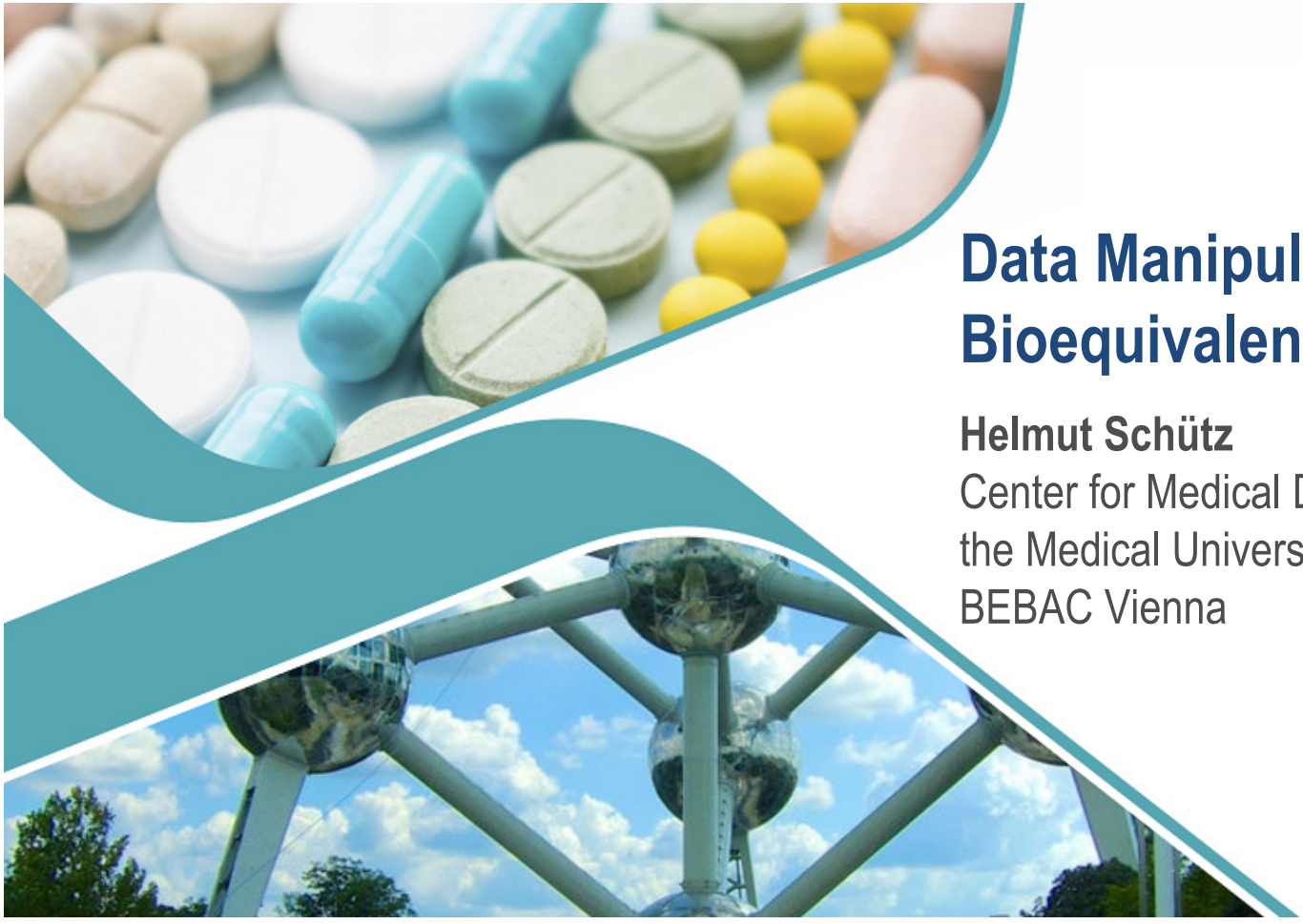
MEDICAL UNIVERSITY OF VIENNA

BEBAC



Data Manipulation in Bioequivalence

Helmut Schütz
Center for Medical Data Science of the Medical University of Vienna
BEBAC Vienna

