

BeST – Bioequivalence Study Template

Pekka Heikkilä, StatFinn Oy, Turku, Finland
Tuomas Kemppainen, StatFinn Oy, Turku, Finland
Tiina Kirsilä, StatFinn Oy, Turku, Finland
Pasi Korhonen, StatFinn Oy, Espoo, Finland

ABSTRACT

BeST product automatically creates a complete pharmacokinetic analysis report from concentration data. This paper is presenting the challenges and solutions of packaging the service within a SAS software tool. BeST is initially set up for reporting bioequivalence studies according to the ICH guideline (CPMP/EWP/QWP/1401/98), but because of the modular approach it can be used in wide range of pharmacokinetic studies.

The main steps of the process are the following:

1. Receiving concentration data
2. Preparation of analysis datasets
3. Calculation of pharmacokinetic parameters
4. Generation of appendices
5. Creation of pharmacokinetic analysis report

The data is received via secure web portal and thereafter all process steps are carried out with the SAS® software. Full audit trail is kept throughout the whole process within a version controlled platform. In this paper the basic principles of each module are summarized. Especially solutions used in the SDTM dataset creation, PK parameter calculation and writing of report using SAS software are described.

INTRODUCTION

One of the most frequently occurring study types in drug development is a pharmacokinetic two-by-two cross-over study carried out to evaluate the bioequivalence of two medicinal products. The bioequivalence studies are strictly regulated [1] and are therefore repeated in similar and standardized manner. Bioequivalence is achieved if the rate and extent of absorption for the two products are similar within predefined acceptance limits.

Pharmacokinetic parameters are derived from the concentration-time curve to assess the rate and extent of absorption and elimination. The area under the concentration time curve (AUC) reflects the extent of exposure. The maximum plasma concentration (C_{max}) or peak exposure and the time to maximum plasma concentration (T_{max}) are parameters that are influenced by absorption rate.

In this paper the whole process from concentration data to a final pharmacokinetic analysis report including pharmacokinetic calculations and all necessary outputs for appendices is described. There have been numerous challenges during the development of the product and we describe the solutions chosen to solve the issues. It is worth to mention that whole process is implemented within the SAS software.

PROCESS

To upload the data, clients are requested to log on to a web interface where they can manage their projects, upload data and download reports. The web interface does not store any project related data which is uploaded directly to separate data server. The same applies for all deliverables which are uploaded directly to client from the data server. All connections are secured with the SSL encryption. All items in the server are under version control including data, programs, logs, outputs and reports.

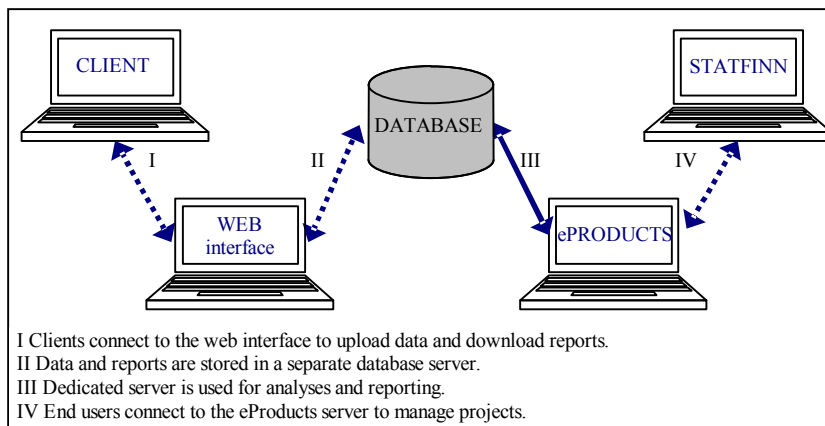


Figure 1 Overview of upload and download process

After the data has been downloaded to the database the BeST tool can be launched from a dedicated eProducts server. BeST is a modularized package of SAS macros for data derivation, pharmacokinetic parameter calculation, data presentation, statistical analysis and reporting. The first step is data derivation where data are imported to the system and analysis datasets are created. The export data in SDTM format [2] are also created during this step. The second step is to calculate the pharmacokinetic parameters with a summary of calculations and individual plots. In the third step all individual outputs including descriptive summary statistics, formal statistical analysis, graphs and listings are generated. Before creation of the report all individual outputs are combined as appendices following the ICH Topic E 3 guideline [3]. As a last step a complete pharmacokinetic analysis report is generated.

