Validation of Bioequivalence Test Performed by BE R package

Sungpil Han shan@acp.kr

2018-10-23

Contents

1 Introduction

`BE` R package (Bae 2018) can analyze bioequivalence study data with industrial strength. The current version `BE` performs bioequivalency tests for several variables of a 2x2 crossover study in a data file. To establish BE, the calculated confidence interval should fall within a BE limit, usually 80-125% for the ratio of the product averages. (Drug Evaluation et al. 2001; Chow and Liu 2008; Hauschke, Steinijans, and Pigeot 2007) In this document, the author performed validation of bioequivalence test performed by `BE` R package.

2 Methods

2.1 Datasets

A simulated dataset, `NCAResult4BE` is shown in Appendix A.

2.2 Calculation `BE`

The required R packages are following.

```r
library(BE)        # install.packages("BE", repos="http://r.acr.kr")
library(tidyverse) # install.packages("tidyverse")
```

`tab_r_be_results` function uses `BE::test2x2()` to calculate the 90% confidence interval.

```r
tab_r_be_results <- function(parameter){
  BE::test2x2(NCAResult4BE, parameter)[[4]] %>%
  as.data.frame() %>%
  mutate(Analysis = 'R: BE package') %>%
  select(Analysis, `Lower Limit`, `Point Estimate`, `Upper Limit`)
}
```

2.3 Calculation `SAS`

`SAS PROC GLM` and `SAS PROC MIXED` in SAS version 9.0 with SEQ (sequence), TRT (treatment), SUBJ (subject), PRD (period) variables, and LNAUCL or LNCMAX denoting the response measure.
A part of examplary SAS program statement is following and the full SAS scripts are appended in Appendix B.

```sas
PROC GLM DATA=BE OUTSTAT=STATRES; /* GLM use only complete subjects. */
CLASS SEQ PRD TRT SUBJ;
MODEL LNAUCL = SEQ SUBJ(SEQ) PRD TRT;
RANDOM SUBJ(SEQ)/TEST;
LSMEANS TRT /PDIFF=CONTROL('R') CL ALPHA=0.1 COV OUT=LSOUT;
```

```sas
PROC MIXED DATA=BE; /* MIXED uses all data. */
CLASS SEQ TRT SUBJ PRD;
MODEL LNAUCL = SEQ PRD TRT;
RANDOM SUBJ(SEQ);
ESTIMATE 'T VS R' TRT -1 1 /CL ALPHA=0.1;
ODS OUTPUT ESTIMATES=ESTIM COVPARMS=COVPAR;
```

A function, `tab_sas_proc_results()` reads SAS analysis results exported to Microsoft Excel files (`.xls`) and converted to comma separated version file (`.csv`). It returns a data frame of 90% confidence interval calculated either `PROC GLM` or `PROC MIXED` in SAS version 9.0.

```r
tab_sas_proc_results <- function(filename, skip_no, analysis_name){
  read_lines(filename, skip = skip_no, n_max = 2) %>%
  paste(collapse='\n') %>% read_csv() %>%
  mutate(Analysis = analysis_name) %>%
  select(Analysis, `Lower Limit` = LL, `Point Estimate` = PE, `Upper Limit` = UL)
}
```

3 Results

3.1 AUC$_{\text{last}}$

Comparison between BE and SAS is shown in Table 1.

```
AUClast_R_BE <- tab_r_be_results("AUClast")

AUClast_proc_glm <- tab_sas_proc_results('sas/SAS_results_AUClast.csv',
                                          skip = 206, 'SAS: PROC GLM')
AUClast_proc_mixed <- tab_sas_proc_results('sas/SAS_results_AUClast.csv',
                                            skip = 278, 'SAS: PROC MIXED')

# Combine all analyses of AUClast
bind_rows(AUClast_R_BE, AUClast_proc_glm, AUClast_proc_mixed) %>%
  kable(longtable = TRUE, booktabs = TRUE, caption = 'AUClast', digits = 5)
```

Table 1: AUClast

Analysis	Lower Limit	Point Estimate	Upper Limit
R: BE package	0.88944	0.95408	1.02341
SAS: PROC GLM	0.88944	0.95408	1.02341
SAS: PROC MIXED	0.88944	0.95408	1.02341

