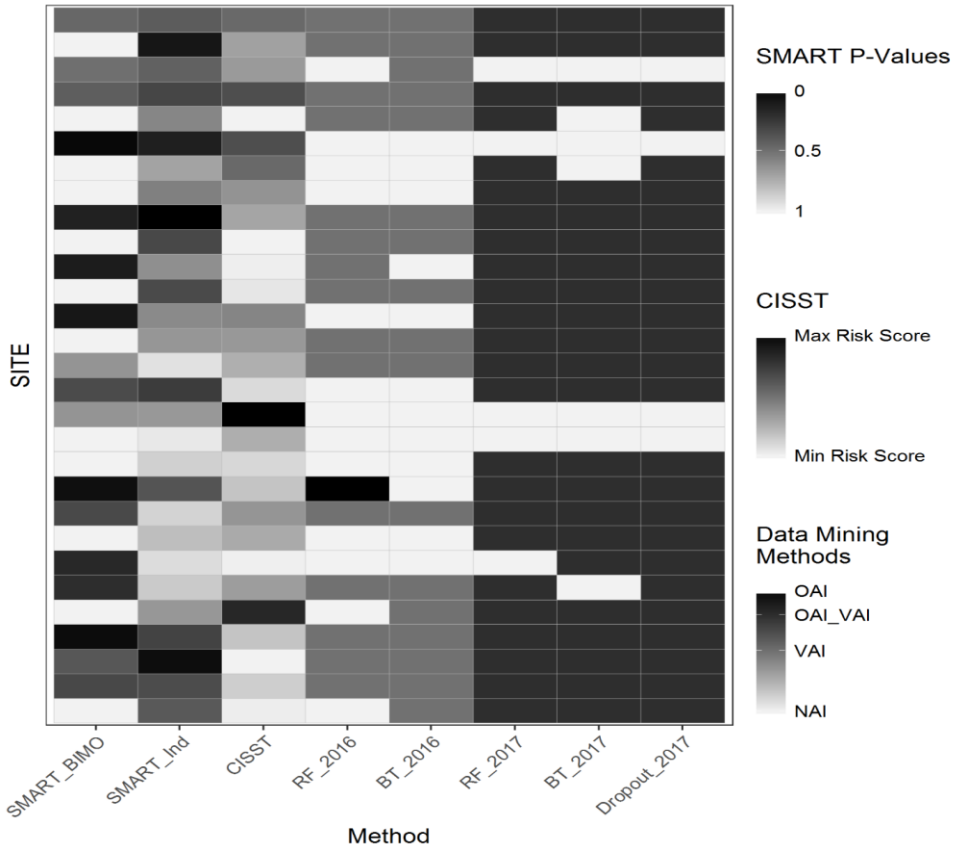
. Summary of methods used to predict site inspection outcomes.

			Data mining	
	SMART™	CISST	2016	2017
Description	Detects outliers using distributional assumptions about the data.	Expert opinions used to develop a risk-based model.	Historical data used to train classification models for prediction.	Historical data used to train classification models for prediction.
Predictions	Uses p-value to identify atypical sites.	Assigns risk score to each site.	NAI, VAI, or OAI.	NAI or VAI/OAI

Journal of biopharmaceutical statistics. 2019 Aug 30:1-4.

Comparison of methods for clinical investigator site inspection





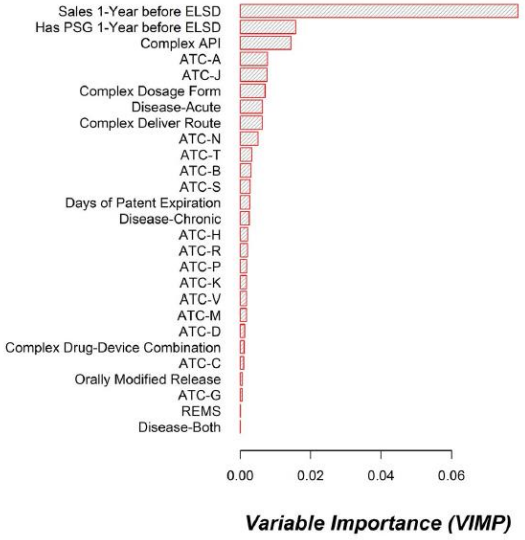
Journal of biopharmaceutical statistics. 2019 Aug 30:1-4.

