

Implementation of bootstrapping in a replicative bioequivalence study with a highly variable drug

Feryal Tanriverdio, Boehringer Ingelheim Pharma GmbH & Co KG, Biberach, Germany

Arne Ring, Boehringer Ingelheim Pharma GmbH & Co KG, Biberach, Germany

ABSTRACT

For comparing the pharmacokinetic parameters of two formulations of the same substance, bioequivalence studies are used. For these studies, the FDA recommends to analyse them using the Average Bioequivalence-method.

For highly variable drugs, it is difficult to apply this method, as a large population size is needed to show that the confidence interval of the ratio of the average lies completely within the range of 80%-125%. In one of our trials, we used the Scaled average bioequivalence-method.

This method is technically more demanding, as the analysed statistics cannot anymore be assumed to be normally distributed. Therefore we applied a bootstrap method.

The aim of this paper is to show how the programming of this study was implemented, focusing specifically on the bootstrap methodology. Problems faced while programming, concerning the processing time and data structure will be shown and the solutions that were found for these problems.

INTRODUCTION

During the clinical development of a drug substance, different formulations get explored (composition of the tablets, dragees, etc). In this process the question concerning the safety and efficacy of the new formulation is posed, which will be acquired in so-called bioavailability – or respectively – in bioequivalence studies.

If an agent has a high pharmacokinetic variability, then a large sample size is required, in order to fulfil the traditional criterion of average bioequivalence.

A possible way out is the „scaled average bioequivalence“ – the bioequivalence limits get expanded as the variability of the reference formulation increases. This method requires a bigger effort in the implementation, as the analysed statistics are not normally distributed. Special linearizations or bootstrap methodology are possible solutions for this.

In the presented study it was the first time that at Boehringer Ingelheim the SABE-method was used. The evaluation had to be presented in a validated way, shortly after obtaining the pharmacokinetic endpoints (C_{max} and AUC). To ensure that the results were delivered within the time frame, the data of a completed bioequivalence study with similar design was copied and adapted to the actual data structure. On the basis of this data, the programming of the evaluation started before the actual data was received.

In the following, the problems in regards of programming – especially concerning the bootstrap methodology – and possible solutions will be illustrated.

BIOEQUIVALENCE ASSESSMENTS

Bioequivalence studies are an important appliance of testing oral administered formulations. If a new formulation of a pharmaceutical product is developed, it must be proved that it is equivalent to the already tested (e.g. during a clinical development) or approved (after the termination of a patent) product. For this purpose, pharmacokinetic parameters resulting from a concentration-time-curve of the actual (reference) formulation and the new (test) formulation are compared. The results of the concentration-time-curve after application of the test and reference formulation don't need to be identical, but similar inside of defined limits.

The assumption thereby is that if the bioequivalence of the pharmacokinetic profiles is similar between reference and test product, then also the safety and efficacy of the two formulations is similar. In different countries, the regulations for marketing authorisation of pharmaceutical drugs are realised on a similarly based concept.

According to the traditional method “average bioequivalence” (ABE), two formulations are assumed being bioequivalent, when the calculation of a 90% confidence interval for the ratio of the geometric means (average) of the

PhUSE 2007

measures for the test and reference product are completely inside the defined bioequivalence limits of 80-125% (two-sided test procedure [1]).

$$0.8 \leq \frac{g\mu_T}{g\mu_R} \leq 1.25 \quad (1)$$

For that purpose so-called „bioequivalence-metrics“ as model-independent values are used for the comparison.

AUC (area under the concentration time curve) as measured value for the extension and $C_{\max}$ (maximum concentration observed) for the velocity of the absorption. For both parameters the geometric mean (average) of the test- ($g\mu_T$) and reference formulation ($g\mu_R$) are calculated (most of the pharmacokinetic parameters are assumed to be log-normal-distributed).

When logarithmising the equation (1), then the calculation of the estimator and the confidence interval can be obtained with the aid of a linear model, because then a normal distribution forms the basis of the difference of both means. The new obtained limits are then $\ln(-\theta_A) = -0.223$ and $\ln(\theta_A) = +0.223$.

$$\ln(-\theta_A) \leq \ln(\mu_T) - \ln(\mu_R) \leq \ln(\theta_A) \quad (2)$$

BIOEQUIVALENCE ANALYSIS FOR "HIGHLY VARIABLE DRUGS"

For drug products with a highly pharmacokinetic variability the ABE method is not a good choice, because in order to achieve a small confidence interval, the number of subjects required increases.

