Detection of data manipulation in bioequivalence trials

- How is data manipulated?
- How is fraud detected?
- How does the software tool work?
- The discovery process
- Software tool results

Abbreviations:

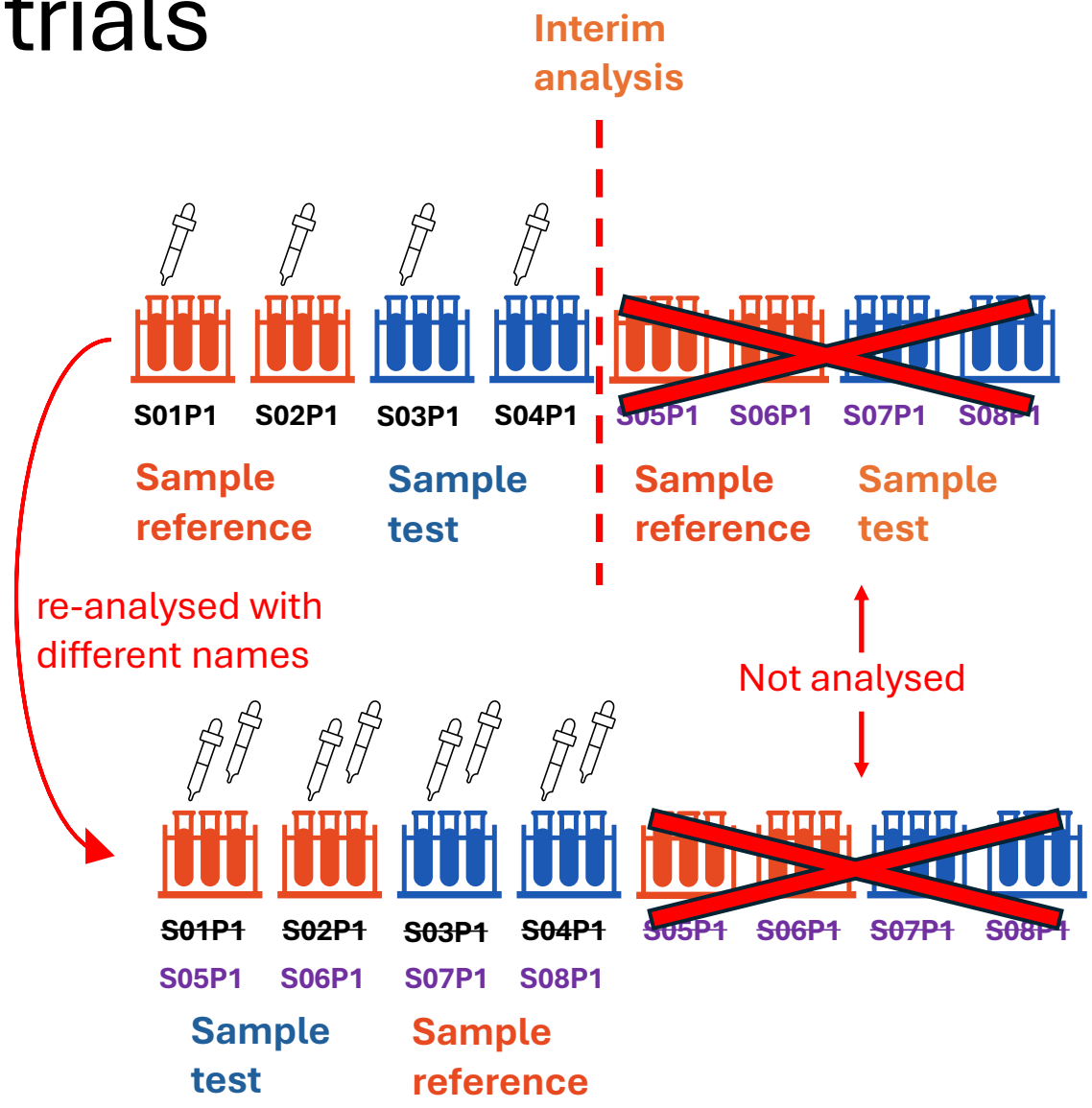
BE	bioequivalence *(Bioequivalence studies are a type of Phase I study in the development of generic medicines)*
PK	pharmacokinetic

Data manipulation in BE trials

How is data manipulated?

Interim-analyses after a portion of subject PK-data was analyzed. If ratio of test/reference plasma concentration is out of range fraud starts:

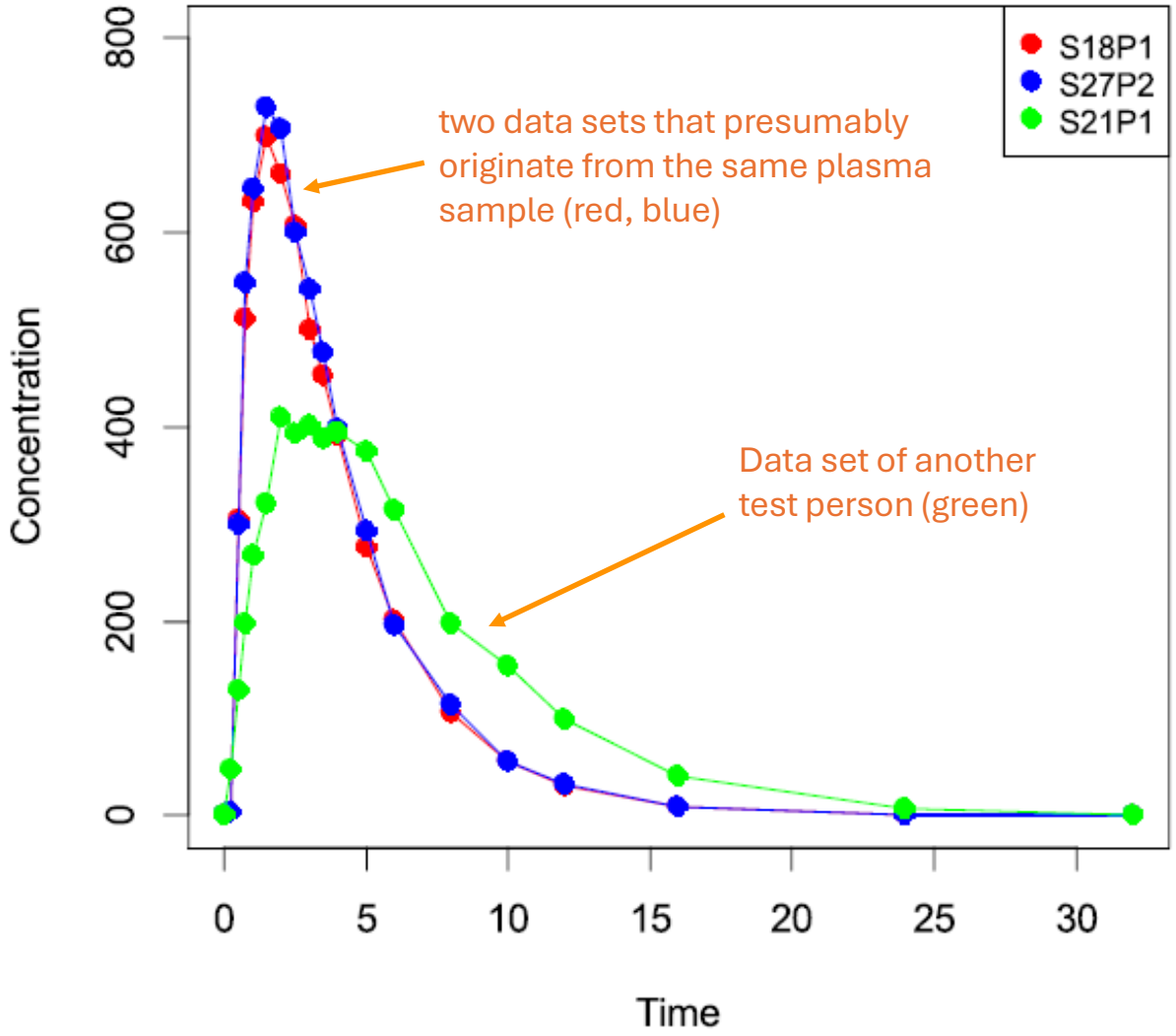
- Initially analyzed samples will be re-analyzed under new subject identity
- Typically, with Test- and Reference-Product switched (to have similar plasma concentrations)



Data manipulation in BE trials

How is fraud detected?

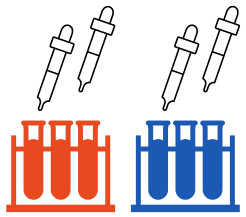
- Since 2020 software program with 2 different routines (Buster* and SaToWIB*) are used by health authorities
- Confirmation of signs alone (e.g. similarity of PK profiles) without solid evidence justifies questioning the validity of data
- Overlapping plasma time concentration profile*



*Fuglsang, A., 2021. Detection of data manipulation in bioequivalence trials. European Journal of Pharmaceutical Sciences. 156 (2021) 105595

Software tool

How does the software tool work?



- If the same plasma sample is injected into the analyzer twice (e.g., first time with the originator name and second time with the generic name), the results of these identical samples will be very similar.

- Software compares all plasma measurements for each subject and period, calculating a similarity-score* that extracts the most similar plasma sample results.

Score
0,14705
0,213025
0,218007
0,21997
0,23856
0,240832
0,255916
0,262737
0,265732

- The vials of samples that might have been re-injected should be checked at a site audit. If vials have been re-injected the amount of plasma inside the vials will be less compared to vials that have not been re-injected.

*Fuglsang, A., 2021. Detection of data manipulation in bioequivalence trials. European Journal of Pharmaceutical Sciences. 156 (2021) 105595

Software tool

The discovery process work

- Process starts with raw-data from site
- Data format should be tabular (e.g. .xlsx, .csv)
- Tidy data to consistent structure (Chapter 2 of the code)
- Run data through software
- Software creates list (Excel-sheet) with **most similar sample pairs** and shows **four most similar pairs in plots** (PDF)
- Auditor checks most similar vials stored at site. Re-analysed vials contain less plasma or plasma that has been diluted.

Software tool

Software tool result: list with most similar sample pairs

Subject

Period

Treatment

	A	B	C
1	Profile 1	Profile 2	Score
2	9 | 1 | B	9 | 2 | A	0,14705
3	1 | 1 | A	2 | 2 | A	0,213025
4	15 | 1 | A	16 | 1 | B	0,218007
5	15 | 1 | A	18 | 1 | A	0,21997
6	7 | 2 | A	20 | 1 | B	0,23856

the smaller the score, the more similar the plasma samples are

(a score of 0 would occur if all data points are identically, which is not possible due to measurement variability)

Software tool

Software tool result: most similar pairs in plots