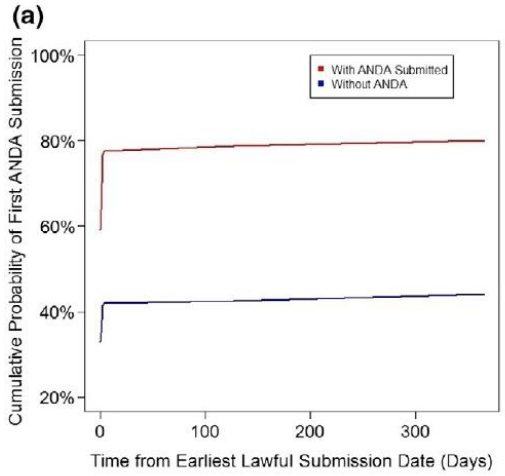
Office of Generic Drugs/Office of Research and Standards/Division of Quantitative Methods and Modeling

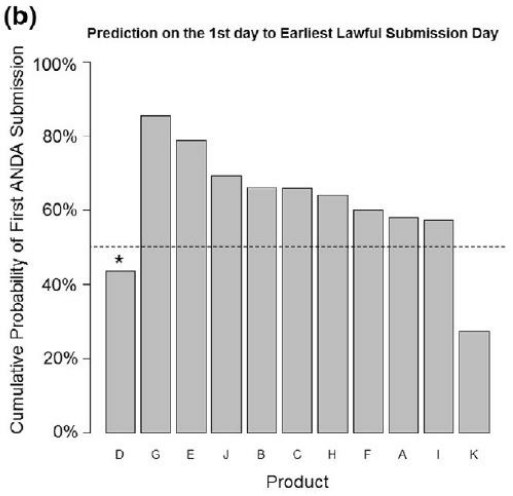
- Machine learning (ML) methodology to predict Abbreviated New Drug Application (ANDA) submissions
- Application of ML for Time-to-Event analysis
- Equivalence Testing of Complex Particle Size Distribution Profiles

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



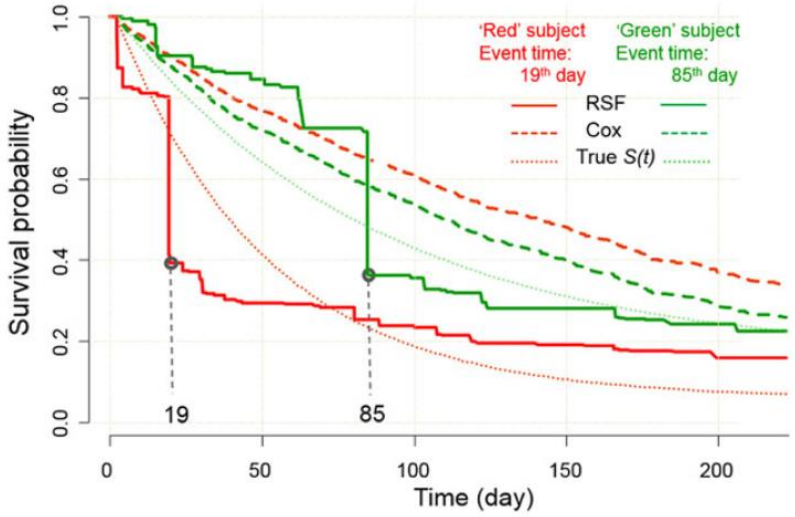


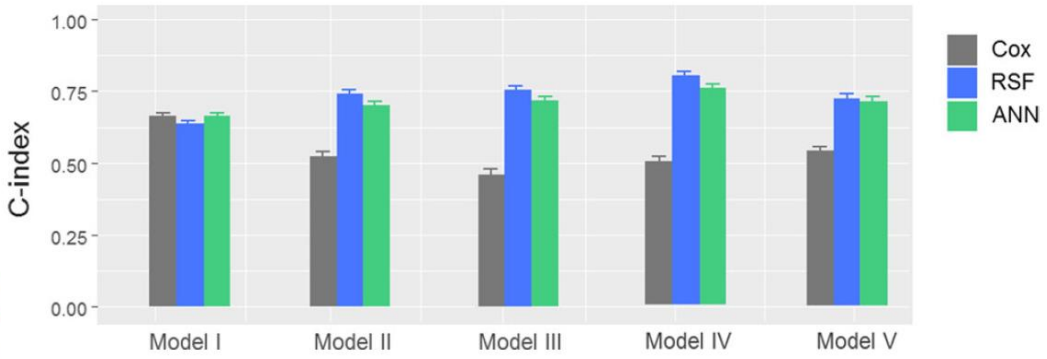


Clin Pharmacol Ther. 2019 Jul;106(1):174-181. doi: 10.1002/cpt.1479.