Drug products exhibiting intra-subject variability greater than 30% CV (coefficient of variation) are considered as highly variable. In the last few years there have been many discussions about methodological approaches that are able to increase the power of the study while keeping the α -level.

The issue of highly variable drugs/products in bioequivalence is sophisticated, but there is no binding regulatory declaration about the acceptance of these methods at this time.

In the presented study the „scaled average bioequivalence (SABE)“ method was applied. Its concept leans on the ABE method, but the acceptance limits get expanded with the increase of the intra-subject variability of the reference formulation [3]:

$$M_{as} = \frac{(\mu_T - \mu_R)^2}{\sigma_{WR}^2} < \theta_{as}^2 \quad (3)$$

M_{as}	= average, scaled bioequivalence measure
μ_T	= the mean of the test formulation
μ_R	= the mean of the reference formulation
σ_{WR}^2	= within-subject variance of the reference formulation
θ_{as}^2	= $\theta_A / \sigma_{w0} = \log(1.25)^2 / 0.25^2 \sim 0.7968871$

Unlike the ABE method, we need a study with repeated administration of the same formulation (replicate design), in order to be able to estimate σ_{WR}^2 . Figure 1 points out the principle. On the abscissa is the variance of the reference formulation. The ABE limits are horizontal, whereas the SABE limits intersect at the zero point and inflate with increasing σ_{WR}^2 .

At the time point of study evaluation, the assumption was that a drug product with a CV greater than 25-30 % can be considered as highly variable. Therefore σ_{w0} was defined as 0.25.

PhUSE 2007

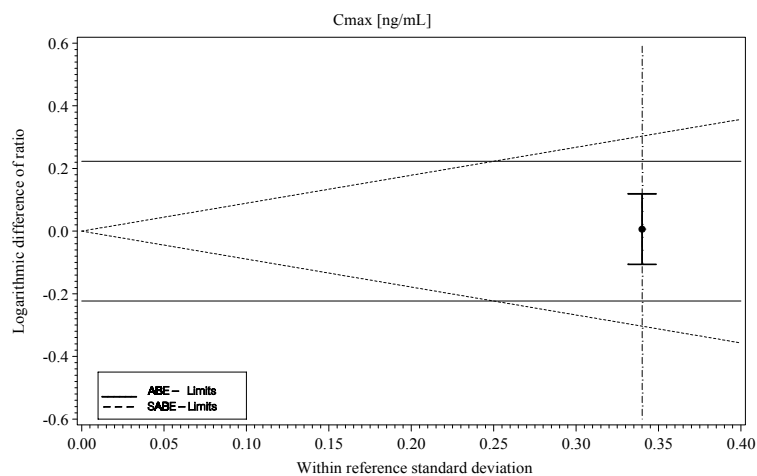


Figure 1: The ABE and SABE-limits, and the 90%-confidence interval of the T/R-ratio. The standard deviation of the reference formulation is here 0.34, σ_{w0} was defined as 0.25.

A major problem that comes with the SABE-method is, that the statistic M_{as} is not normally distributed any longer. Therefore there are suggestions to calculate special linearizations or bootstrap-methods. Although both of the methods were explored in the past, their regulatory acceptance is still unknown.

PRESENTED STUDY

BACKGROUND OF THE STUDY

The formulations of the drug agents, whose bioequivalence we wanted to prove in the study, have a high pharmacokinetic intra- and interindividual variability. In several former 2x2 crossover-studies the intraindividual ANOVA-CVs (between test and reference) of AUC and C_{max} were detected on a value of 45%, so that we have here the assumption of a highly variable drug. The intraindividual variability of repeated administration was not explored. Due to the existing result, it could be foreseen, that for an analysis of a 2x2 crossover-study with the ABE-method, more than 140 subjects would be needed in order to achieve a power of 90%.

STUDYDESIGN

The presented study was therefore conducted with a replicative design with the two sequences RTRT and TRTR. Due to this design, a smaller sample size is needed. The reference and test substances are administered to the subjects in a blinded fashion. The primary endpoints are AUC and C_{max} of two analytes of a drug agent. The sample size estimation showed a power of 90% with the SABE-method with a sample size of 60 subjects. With the ABE-method only a power of <65% could be achieved with this sample size. To be prepared for any dropouts, 66 subjects, (in 4 study cohorts) were included in the study.