3.2 C$_{\text{max}}$

Comparison between BE and SAS is shown in Table 2.

```
Cmax_R_BE <- tab_r_be_results("Cmax")

Cmax_proc_glm <- tab_sas_proc_results('sas/SAS_results_Cmax.csv',
                                       skip = 206, 'SAS: PROC GLM')
Cmax_proc_mixed <- tab_sas_proc_results('sas/SAS_results_Cmax.csv',
                                         skip = 278, 'SAS: PROC MIXED')

# Combine all analyses of Cmax
bind_rows(Cmax_R_BE, Cmax_proc_glm, Cmax_proc_mixed) %>%
  kable(longtable = TRUE, booktabs = TRUE, caption = 'Cmax', digits = 5)
```

Table 2: Cmax

Analysis	Lower Limit	Point Estimate	Upper Limit
R: BE package	0.90136	0.97984	1.06515
SAS: PROC GLM	0.90136	0.97984	1.06515
SAS: PROC MIXED	0.90136	0.97984	1.06515

4   Conclusion

There is no discrepancy between results from BE and SAS. We also performed multiple analyses with the actual clinical trial datasets and have found no differences (data not shown: confidential). Noncompartmental analysis performed by the open-source R package, NonCompart can be **qualified and validated** enough to acquire the identical results of the commercial statistical software, SAS.

Please report issues regarding validation of the R package to https://github.com/asancpt/BEreport/issues.

Affiliation:
Sungpil Han
M.D/Ph.D, Resident
Department of Clinical Pharmacology and Therapeutics,
Asan Medical Center, University of Ulsan,
Seoul 05505, Republic of Korea
E-mail: shan@acp.kr
URL: www.github.com/shanmdphd

A   Raw data

Table 3: The raw data used for analysis by BE and SAS PROC

SUBJ	GRP	PRD	TRT	AUClast	Cmax	Tmax
1	RT	1	R	5018.927	1043.13	1.04
1	RT	2	T	6737.507	894.21	1.03
2	TR	1	T	4373.970	447.26	1.01
2	TR	2	R	6164.276	783.92	1.98
4	TR	1	T	5592.993	824.42	1.97
4	TR	2	R	5958.160	646.31	0.97
5	TR	1	T	3902.590	803.70	0.80
5	TR	2	R	4620.156	955.30	0.74
6	RT	1	R	3735.274	995.34	1.02
6	RT	2	T	4257.802	816.33	1.00
7	RT	1	R	4314.993	608.99	0.95
7	RT	2	T	5030.372	806.57	0.74
8	RT	1	R	6053.098	1283.67	0.72
8	RT	2	T	5790.067	822.95	1.03
9	RT	1	R	4602.582	679.39	0.74
9	RT	2	T	6042.462	556.55	0.98
10	RT	1	R	8848.988	1136.91	1.03
10	RT	2	T	7349.822	1082.79	0.97
11	TR	1	T	3054.096	547.73	2.02
11	TR	2	R	4719.175	984.69	0.54
13	RT	1	R	4828.682	615.17	1.00
13	RT	2	T	4175.434	692.26	0.97
14	RT	1	R	4566.275	864.56	1.03
14	RT	2	T	5042.649	1122.75	0.75
15	TR	1	T	4950.980	719.40	0.97
15	TR	2	R	4959.554	660.17	0.96
16	RT	1	R	4577.432	609.64	3.01
16	RT	2	T	4773.723	807.65	1.01
17	RT	1	R	6462.652	861.56	2.02
17	RT	2	T	5246.032	1187.75	0.73
18	TR	1	T	4754.625	919.87	0.77
18	TR	2	R	3214.809	1042.84	0.53
19	TR	1	T	7619.304	1089.84	3.00
19	TR	2	R	5210.569	1127.94	2.04
20	TR	1	T	5063.471	1191.46	0.71
20	TR	2	R	6406.634	1069.19	1.00
21	RT	1	R	5580.289	742.67	0.97
21	RT	2	T	6304.119	447.85	0.99
22	RT	1	R	4398.887	682.73	2.02
22	RT	2	T	3760.359	669.01	1.04
23	TR	1	T	5141.165	937.02	0.51
23	TR	2	R	5835.275	894.72	1.04
24	TR	1	T	4343.439	713.57	1.03
24	TR	2	R	2848.448	811.83	0.71
25	TR	1	T	3983.260	1160.32	0.73