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

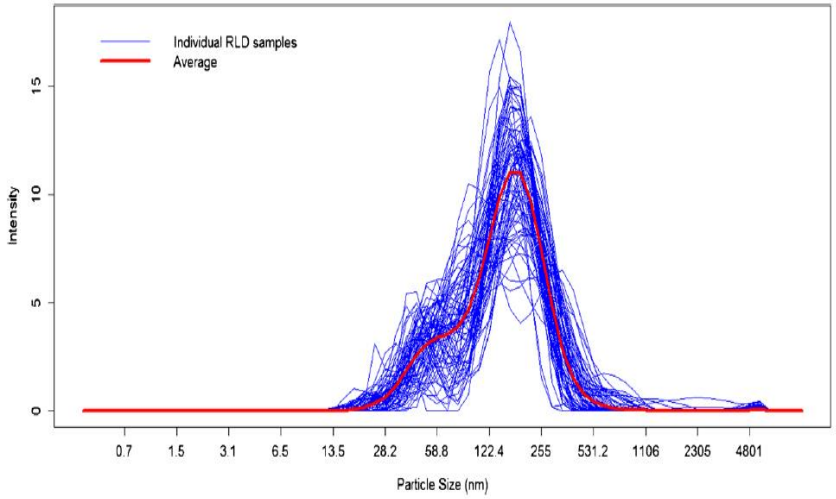


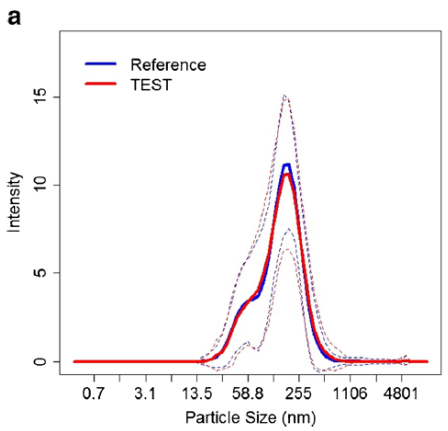


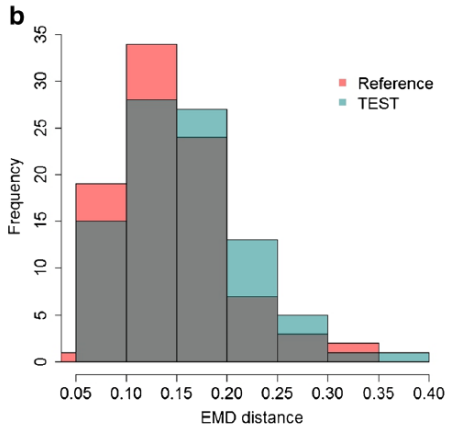
Clin Transl Sci. 2018 May;11(3):305-311. doi: 10.1111/cts.12541.