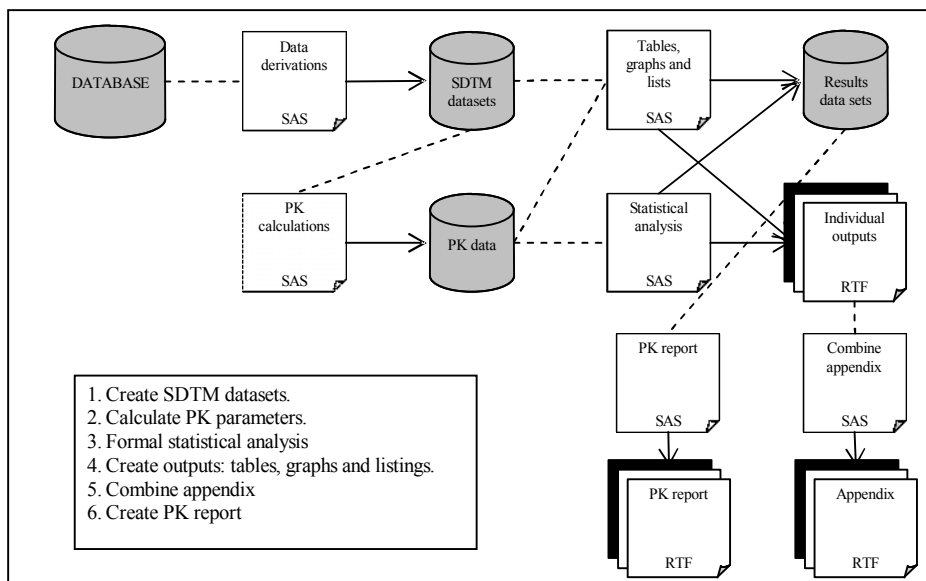


Figure 2 Overview of the SAS system design from the input data to the final report.

PhUSE 2010

DATA DERIVATION

In this chapter we describe what the minimum required input data for the process is and how the data are derived to SDTM datasets. We decided to follow the CDISC standards as far as possible to support the regulatory submission. This also allows us to further develop the system and gives access to written data specification. All datasets are exported using the SDTM version 3.1.2 [4].

CONCENTRATION DATA (PC)

The main source data is concentration data. The required variables are unique identifier for each subject (USUBJID), description of the sample (ETCD, PCTEST), collection time of the sample (PCDTC), obtained result (PCORRES) and reference time (PCRFTDTC). Some derivation is needed to achieve the format defined in SDTM specifications which are summarized below.

Table 1 SDTM data specifications of PC data and derivations from input data

Variable name	Variable label	Data derivation	Comment
STUDYID	Study Identifier	=TA.STUDYID	
DOMAIN	Domain Abbreviation	="PC"	
USUBJID	Unique Subject Identifier	="STUDYID" "_" "SUBJID"	SUBJID is collected.
PCSEQ	Sequence Number	_N_	
PCTESTCD	Pharmacokinetic Test Short Name	="ETCD" "_" "PCTESTC"	PCTESTC is collected as abbreviation for PCTEST
PCTEST	Pharmacokinetic Test Name	Input	
PCORRES	Result or Finding in Original Units	Input	
PCORRESU	Original Units	Input	
PCSTRESC	Character Result/Finding in Std Format	=PCORRES	Assuming that standard format is collected.
PCSTRESN	Numeric Result/Finding in Standard Units	if PCSTRESC='LLOQ' then =.I else =INPUT(PCSTRESC, best12.)	
PCSTRESU	Standard Units	=PCORRESU	
PCSPEC	Specimen Material Type	="PLASMA"	Assuming that specimen for sample is plasma.
PCLLOQ	Limit of Quantitation	Input	
VISITNUM	Visit Number	Input	Should be the same as in TA.
VISIT	Visit Name	Input	Should be the same as in TA.
PCDTC	Date/Time of Specimen Collection	Derived. See below.	Separate values for year, day, month, hour and minutes.
PCDY	Actual Study Day of Specimen Collection	=PCDTC – first.PCRFTDTC + 1	
PCTPT	Planned Time Point Name	Input	Should use hh:mm format
PCTPTNUM	Planned Time Point Number	Derived	Order number of PCTPT.
PCELTM	Planned Elapsed Time from Time Point Ref	Derived. See below.	Derived from PCTPT.
PCRFTDTC	Date/Time of Reference Point	Derived. See below.	Separate values for year, day, month, hour and minutes.