Background

Data manipulation detected in the past

- Ranbaxy (2004 – 2008), GVK Bio (2014), Semler (2016), Panexcell (2019), Synchron Research (2022), Synapse (2023)

Various ‘methods’ used by the CROs

- Only the reference administered
- Fake sequences, e.g. TT | RR
- Unblinded interim analysis and – if BE unlikely due to T/R-ratio
  - Swap the code of T and R in subsequent subjects
  - If T/R-ratio in the interim is very ‘bad’, additionally dilute T- or R-samples
- Analyze backup samples of yet another study

Risk

Regulatory agencies use an arsenal of tools to detect fraud

- In the ideal situation a whistleblower gives details, which helps inspectors (Ranbaxy and GVK cases)
- Software
  - T/R-ratios of C_{max} vs analytical batch (Excel)
  - FDA's 'DABERS' (Data Anomalies in BioEquivalence R Shiny)
  - Fuglsang's 'Buster' and 'SaToWIB' routines (R)
  - BEBAC's 'FraudDetection' (R)

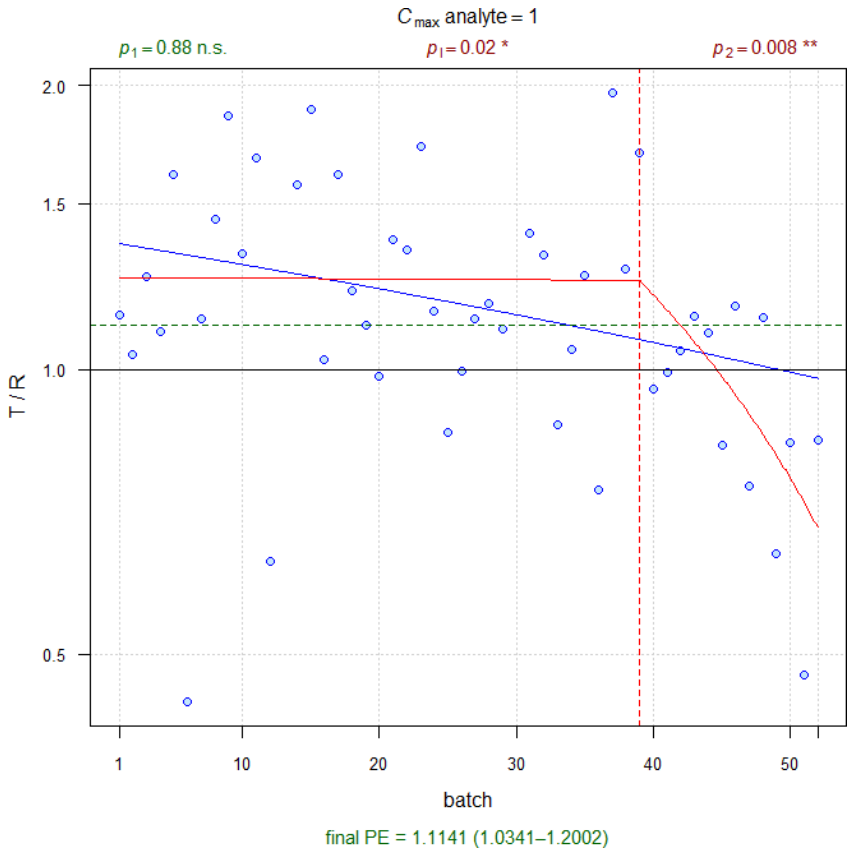
Do not wait for a regulatory action

- Request full data of the CRO before submission
- Assess the data by various approaches to detect a signal of potential manipulation
- Consider a thorough audit