ANALYSIS PLAN

The primary analysis will be the assessment of scaled average bioequivalence (SABE). In addition the unscaled average bioequivalence (ABE) is planned, because this method is known to be accepted by the authorities. As described in the introduction, the SABE will be performed with the help of bootstrap-method as well as with a special linearization. As the equation (3) contains a squaring of the limit, only a one-sided test will be applied. As a sensitivity analysis, the ABE was also calculated with the same bootstrap datasets as in the primary analysis.

IMPLEMENTATION

In this study the evaluation had to be finished one week after receiving the analysis data (calculated AUC and C_{max}), whereas a lot of displays for presenting the results were planned.

Because this time frame was too short to program the complete evaluation, the programming had to started earlier, basing on test data (frontloading concept). The complete programming of analysis should be finished before the final data were available.

The programming was based on SAS® version 8.2

PhUSE 2007

SIMULATION OF THE DATA STRUCTURE

The programming was based on a previous replicate design study that was already analysed with the ABE-method, but not with SABE. This study had a similar sample size (68 vs. 66 subjects), but the study design was different, as the sequences were TRRT and RTTR. This had to be adapted accordingly.

The ABE results from this study could be used for validation purposes of the new programming.

ANALYSIS DATA SETS - ADS

At Boehringer Ingelheim Pharma GmbH & Co. KG the data is stored in an Oracle*Clinical database. The evaluation is based on this data, but on so-called Analysis Data Sets (ADS), which are created specifically for the analysis. The ADS contain the derived analysis variables and are permanently stored and form the basis of the evaluation.

The main advantage of the ADS is the structured way of data handling. All derived study endpoints are derived only once within the ADS. For the presentation of data, almost no manipulation or derivations are necessary – they are replaced by simple data selection from the ADS. This simplifies the presentation in listings and tables, but also the validation of the results.

Another benefit is when analysing several studies within the same project as all ADS have the same data structure and a mandatory set of variables. So the data from the simulation study could be adapted easily to the actual study. All evaluations concerning the efficacy are based on these ADS.

CONTROLLING THE DATA SOURCE

During development of the programs all programs should refer to the ADS with the simulated data. After getting the data, the original ADS should be the data source. To avoid a mix-up of the both data sources later on, two LIBNAME-statements were created: one that includes the source for the simulated data, and one for the original data. As a lot of programs were created, the data source should be controlled from a central place.

An easy way to do this is to create a dynamically controlling of the path with the help of a central program. Here both LIBNAME-Statements should be defined. The name of the actually selected LIBNAME-Statement should then be assigned to a global macro-variable. The programs then access only the macro-variable, so that they don't need to be changed any more.

```
central program:      %GLOBAL inads;
                     LIBNAME ads_sim      "C:\KSFE\simu\";
                     LIBNAME ads          "C:\KSFE\ads\";
                     %LET inads = <ads , ads_sim> ;

accessing programs:  DATA adsAcces;
                     SET &inads..daten;
                     RUN;
```

EVALUATION ON MACRO-BASIS

The great amount of displays based on only a few display templates, because the display should look similar for the different endpoints and analyses. Altogether 124 tables and 72 figures were planned, whereas only 21 display templates were needed. Therefore one macro was created for each display template, so that the tables could be displayed only with macro calls.

SPECIAL METHOD: BOOTSTRAP

As mentioned before, the SABE-Metric M_{as} cannot be assumed to be normally distributed. Therefore the confidence intervals are estimated with the help of bootstrap methodology. When applying bootstrap methodology, an original dataset with the sample size n is repeatedly simulated with the same sample size via random sampling with replacement. From each bootstrap-dataset the variability of the reference-formulation (σ_{WR}^2) and the T/R-ratio of the bioequivalence-metric are estimated, in order to calculate M_{as} according to equation (3).

The distribution parameters and the SABE confidence intervals are assessed with the aid of the percentile method.

GENERATING OF RANDOM, BUT REPRODUCIBLE BOOTSTRAP SAMPLES

The technical demand of generating the subject-numbers was that the bootstrap samples should be random, but reproducible. In SAS® there is, among other things, the function RANUNI(seed) available, this can be used to generate from the uniform distribution on the interval (0, 1) a stream of random numbers. Seed is the initial starting point of this function. When using a seed greater than 0, the stream of random numbers can be replicated by using the same DATA step. The seed-number used by the function needed to be randomly. This was achieved by rolling a 10-sided dice for each of the four endpoints. The rolled number is random, but with the help of this number, the bootstrap samples were reproducible.