25	TR	2	R	3476.389	769.63	0.78
27	TR	1	T	5772.972	1219.56	0.99
27	TR	2	R	7673.260	1063.29	1.03
28	RT	1	R	5679.039	650.24	1.00
28	RT	2	T	5160.875	891.63	1.05
29	TR	1	T	4800.455	770.63	2.02
29	TR	2	R	5772.925	738.17	1.04
30	RT	1	R	4722.324	1034.11	0.77
30	RT	2	T	2896.939	569.22	1.03
31	RT	1	R	8032.393	1043.82	1.98
31	RT	2	T	6076.359	1141.43	0.96
32	TR	1	T	4245.372	608.93	2.97
32	TR	2	R	4745.770	539.66	2.04
33	TR	1	T	3648.195	856.18	0.76
33	TR	2	R	3356.777	647.95	0.98
34	TR	1	T	5015.499	739.42	0.96
34	TR	2	R	6325.746	682.41	1.99
35	RT	1	R	6259.347	1020.55	1.96
35	RT	2	T	5802.468	835.87	2.04
36	RT	1	R	4669.384	682.87	3.01
36	RT	2	T	3783.584	729.63	1.00

B   SAS Scripts and results

To run these scripts, the dataset `NCAResult4BE` should be exported from R by `write.csv()`.

B.1   C$_{max}$

```
DATA BE; /* It will load 91 records. */
  INFILE 'c:\Users\mdlhs\asancpt\BEreport\sas\NCAResult4BE.csv' FIRSTOBS=2 DLM=",";
  INPUT SUBJ $ SEQ $ PRD $ TRT $ AUClast Cmax Tmax;
  IF CMAX =< 0 THEN DELETE;
  LNCMAX = LOG(Cmax);
  LNAUCL = LOG(AUClast );

PROC PRINT; RUN;

PROC GLM DATA=BE OUTSTAT=STATRES; /* GLM use only complete subjects. */
  CLASS SEQ PRD TRT SUBJ;
  MODEL LNCMAX = SEQ SUBJ(SEQ) PRD TRT;
  RANDOM SUBJ(SEQ)/TEST;
  LSMEANS TRT /PDIFF=CONTROL('R') CL ALPHA=0.1 COV OUT=LSOUT;
RUN;

PROC PRINT DATA=STATRES; RUN;
PROC PRINT DATA=LSOUT; RUN;

DATA STATRES;
  SET STATRES;
  IF _TYPE_='ERROR' THEN CALL SYMPUT('DF', DF);

DATA LSOUT;
  SET LSOUT;
  IF TRT='R' THEN CALL SYMPUT('GMR_R', LSMEAN);
  IF TRT='T' THEN CALL SYMPUT('GMR_T', LSMEAN);
  IF TRT='R' THEN CALL SYMPUT('V_R', COV1);
  IF TRT='T' THEN CALL SYMPUT('V_T', COV2);
  IF TRT='T' THEN CALL SYMPUT('COV', COV1);

DATA LSOUT2;
  LNPE = &GMR_T - &GMR_R;
  DF = &DF;
  SE = SQRT(&V_R + &V_T - 2*&COV);
  LNLM = TINV(0.95, DF)*SE;
  LNLL = LNPE - LNLM ;
  LNUL = LNPE + LNLM;
  PE = EXP(LNPE);
  LL = EXP(LNLL);
  UL = EXP(LNUL);
  WD = UL - LL;

PROC PRINT DATA=LSOUT2; RUN;

PROC MIXED DATA=BE; /* MIXED  uses all data. */
```

```
  CLASS SEQ TRT SUBJ PRD;
  MODEL LNCMAX = SEQ PRD TRT;
  RANDOM SUBJ(SEQ);
  ESTIMATE 'T VS R' TRT -1 1 /CL ALPHA=0.1;
  ODS OUTPUT ESTIMATES=ESTIM COVPARMS=COVPAR;
RUN;

DATA COVPAR;
  SET COVPAR;
  IF CovParm = 'Residual' THEN CALL SYMPUT('MSE', Estimate);

DATA ESTIM;
  SET ESTIM;
  MSE = &MSE;
  LNLM = (Upper - Lower)/2;
  PE = EXP(Estimate);
  LL = EXP(Lower);
  UL = EXP(Upper);
  WD = UL - LL;

PROC PRINT Data=ESTIM; RUN;
```

```
read_csv('sas/SAS_results_Cmax.csv') %>%
  knitr::kable(longtable = TRUE, booktabs = TRUE, caption = 'Cmax')
```