2×2×2, n=50, suspected interim analysis

Simple approach: Assessing the T/R-ratios

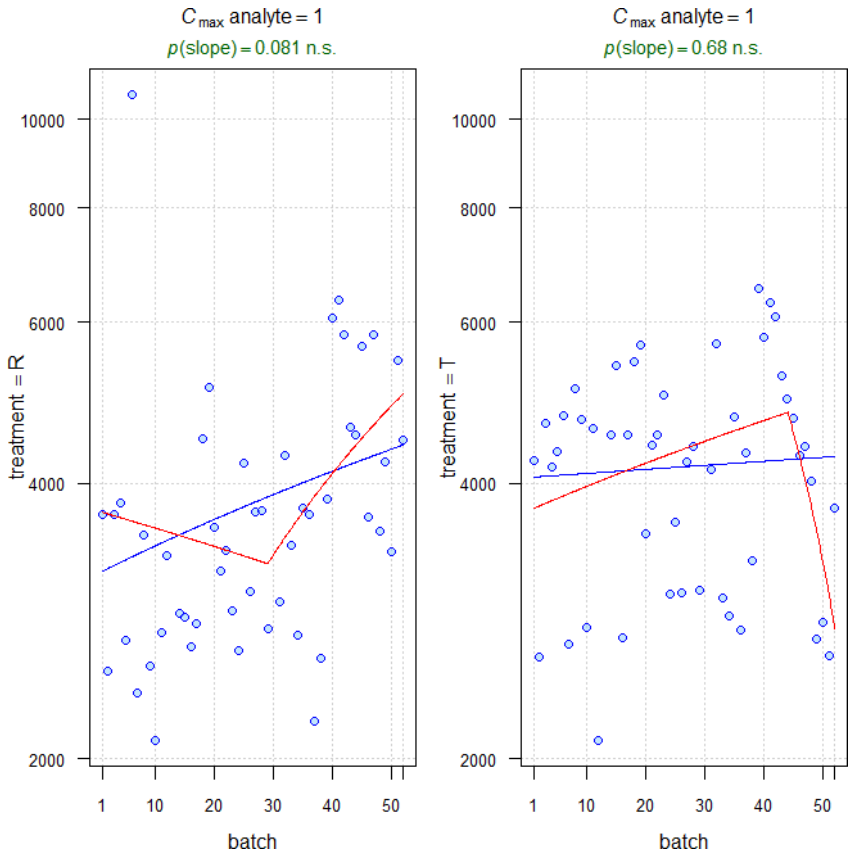
- Blue line linear regression
- Red line segmented (‘stick’) regression
- Starting with batch 39 (red dashed line) the T/R-ratios are significantly lower than before and thus ‘save’ the otherwise failing study
- BE in the final analysis



2×2×2, n=50, suspected interim analysis

C_{max}-values by treatment: Any differences, trend?

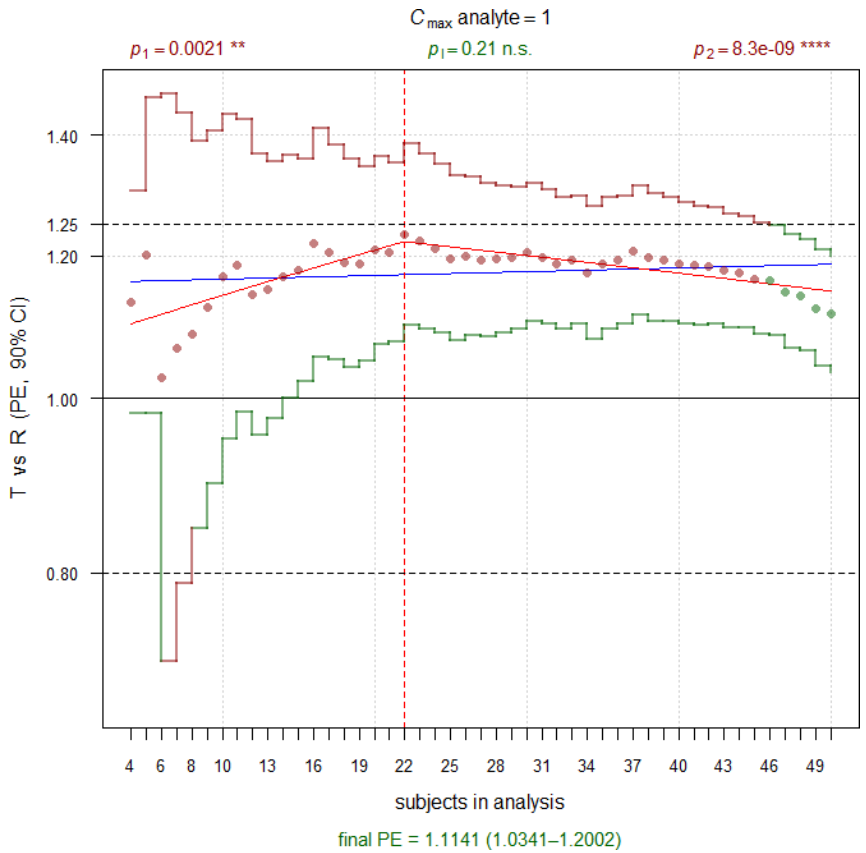
- A similar pattern like before
- Was T swapped with R in the later batches?
- Were the T-samples even diluted to 'improve' the T/R-ratio?