PhUSE 2007

GENERATING OF ALL SUBJECT-NUMBERS IN ONE DATASET

The aim was to create 2000 bootstrap samples with 66 subjects for each endpoint. In order not to create each of them separately, because then a new seed-number would be needed for each of these, all bootstrap samples were created in one dataset and then divided in 2000 datasets. That means that a stream of $2000 * 66$ (132.000) subject numbers must be created in one dataset.

As we wanted to have an equal number of patients for both treatment groups for our bootstrap, the original dataset was split in two, according to the treatment groups we had. Then these two datasets were resampled. That means we had a dataset with $2000 * 33$ patients for the first treatment group and the same for the other. Each bootstrap sample was identified with a variable (*datas*) in order to assign the splitting rule for the 2000 datasets. The resampled datasets were set together again in the end (*bootges*). This is shown in the following SAS® -macro *bootc*:

```
%MACRO bootc (datin = , samp = , tpatt1 = , tpatt2 = , obs1= , obs2 = , seedm = );

  %LOCAL i ptcnt;
  %LET ptcnt = 0;

  %DO i=1 %TO 2;
    PROC SORT DATA = &datin(KEEP = tpatt ptno) OUT=datin&i (WHERE=(tpatt="&&tpatt&i")) NODUPKEYS;
    BY tpatt ptno;
    RUN;
  %END;

  DATA tpatt1 tpatt2 ;
    RETAIN seed_1 &seedm;
    %DO i=1 %TO 2;
      PUT "trt nr &i";
      DO datas=1 TO &samp;
        DO ptnnew= %EVAL( &ptcnt + 1) TO %EVAL(&ptcnt + &&obs&i);
          selpt&i = ceil (ranuni (seed_1) * total&i);
          PUT selpt&i ;
          SET datin&i point = selpt&i nobs = total&i;
          OUTPUT tpatt&i;
        END;
      END;
    %END;
    STOP;
  RUN;

  DATA bootges;
    SET tpatt1 tpatt2;
  RUN;
  ...
%MEND bootc;

%bootc (datin=kpkpfas, samp=2000, tpatt1=TRTR, tpatt2=RTRT, obs1=33, obs2=33, seedm=7225449 );
```

In the log you can see how the macro gets resolved for the dataset part:

```
MPRINT(BOOTC): DATA tpatt1 tpatt2 ;
MPRINT(BOOTC): RETAIN seed_1 7225449;
MPRINT(BOOTC): PUT "trt nr 1";
MPRINT(BOOTC): DO datas=1 TO 2000;
MPRINT(BOOTC): DO ptnnew= 1 TO 33;
MPRINT(BOOTC): selpt1 = ceil (ranuni (seed_1) * total1);
MPRINT(BOOTC): PUT selpt1 ;
MPRINT(BOOTC): SET datin1 point = selpt1 nobs = total1;
MPRINT(BOOTC): OUTPUT tpatt1;
MPRINT(BOOTC): END;
MPRINT(BOOTC): END;
MPRINT(BOOTC): PUT "trt nr 2";
MPRINT(BOOTC): DO datas=1 TO 2000;
MPRINT(BOOTC): DO ptnnew= 1 TO 33;
MPRINT(BOOTC): selpt2 = ceil (ranuni (seed_1) * total2);
MPRINT(BOOTC): PUT selpt2 ;
MPRINT(BOOTC): SET datin2 point = selpt2 nobs = total2;
MPRINT(BOOTC): OUTPUT tpatt2;
MPRINT(BOOTC): END;
MPRINT(BOOTC): END;
MPRINT(BOOTC): STOP;
MPRINT(BOOTC): RUN;
```

PhUSE 2007

```
trt nr 1
21
18
18 ...

trt nr 2
11
22
20 ...
```

→ The interval can be expanded to the interval [1;33] with the multiplication of the total number of subjects (variable *total&i*) and the CEIL-function. The variable „datas“ is sequentially incremented after 66 resampled subjects in order to assign in the huge dataset ‚bootges‘ the splitting rule for the 2000 bootstrap-datasets.