PhUSE 2010

STUDY DESIGN DATA (TA)

In the study design data we are collecting information regarding periods (VISITNUM), treatments (ETCD) and treatment sequences (ARMCD). This is the information needed to analyze the data according design of the study. The description fields (ELEMENT and VISIT) are used as input macro variables while generating outputs and report.

Table 2 SDTM data specifications for TA data and derivations from input data

Variable name	Variable label	Data derivation	Comment
STUDYID	Study Identifier	Input	
DOMAIN	Domain Abbreviation	= "TA"	
ARMCD	Planned Arm Code	Input	
ARM	Description of Planned Arm	Derived	Combining treatments according to the sequence order using ARMCD, VISITNUM, ETCD and ELEMENT.
TAETORD	Order of Element within Arm	=VISITNUM	Collected as input data
ETCD	Element Code	Input	
ELEMENT	Description of Element	Input	
EPOCH	Epoch	=VISIT	Collected as input data

OTHER DATA

In addition to the minimum required data PC and TA described above, additional information, like name of the sponsor, active ingredient, product names etc., are collected form the web portal. This additional information is used as meta information for the report. The BeST tool can also report other datasets (DM, VS, PE, AE etc...) and use calculated PK parameters as its input, but this is out of scope of this paper.

DATETIME FORMAT

In the SDTM specification, datetime format is set to ISO 8601 [5]. Shortly this means that the date is presented as "year-month-day" and time as "hour:minute". Date and time are separated with "T". For example, the time "*quarter past one in the afternoon on 13th February in year 2010*" is presented as "2010-02-13T13:15". We decided to use this format throughout the application. However, we are not requiring the input in this format for user convenience. Instead we require separate fields for year, month, day, hour and minute to avoid misinterpretation.

After combining the fields into SAS datetime format one can simply use `informat` statement to have datetime presented in ISO 8601 format. The core SAS code is presented below.

```
pcdtc_date=INPUT (COMPRESS (pcdtc_yyyy||pcdtc_mm||pcdtc_dd),YYMMDD8.);
pcdtc_time=INPUT (pcdtc_tm,TIME.);
pcdy=pcdtc_date*24*60*60 + pcdtc_time;
pcdtc=PUT (pcdy_,E8601DT.);
```

The elapsed time in ISO 8601 is a bit different compared to what we have used to seen. As a prefix we have a letter "P" followed by the duration related to the reference time in days, hours and minutes. For example "thirty one and half hour after the start of treatment" is presented as P1D7H30M0S. Elapsed time prior the reference time is prefixed with "-".

After calculation of the difference between reference time and sample collection time we can use `informat` and `format` statements to transform the data to desired ISO 8601 format. The core SAS code is presented below.

```
pceltm_=tranwrd (compress ('PT' || put (elapsed_,time.)), 'PT-', '-PT');
pceltm=put (input (pceltm_, $N8601B.), N8601E.);
```

PHARMACOKINETIC PARAMETER CALCULATIONS

The pharmacokinetic parameters are calculated from plasma concentration measurements over time based on non-compartmental analysis [6]. Maximum concentration in the plasma (C_{max}) is obtained directly from the data being the highest plasma concentration value as is the time when C_{max} is reached (T_{max}). The last observed concentration above limit of quantification (C_{last}) and the time when this is reached (T_{last}) are also obtained directly from the data.

PhUSE 2010

The area under the plasma concentration curve (AUC_t) is calculated using the linear trapezoidal rule [7]. The trapezoidal rule is a method used to numerically approximate the area under a curve. The observed concentrations are connected by straight lines which generate segments with an area of square and a triangle. The sum of these areas gives the final numerical result of the area under the curve. This trapezoidal rule can be expressed as follows:

$$AUC_t = \sum_{i=0}^{n-1} \frac{t_{i+1} - t_i}{2} (c_i + c_{i+1})$$

Derivation of AUC_t from the PC data can be done with the following SAS code.

```
DATA auc;
SET pc;
  x = elapsed;
  y = pcstresn;
  xpre = LAG(x);
  ypre = LAG(y);
  xdif = x - xpre;
  auc = xdif * (( y + ypre ) / 2);
RUN;
```