2×2×2, n=50, suspected interim analysis

BE assessed with an increasing number of subjects analyzed

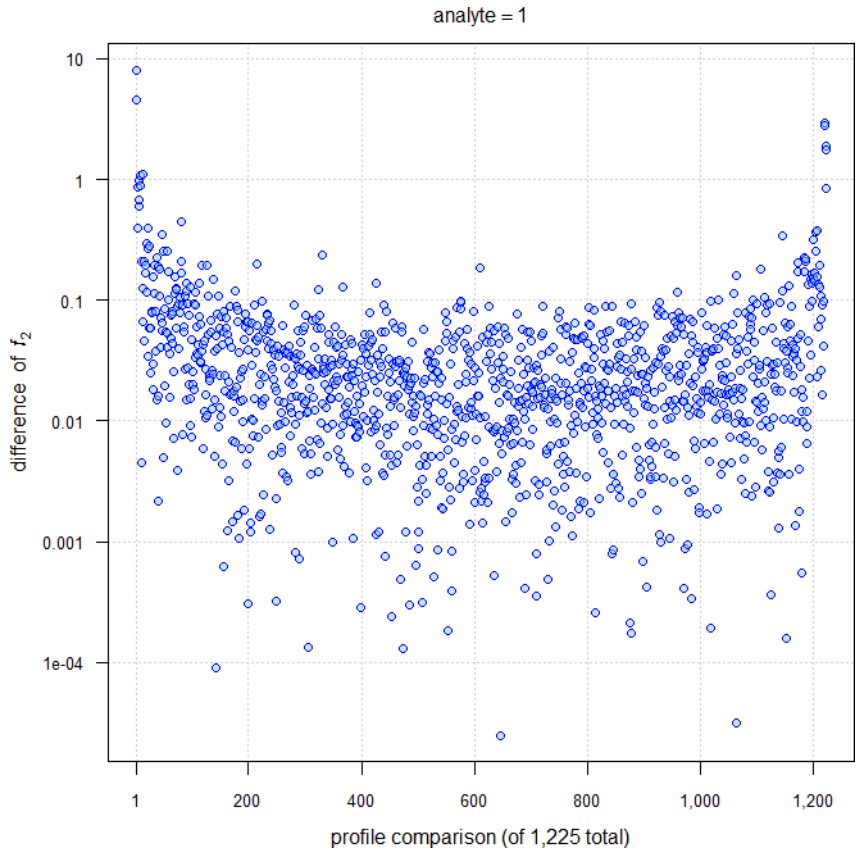
- Circles point estimates
- Stairs 90% confidence intervals
  (red if outside BE margin, green if passing BE)
- Both segments of PE re-gressions are significant
- Did the manipulation start already earlier than we assumed?