	DATAS	PTNNEW	Patient number	TPATT
58	1	58	59	TRTR
59	1	59	19	TRTR
60	1	60	11	RTRT
61	1	61	28	TRTR
62	1	62	5	TRTR
63	1	63	31	TRTR
64	1	64	66	TRTR
65	1	65	65	RTRT
66	1	66	25	RTRT
67	2	1	16	TRTR
68	2	2	11	RTRT
69	2	3	57	TRTR
70	2	4	28	TRTR

Figure 2: abstract of the dataset with all 132.000 drawn subject-numbers

SPLITTING THE DATASETS

The division of this huge dataset required a lot of processing time. Initially a WHERE-Statement in a dataset was used, that was handled with a Do-Loop-processing.

```
%DO j = 1 %TO 2000;
  DATA boot&j;
    SET bootges;
    WHERE datas=&j;
  RUN;
%END;
```

With merging of the original dataset (see next chapter “MERGE compared to PROC SQL”) and the calculation of the test-statistic, 16 hours per endpoint were needed. Having only a 1 week time period for evaluation, this was not acceptable. The reason for this is that the program looks 2000 times sequentially for the variable *datas = &j* in the huge dataset *bootges*. Having a cache problem, in the beginning of the loop the response time of the program is very quick, getting slower the higher the index is. In the beginning the cache is empty and gets filled with every loop-step, so less capacity remains towards end. The figures (3) - (5) demonstrate this: for the index-number = 1 the selection of the first dataset needs 0.01 seconds (figure 3) and for the index-number 612 increases to 6.57 seconds (figure 4). When selecting the datasets directly, that means outside the loop, all selections were fast: *datas = 1* needed 0.01 seconds and *datas = 612* 0.03 seconds, as shown in figure (5).

PhUSE 2007

```

Command ==>
NOTE: 39779 line(s) written to external file.

NOTE: There were 66 observations read from the data set WORK.BOOTGES.
NOTE: The data set WORK.BOOTGESX has 66 observations and 5 variables.
NOTE: DATA statement used:
      real time    0.01 seconds
      cpu time     0.01 seconds

NOTE: The data set WORK.MB1 has 2040 observations and 34 variables.
NOTE: DATA statement used:
      real time    0.21 seconds
      cpu time     0.12 seconds

NOTE: There were 66 observations read from the data set WORK.BOOTGES.
NOTE: The data set WORK.BOOTGESX has 66 observations and 5 variables.
NOTE: DATA statement used:
      real time    0.03 seconds
  
```

Figure 3: selection from the huge dataset: index-number is 1

```

      cpu time     0.13 seconds

NOTE: There were 66 observations read from the data set WORK.BOOTGES.
NOTE: The data set WORK.BOOTGESX has 66 observations and 5 variables.
NOTE: DATA statement used:
      real time    6.57 seconds
      cpu time     0.01 seconds

NOTE: The data set WORK.MB612 has 2064 observations and 34 variables.
  
```

Figure 4: selection from the huge dataset: index-number is 612

```

Command ==>
NOTE: 39779 line(s) written to external file.

3426
3427 data nj2;
3428 set nj;
3429 where datas=1;
3430 run;

NOTE: There were 66 observations read from the data set WORK.NJ.
NOTE: The data set WORK.NJ2 has 66 observations and 5 variables.
NOTE: DATA statement used:
      real time    0.01 seconds
      cpu time     0.01 seconds

3431
3432 data nj3;
3433 set nj;
3434 where datas=612;
3435 run;

NOTE: There were 66 observations read from the data set WORK.NJ.
NOTE: The data set WORK.NJ3 has 66 observations and 5 variables.
NOTE: DATA statement used:
      real time    0.03 seconds
      cpu time     0.03 seconds
  
```

Figure 5: Direct selection from the huge dataset for the index-numbers 1 and 612

The time problem could be solved by using PROC SQL:

```

%DO j = 1 %TO 2000;
  PROC SQL;
    CREATE TABLE boot&j AS
      SELECT a.*, b.ptnonew, b.datas
      FROM &datin a, bootges b
      WHERE b.datas = &j AND a.ptno = b.ptno;
  QUIT;
%END;
  
```

The result was that the computing time could be reduced from 16 hours to 30 minutes for each endpoint, although also here a looping is used. But PROC-SQL-procedure ends with a QUIT-statement and clears at the same time the cache. It would have been an alternative to create an index on the variable datas. This would have had the same time results like the PROC-SQL-Statement.

```

PROC SQL;
  CREATE INDEX datas ON bootges(datas);
QUIT;
  