The elimination rate constant (K_{el}) describes the rate of decrease in concentration per unit time which is estimated from the log-linear terminal part of the concentration-time curve, as the slope of natural logarithm of concentration against time. K_{el} is calculated using Huber's M robust regression [8]. Data points that belong to the terminal log-linear phase are evaluated based on coefficient of determination (R²). For each patient the highest R² is selected and the corresponding regression model is used in the estimation of K_{el}. Core SAS code for the robust regression is presented below.

```
PROC ROBUSTREG METHOD=M(CONVERGENCE=COEF(EPS=0.0001) MAXITER=999999);
MODEL log_conc = elapsed;
RUN;
```

The elimination half life (T_½) of a substance is the time taking to half the blood plasma concentration which is calculated using the equation:

$$T_{1/2} = \ln(2) / K_{el}.$$

To estimate the AUC beyond last observation (C_{last}) is handled as a zero-time concentration following an intravenous bolus and that kinetics of the drug does not change beyond C_{last}. The area under the concentration curve extrapolated to infinity (AUC_∞) is then calculated using equation:

$$AUC_{\infty} = AUC_t + (C_{last} / K_{el}).$$

Finally the extrapolated proportion (AUC_{extrap}) of AUC_∞ is calculated using the equation:

$$AUC_{extrap} = 100 * [(C_{last} / K_{el}) \div AUC_{\infty}].$$

STATISTICAL ANALYSIS

The assessment of bioequivalence is based upon 90% confidence intervals for the ratio of the population geometric means (test/reference) for the pharmacokinetic parameter [9]. This method is equivalent to using two one-sided tests, each at the 5% significance level. A confidence interval for the difference on the log scale between formulations is obtained from analysis of variance model. This confidence interval is then back-transformed to obtain the desired confidence interval for the ratio on the original scale. Below is the core SAS code for the statistical analysis and geometric means ratio for the bioequivalence.

```
PROC MIXED DATA=pp ORDER=internal;
CLASS usubjid armcd visitnum etcd;
MODEL log_var=armcd visitnum etcd;
RANDOM usubjid(armcd);
ESTIMATE 'Test / Reference' etcd -1 1 / ALPHA=0.1 CL;
LSMEANS etcd / ALPHA=0.1 CL;
RUN;
```

PhUSE 2010

```
DATA gm;
SET est1;
  gm_est=PUT(EXP(estimate),10.3);
  gm_lower= PUT(EXP(lower),10.3);
  gm_upper= PUT(EXP(upper),10.3);
RUN;

DATA est;
SET est2;
  r_est=TRIM(COMPRESS(put(100*EXP(estimate),10.2)));
  r_lower=TRIM(COMPRESS(put(100*EXP(lower),10.2)));
  r_upper=TRIM(COMPRESS(put(100*EXP(upper),10.2)));
RUN;
```

REPORTING

All individual outputs tables, graphs and listings generated earlier are finally combined into appendices and the pharmacokinetic analysis report is written. The challenges conquered during these stages are presented next.

COMBINING APPENDICES

All outputs are first created as individual RTF (Rich Text Format) outputs using ODS RTF statements [10]. The combination of an appendix is done by reading in the RTF document as text and then printing them to single file and adding title page, table of contents, page numbering, header and footer.

Immediately after the RTF start page tag, the title page is created in the beginning of the file embedding some collected meta- information as presented in the following SAS code.

```
DATA combine;
LENGTH text $1024.;

/* Start of RTF document */
text="{\rtf1\ansi\ansicpg1252\uc1\deff0\deflang1035\deflangfe1033"; output;

/* Title page */
text="\qc\line\line\line\line\line\line\line\line \fs36 \b Pharmacokinetic
Analysis Report - Appendix XX"; output;
text="\qc\line\line\line\b Tables, Figures and Listings"; output;
text="\qc\line\line\line\fs24 \b Study: &STUDY."; output;
text="\qc\line\line\line\b0 Client: &SPONSOR."; output;
text="\qc\line\line\line Confidential \sect\sectd\pard"; output;

/* Table of contents */
text='{ \b Table of contents \par \par} {\field{\*\fldinst {\TOC \f \h}}}
\sect'; output;

RUN;
```

We also need to set up a page break datasets which will be added between each output.

```
DATA rtf_brk;
LENGTH txt $1024.;
text='\sect\sectd\sbkpage\plain\ansi';
RUN;
```