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







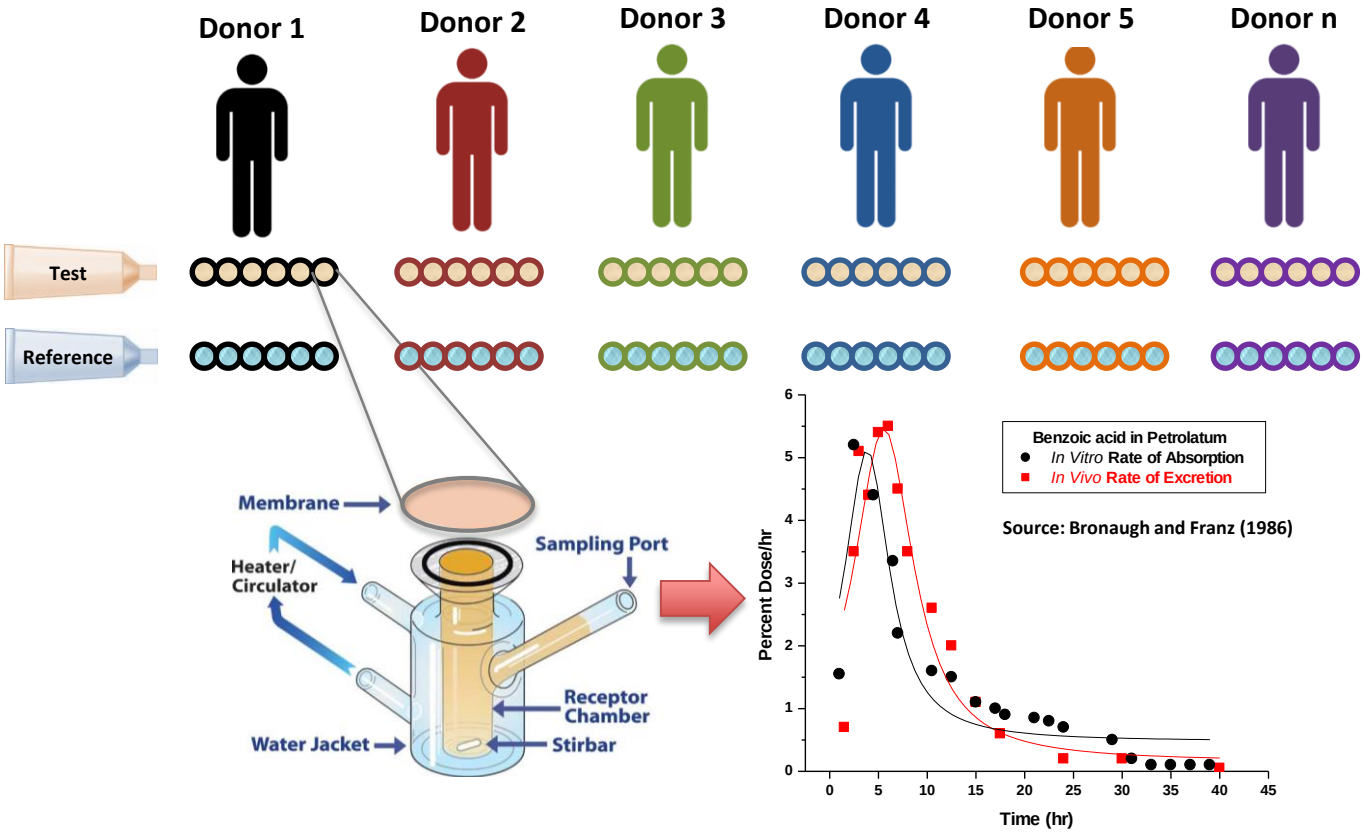
AAPS J. 2018 Apr 12;20(3):62. doi: 10.1208/s12248-018-0212-y.

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

*Package 'SABE'**

**Scaled Average BioEquivalence*

IVPT Study Design



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)

BE assessment

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 (1-2α) *100% confidence interval:

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

is contained within the limits $[\frac{1}{m}, m]$

BE assessment

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 $\theta = \frac{(\ln(m))^2}{(0.25)^2}$

BE assessment

Based on the this criterion, the two products are declared equivalent if

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

2. The upper 95% bound of the scaled confidence interval is ≤ 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)

R-package
'SABE'

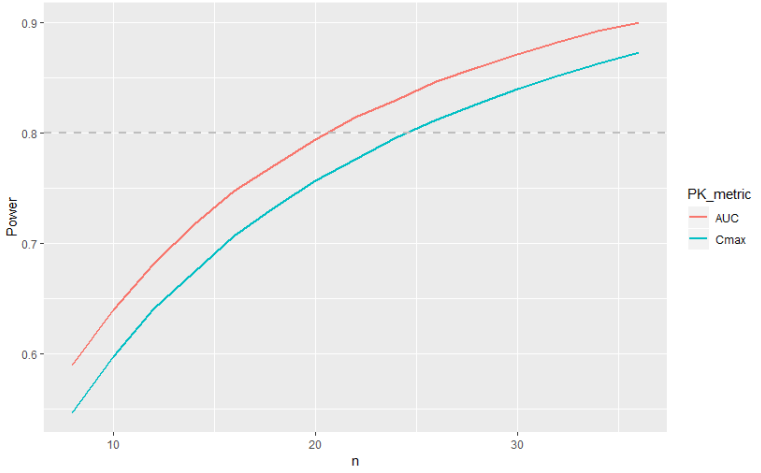
Bioequivalence assessment

IVPT.outcome(DataSet)

pk_metric	T/R Ratio	Unscaled 90% CI LL	Unscaled 90% CI UL	Swr	Scaled Upper Bound
AUC	1.00860	0.6416316	1.755730	1.650961	-1.328058
Cmax	1.11192	0.7576997	1.611803	1.573147	-1.419273