```

PhUSE 2007

MERGE COMPARED TO PROC SQL

Another reason why the programming for merging the data was done with PROC SQL is that the created bootstrap samples only include the subject-numbers, but no data. Therefore the subject information needs to be re-merged to the existing subject-numbers. Each subject has 4 records of information. Because of randomly sampling with replacement a subject can occur more than once in the bootstrap dataset. So it can occur that the subject information must be merged several times into the bootstrap dataset:

Table 1: Bootstrap-dataset with subject-numbers

seed	ptnnew	ptno	tpatt	datas
7725449	1	16	TRTR	1
7725449	2	1	TRTR	1
7725449	3	18	RTRT	1

Table 2: Original-dataset with several records for one subject

ptno	tpatt	Analyt	Endpoint	y	...
1	TRTR	A1	C _{max}	Y1	...
1	TRTR	A2	C _{max}	Y2	...
1	TRTR	A1	AUC	Y3	...
1	TRTR	A2	AUC	Y4	...

This procedure (Cartesian product) is easier to realize with PROC SQL than in a datastep, as shown in the following examples:

Dataset 'a' includes subject number '1' twice.

```
DATA a;
  ptnnew=1; ptno=1; OUTPUT;
  ptnnew=2; ptno=1; OUTPUT;
RUN;
```

	PTNNEW	PTNO
1	1	1
2	2	1

Figure 6: Dataset 'a'

In the original dataset 'b' the subject-number '1' has two records, which needed to be merged twice, as it occurs twice.

```
DATA b;
  ptno=1; value = 48.5 ; OUTPUT;
  ptno=1; value = 0 ; OUTPUT;
RUN;
```

	PTNO	VALUE
1	1	48.5
2	1	0

Figure 7: Dataset 'b'

Using a SQL-Statement (dataset 'c') a Cartesian product can be created easily. Joining both tables over the subject-number would give the desired results:

```
PROC SQL;
  CREATE TABLE c AS
  SELECT a.ptnnew, b.*
  FROM b, a
  WHERE b.PTNO = a.ptno ;
QUIT;
```

	PTNNEW	PTNO	VALUE
1	1	1	48.5
2	2	1	48.5
3	1	1	0
4	2	1	0

Figure 8: Dataset 'c': with PROC SQL

PhUSE 2007

With a simple MERGE (dataset ,d') in a dataset, one would not get the right results:

```
DATA d;  
  MERGE a(IN=in1) b(IN=in2);  
  BY ptno;  
  IF in1 AND in2;  
RUN;
```

	PTNNEW	PTND	VALUE
1	1	1	48.5
2	2	1	0

Figure 9: Dataset 'd': with DATA-STEP and Merge

In order to program a Cartesian product in a Dataset – accordingly to the PROC SQL – the subject-numbers must be selected one by one: In the following there is an example code for this:

```
DATA f;  
  DROP ptnox;  
  DO p1=1 TO nobs1;  
    SET a(RENAME=(ptno=ptnox))  
        NOBS=nobs1 point=p1;  
    DO p2=1 TO nobs2;  
      SET b NOBS=nobs2 point=p2;  
      IF ptno=ptnox THEN OUTPUT;  
    END;  
  END;  
  STOP;  
RUN;
```

DELETING THE TEMPORARY WORK-LIBRARY

After creation of the 2000 bootstrap datasets, the ratios of geometric mean of the test and reference metrics is calculated. The M_{as} -metric is then calculated with the help of the within-subject variance of the reference formulation according to equation (3) and then permanently stored in a dataset. From this permanent file the confidence interval was calculated using the percentile method. The WORK-library now consists of 2000 bootstrap datasets and 2000 datasets for calculations for each endpoint. This folder should be cleaned up, in order to prevent access from other bootstrap datasets generating programs. When deleting the WORK-library with the PROC DATASETS Procedure and the Kill-Option, the program crashed in our case.

```
PROC DATASETS LIB=work MEMTYPE=DATA NOLIST KILL;  
QUIT;
```

The reason for this crashing could not be localised during the implementation period, but we assumed, that cache problems caused it. With the use of an x-command a work-around for this was programmed. This executes instructions on operating-system level. After identifying the physical path of the WORK-library, the DEL-command can delete the content of the path:

```
DATA _null_;  
  LENGTH str2 $500;  
  str2='del ' !! PATHNAME('WORK') !! '\*.sas7bdat';  
  str2=QUOTE(TRIM(str2));  
  CALL SYMPUT('CMD',TRIM(str2));  
  PUT str2=;  
RUN;  
  
%PUT &cmd;  
"del C:\MedSAS\SASWork\_TD2388\*.sas7bdat"  
  
X &cmd;
```

FURTHER METHODOLOGICAL APPLICATIONS

After finishing the planned analysis for the study report, further methodical analyses were made in order to gain information for further similar trials. For example the impact of covariates on the estimation of confidence intervals was explored, e.g. gender or weight.