PhUSE 2010

Then we read each output in as an ASCII file and remove unnecessary RTF tags.

```
DATA rtf_in;  
LENGTH text $1024.;  
INFILE "&rtfin_path\&current" LRECL=1024;  
INPUT;  
    text=_infile_;  
RUN;  
  
DATA rtf_in2;  
LENGTH filename $128. text $1024.;  
SET rtf_in;  
  
/* Remove \rtf1, {, and } from the RTF output and combine all files */  
text= TRANWRD (text, '{\rtf1','');  
IF last.file_num THEN DO;  
    IF LENGTH(text)<=1 THEN text='';  
    IF LENGTH(text)>1 THEN text=SUBSTR(text,1,LENGTH(text)-1);  
END;  
  
/* Delete unnecessary enter after page break*/  
text=TRANWRD(text, '{\page\par}','{\page}');  
RUN;
```

Finally we combine the files and add page breaks between the outputs.

```

/* Combined rtf file */
DATA combine;
    %IF &file_num=1 %THEN %DO; SET COMBINE rtf_in2;          %END;
    %ELSE                  %DO; SET COMBINE rtf_brk rtf_in2; %END;
RUN;

```

At the end we should have an appendix document which contains title page, table of contents and all outputs.

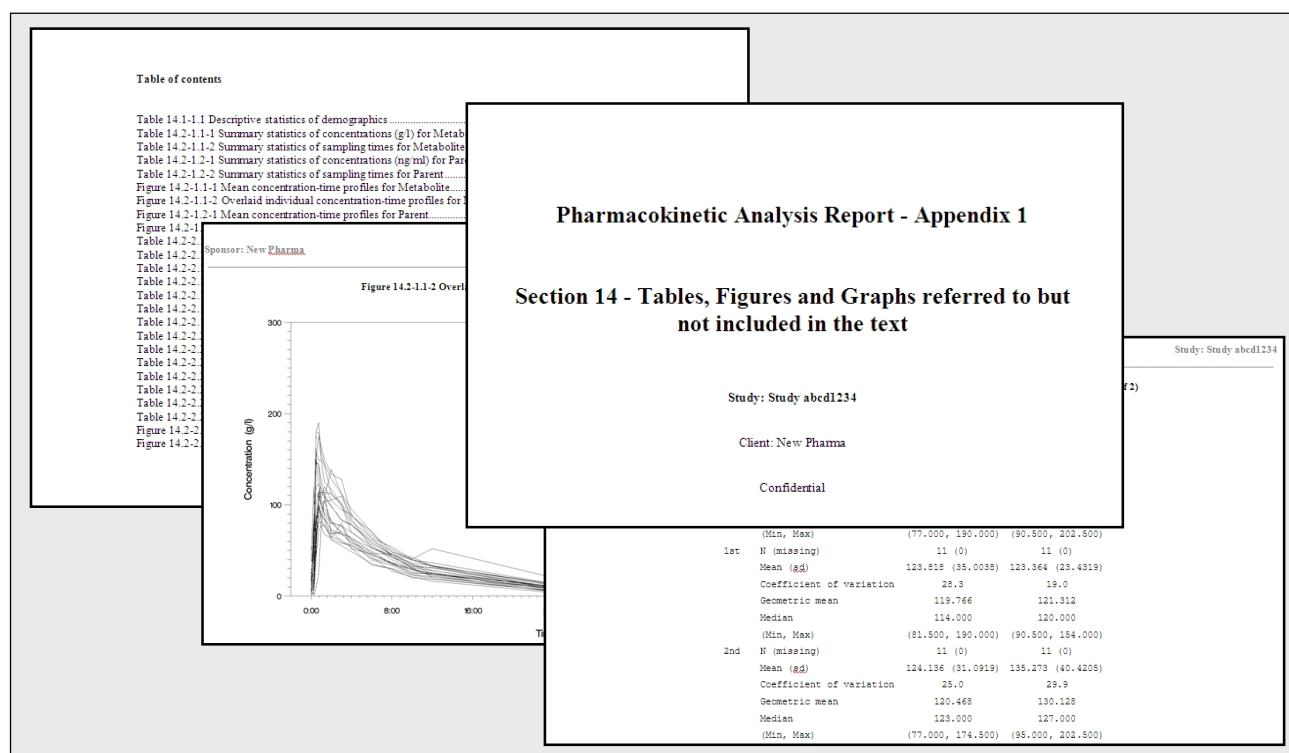


Figure 3 Screenshots of Appendix layout

PHUSE 2010

WRITING PHARMACOKINETIC ANALYSIS REPORT

Finally, writing of the report is combination of writing predefined text by replacing keywords with meta-data parameters (such as sponsor, active ingredient etc...) and creating tables and graphs within the text. The title page and table of contents are generated similarly as described in the previous section. After the title page each section is generated as a data set containing the predefined text with additional meta-data information embedded into the text.