2×2×2, n=50, suspected interim analysis

Comparison of similarity of plasma profiles by f_2

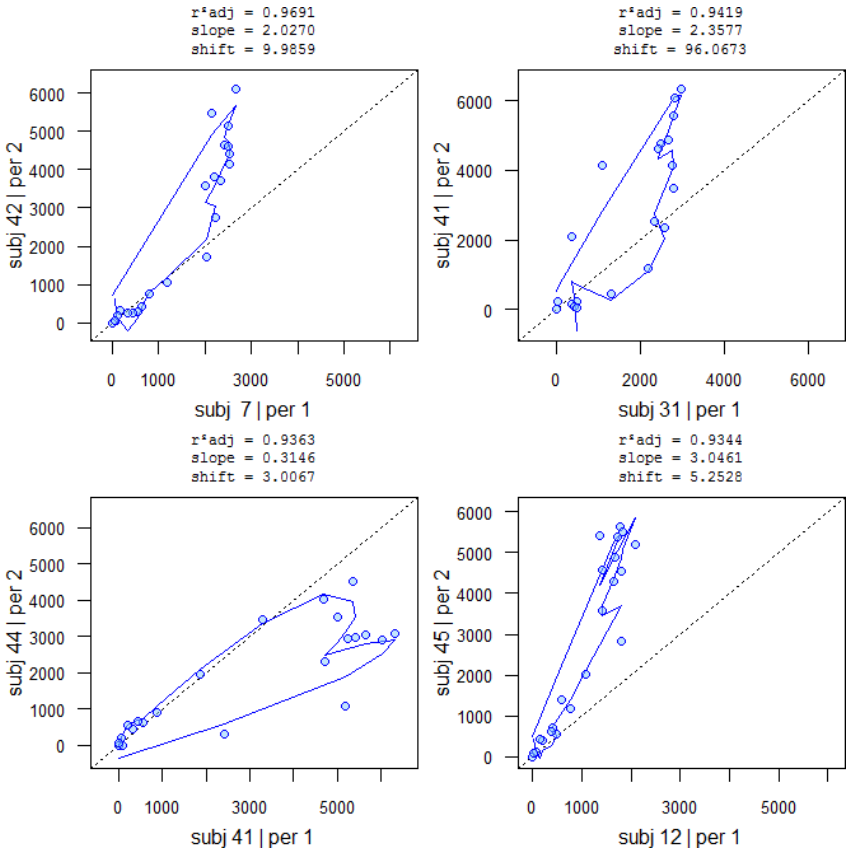
- Each profile with any other (irrespective of the treatment)
- Profiles with very small differences in their f_2-values are suspect



2×2×2, n=50, suspected interim analysis

Correlation of plasma profiles by the measured concentration

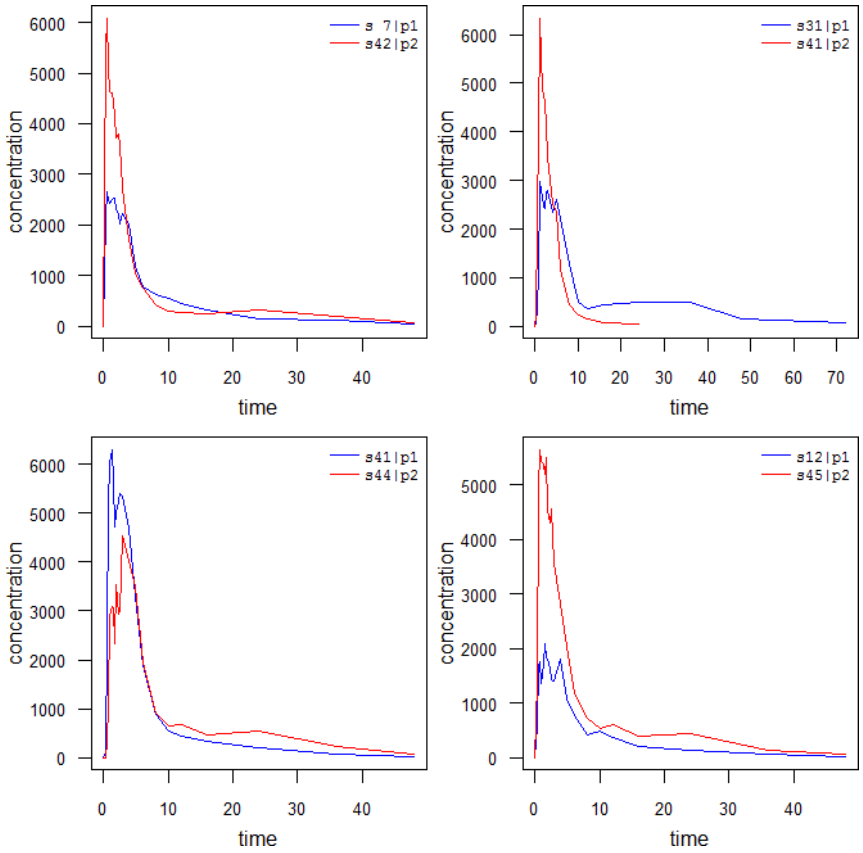
- Each profile with any other (irrespective of the treatment)
- We have to take the time into account – otherwise similar profiles with different lag-times will be falsely appear highly correlated
- Highly correlated concentrations are suspect