An interesting examination was the simulation of a study with less subject numbers as in the original trial. For instance when including the first 33 subjects instead of all 66 subjects, the results from the ABE-estimation would not have lead to a successful bioequivalence-study. The confidence interval would have been wider and the point estimator differed clearly from 100 % (see figure 10). However the SABE-analysis would have still been successful.

By structuring the programs in macros and macro calls, it was simple to create outputs for new questions.

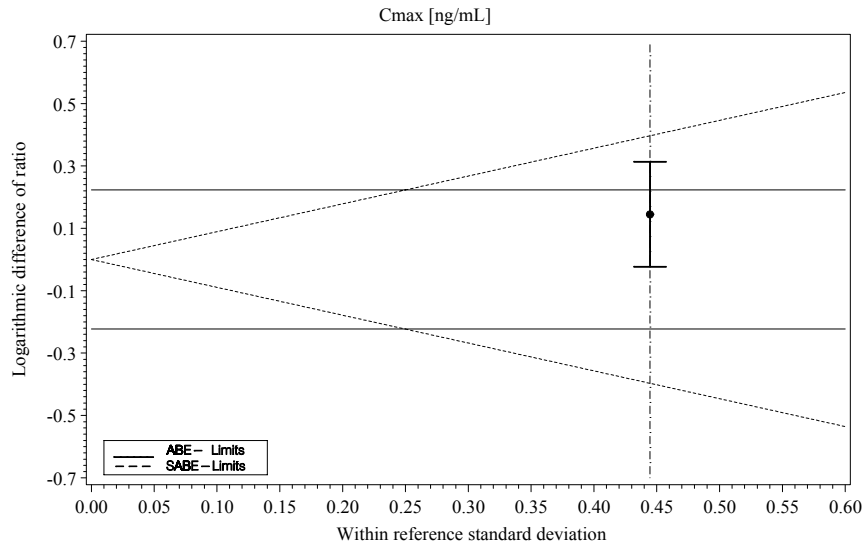


Figure 10: the ABE and SABE-limits, as well as the 90%-confidence-intervall of the T/R-ratio for the first 32 subject besides all same conditions as in figure 1.

CONCLUSION

As planned, the results of the analysis could be presented in a validated way one week after receiving the data – the frontloading on the basis of the simulated datasets was successful. The new analysing methods were tested intensively very early, so the problems that occurred during implementation could be identified and solved.

PhUSE 2007

REFERENCES

- [1] Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *J Pharmacokinet Biopharm.* 1987 Dec;15(6):657-80.
- [2] Guidance for industry: *Bioavailability and Bioequivalence Studies for Orally Administered Drug Products - General Considerations (Revision 1)*. U.S.Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (**CDER**), March 2003.
- [3] Guidance for industry: *Statistical Approaches to Establishing Bioequivalence*. U.S.Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (**CDER**), January 2001.
- [4] Laszlo Tothfalusi and Laszlo Endrenyi *Limits for the Scaled Average Bioequivalence of Highly Variable Drugs and Drug Products*, Pharmaceutical Research, Vol. 20, No. 3, March 2003

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Feryal Tanriverdio

Boehringer Ingelheim Pharma GmbH & Co KG

Birkendorfer Strasse 65

Biberach an der Riss / 88400

Work Phone: (+49) 7351 / 54-94633

Fax: (+49) 7351 / 83-94633

Email: Feryal.Tanriverdio@boehringer-ingelheim.com

Dr. Arne Ring

Boehringer Ingelheim Pharma GmbH & Co KG

Birkendorfer Strasse 65

Biberach an der Riss / 88400

Work Phone: (+49) 7351 / 54-94633

Fax: (+49) 7351 / 83-94633

Email: Arne.Ring@boehringer-ingelheim.com

Brand and product names are trademarks of their respective companies.