For example, the creation of chapter of study type and design can be created as presented in the following core SAS code.

```
/* Start of the chapter */
ODS RTF TEXT="{\par \tc \b 2 Study type and design \tcl1}";

/* First section */
ODS RTF TEXT="{\S={outputwidth=100% just=c} \R/RTF'\qj\par' This &PHASE. study
was a randomized, &BLINDSTATUS., two-period, single-dose cross-over trial with
at least &WASHOUTLENGTH. wash-out period between the treatment periods."};

/* Second section */
ODS RTF TEXT="{\S={outputwidth=100% just=c} \R/RTF'\qj\par' The test and
reference products were as follows:"};

/* Table of product information */
DATA prdoducts;
LENGTH a $10. b $200.;
DO;
  a="Test";
  b="&TESTPRODUCT. &TESTSTRENGTH. &TESTFORMULATION. by &SPONSOR.";
  OUTPUT;
END;
DO;
  a="Reference";
  b="&REFPRODUCT. &REFSTRENGTH. &REFFORMULATION. by &REFSPONSOR.";
  OUTPUT;
END;
LABEL a="Product"
      b="Description";
PROC PRINT DATA=tmp NOOBS LABEL;
RUN;
```

At the last part of the code an in-text-table is created directly to the report. Similarly any other in-text-table can be created. It is also possible to use `proc report` to create in-text-tables. Creating the report differs slightly from creating the appendices, because the report is generated directly between `ODS RTF` statements.

2 Study type and design

This Phase I study was a randomized, Double-blinded, two-period, single-dose cross-over trial with at least 14 days wash-out period between the treatment periods.

The test and reference products were as follows:

Product	Description
Test	Medic 500 mg Tablet by New Pharma
Reference	Drug 500 mg Capsel by Old Pharma

Figure 4 Screenshot of Pharmacokinetic Analysis Report

PhUSE 2010

CONCLUSION

This paper summaries how an automated reporting tool with SAS software can be developed. Development of an automated process has technical challenges, but to combine the different solutions towards a common goal requires also project and team management to succeed.

One of the main achievements in this project is to use statistical methodologies to determine elimination half life to calculate and derive the pharmacokinetic parameters. Several different approaches were compared and their sensitivity was tested against existing studies and with simulation.

Another achievement was the development of a tool which uses only SAS software to produce all calculations, statistics, graphs, appendices and the report directly from the input data. For BeST the only required data is concentration data and design data.

REFERENCES

1. EMEA, CHMP. *Note for Guidance on The Investigation of Bioavailability and Bioequivalence*. CPMP/EWP/QWP/1401/98, 2010.
2. CDISC, Submission Data Standards Team. *Study Data Tabulation Model*. Version 1.2 Final, 2008.
3. EMEA, CPMP. *Note for Guidance on Structure and Content of Clinical Study Report*. CPMP/ICH/137/95, 1996.
4. CDISC Submission Data Standards Team. *Study Data Tabulation Model Implementation Guide: Human Clinical Trials*. Version 3.1.2 Final, 2008.
5. ISO. *Data elements and interchange formats -- Information interchange -- Representation of dates and times*. VISO 8601:2004, 2004.
6. Gibaldi M., Perrier D. *Pharmacokinetics*. Marcel Dekker Inc., 1982.
7. Rowland M., Tozed T. N. *Clinical Pharmacokinetics Concepts and Applications*. Williams & Walkins, 1995.
8. Rousseeuw, R.J. & Leroy, A.M.. *Robust Regression and Outlier Detection*. Wiley, 1987.
9. Chow S.-C., Liu J.-P.. *Design and Analysis of Bioavailability and Bioequivalence Studies*. Chapman & Hall, 2009.
10. Haworth, L. E. Zender, C. L., Burlew, M. M.. *Output Delivery System the Basics and Beyond*. SAS Institute Inc., 2009.

CONTACT INFORMATION

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

Pekka Heikkilä
StatFinn Oy
Linnankatu 9 b A 1
FIN-20100 Turku
Work Phone: +358 50 321 6006
Email: pekka.heikkila@statfinn.fi
Web: www.statfinn.fi

Brand and product names are trademarks of their respective companies.