2×2×2, n=50, suspected interim analysis

Comparison of plasma profiles by the measured concentration

- We suggest to compare suspect highly correlated concentrations visually



R ‘FraudDetection’

Required data

- Sampling schedule
- Concentrations (any number of analytes)
- Analytical batches and / or dates of analysis
- Randomization (currently 2×2×2 and any Williams’ design)

Optional

- Actual sampling time points
- Method used by the CRO to calculate *AUC*
- PK metrics reported by the CRO

Supported data formats

- CSV, XLS(x), ODS, SAS XPT, Phoenix Project file
- CDISC (via Phoenix 8.3.4)

R ‘FraudDetection’

Recalculation by NCA

- C_{max}, t_{max}, C_{last} , t_{last}
- AUC_{0-t} (linear trapezoial or linear-up / logarithmic down)
- Optional
  - λ_z (start- and end-time, number of data points)
  - $AUC_{0-\infty}$ (observed or predicted)
  - Extrapolated fraction

Methods

- Spaghetti (grouped by treatment) and treatment (grouped by subject) plots
- PK metric by treatment vs batch or date of analysis
- T/R-ratios vs batch or date of analysis
- $\log_e$(PK) – mean[$\log_e$(PK)]; runs test

R ‘FraudDetection’

Methods cont’d

- $\log_e(PK_T/PK_R)$ – mean[$\log_e(PK_T/PK_R)$]; runs test
- BE by subjects analyzed (≥4)
  - Plot (PE, 90% CI)
  - Table (MSE, PE, 90% CI, pass|fail)
- MSE of model by subjects analyzed
- Model residuals by subjects analyzed
- Difference factor f_1 by subject
- Similary factor f_2 by subject
- Comparison of f_2 of profiles with any other
  - Plot of differences
  - Table of most and least similar profiles (default 6)
- If data provided by the CRO, comparison of NCA

Problems

Caveat

- Multiple analytes with the same method
  - Might give contradictory outcomes
  - Judgement required

Unresolved

- No statistical method
  (null hypothesis = no manipulation, alternative = manipulation)
  - Only exploratory and subjective
  - Leaves room for interpretion
- Breakpoint of segmented regression
  - Not unique in the different methods
  - Most reliable possibly the BE plot

Problems

Unresolved cont'd

- If study is performed in groups
  - Different PK might be detected by pure chance
  - Should not interpreted as a signal of manipulation
- Comparison of f_2
  - How similar is similar?
- Correlation of plasma profiles
  - Threshold of r^2 (default >0.95)?
  - Threshold of slope for detecting dilutions (default <0.5 and >2)?
- Runs test has low power

Problems

FDA about 'DABERS'

- **Despite its demonstrated effectiveness,** a major drawback is that the pharmacokinetics and pharmacodynamics may be too complicated to describe with a single statistic**. Indeed,** the current practice offers no practical guidelines regarding how similar PK profiles from different subjects can be in order to be considered valid**. This makes it difficult to assess the adequacy of data to be accepted for an ANDA and requires additional information requests to applicants. This project will address the current gap in identifying the data anomalies and potential data manipulations by use of state-of-the-art statistical methods, specifically focusing on machine learning and data augmentation. […] from a regulatory perspective, our project will provide a data driven method that can model complex patterns of PK data to identify potential data manipulations under an ANDA.**

https://www.hhs.gov/sites/default/files/hhs-ai-use-cases-2023-public-inventory.csv

Data Manipulation in Bioequivalence

Thank You!

Helmut Schütz

Center for Medical Data Science

MEDICAL UNIVERSITY OF VIENNA

1090 Vienna, Austria

helmut.schuetz@meduniwien.ac.at

BEBAC

1070 Vienna, Austria

helmut.schuetz@bebac.at

Acknowledgments: Osama Abd Elrahman, Susana Almeida, Anders Fuglsang, Olivier LeBlaye