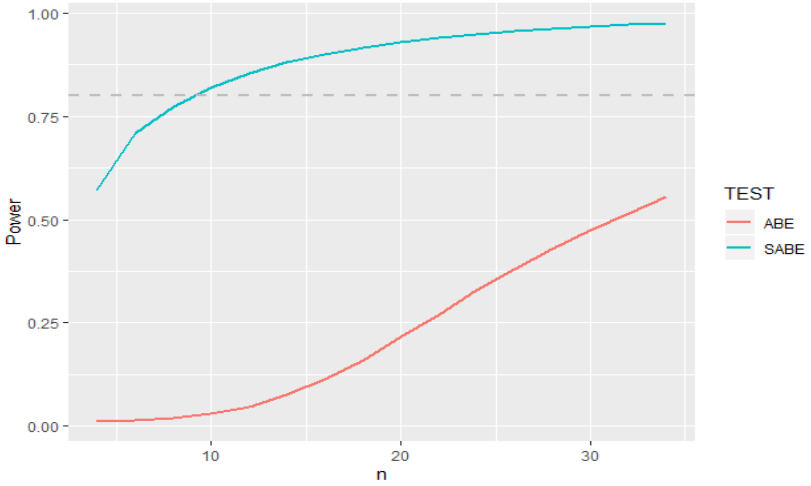
R-package ‘SABE’

Power analysis



Power with respect to PK-metric

Power with respect to BE assessment method



R-package 'SABE'

alphaTest(PE,matrixT,matrixR,n,r,trialn)

SABE	ABE	n
0.03128	0.005038	4
0.03054	0.00245	6
0.02752	0.001334	8
0.02387	0.000756	10
0.02037	0.000432	12
0.01721	0.00024	14
0.01346	0.000128	16
0.01083	9.8e-05	18

References

Hu, M., Jiang, X., Absar, M., Choi, S., Kozak, D., Shen, M., Weng, Y.T., Zhao, L. and Lionberger, R., 2018. Equivalence testing of complex particle size distribution profiles based on earth mover's distance. *The AAPS journal*, *20*(3), p.62.

Gong, X., Hu, M. and Zhao, L., 2018. Big Data Toolsets to Pharmacometrics: Application of Machine Learning for Time-to-Event Analysis. *Clinical and translational science*, *11*(3), pp.305-311.

FDA, Guidance for Acyclovir cream 5%, http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM428195.pdf. (recommended December 2014; revised in December 2016)

Hein, N., Rantou, E., & Schuette, P. (2019). Comparing methods for clinical investigator site inspection selection: a comparison of site selection methods of investigators in clinical trials. *Journal of biopharmaceutical statistics*, 1-14.

Hu, M., Babiskin, A., Wittayanukorn, S., Schick, A., Rosenberg, M., Gong, X., Kim, M.J., Zhang, L., Lionberger, R. and Zhao, L., 2019. Predictive analysis of first Abbreviated New Drug Application submission for new chemical entities based on machine learning methodology. *Clinical Pharmacology & Therapeutics*.

Stead, A. G., Hasselblad, V., Creason, J. P., & Claxton, L. (1981). Modeling the Ames test. *Mutation Research/Environmental Mutagenesis and Related Subjects*, *85*(1), 13-27.

Sellers, K. F., Borle, S., & Shmueli, G. (2012). The COM-Poisson model for count data: a survey of methods and applications. *Applied Stochastic Models in Business and Industry*, *28*(2), 104-116.

https://cran.r-project.org/web/packages/compoisson/compoisson.pdf

Wang, W., & Chow, S. C. (2003). Examining outlying subjects and outlying records in bioequivalence trials. *Journal of biopharmaceutical statistics*, *13*(1), 43-56.

Acknowledgements

Office of Biostatistics/OTS

Stella Grosser
Fairouz Makhlouf
Nam Hee Choi
Sungwoo Choi
Paul Schuette

Office of Generic Drugs

Meng Hu
Liang Zhao


FDA
U.S. FOOD & DRUG
ADMINISTRATION