B.2 AUC$_{last}$

```
DATA BE; /* It will load 91 records. */
  INFILE 'c:\Users\mdlhs\asancpt\BEreport\sas\NCAResult4BE.csv' FIRSTOBS=2 DLM=",";
  INPUT SUBJ $ SEQ $ PRD $ TRT $ AUClast Cmax Tmax;
  IF CMAX =< 0 THEN DELETE;
  LNCMAX = LOG(Cmax);
  LNAUCL = LOG(AUClast );

PROC PRINT; RUN;

PROC GLM DATA=BE OUTSTAT=STATRES; /* GLM use only complete subjects. */
  CLASS SEQ PRD TRT SUBJ;
  MODEL LNAUCL = SEQ SUBJ(SEQ) PRD TRT;
  RANDOM SUBJ(SEQ)/TEST;
  LSMEANS TRT /PDIFF=CONTROL('R') CL ALPHA=0.1 COV OUT=LSOUT;
RUN;

PROC PRINT DATA=STATRES; RUN;
PROC PRINT DATA=LSOUT; RUN;

DATA STATRES;
  SET STATRES;
  IF _TYPE_='ERROR' THEN CALL SYMPUT('DF', DF);

DATA LSOUT;
  SET LSOUT;
```

```sas
      IF TRT='R' THEN CALL SYMPUT('GMR_R', LSMEAN);
      IF TRT='T' THEN CALL SYMPUT('GMR_T', LSMEAN);
      IF TRT='R' THEN CALL SYMPUT('V_R', COV1);
      IF TRT='T' THEN CALL SYMPUT('V_T', COV2);
      IF TRT='T' THEN CALL SYMPUT('COV', COV1);

DATA LSOUT2;
    LNPE = &GMR_T - &GMR_R;
    DF = &DF;
    SE = SQRT(&V_R + &V_T - 2*&COV);
    LNLM = TINV(0.95, DF)*SE;
    LNLL = LNPE - LNLM ;
    LNUL = LNPE + LNLM;
    PE = EXP(LNPE);
    LL = EXP(LNLL);
    UL = EXP(LNUL);
    WD = UL - LL;

PROC PRINT DATA=LSOUT2; RUN;

PROC MIXED DATA=BE; /* MIXED  uses all data. */
    CLASS SEQ TRT SUBJ PRD;
    MODEL LNAUCL = SEQ PRD TRT;
    RANDOM SUBJ(SEQ);
    ESTIMATE 'T VS R' TRT -1 1 /CL ALPHA=0.1;
    ODS OUTPUT ESTIMATES=ESTIM COVPARMS=COVPAR;
RUN;

DATA COVPAR;
    SET COVPAR;
    IF CovParm = 'Residual' THEN CALL SYMPUT('MSE', Estimate);

DATA ESTIM;
    SET ESTIM;
    MSE = &MSE;
    LNLM = (Upper - Lower)/2;
    PE = EXP(Estimate);
    LL = EXP(Lower);
    UL = EXP(Upper);
    WD = UL - LL;

PROC PRINT Data=ESTIM; RUN;
```

C Session Information

```
devtools::session_info()
```

```
## - Session info ---------------------------------------------------------
##   setting  value
##   version  R version 3.5.1 (2018-07-02)
##   os       macOS High Sierra 10.13.6
##   system   x86_64, darwin17.6.0
##   ui       unknown
##   language (EN)
##   collate  en_US.UTF-8
##   ctype    en_US.UTF-8
##   tz       Asia/Seoul
##   date     2018-10-23
##
## - Packages -------------------------------------------------------------
##   package      * version date       lib source
##   assertthat     0.2.0   2017-04-11 [1] CRAN (R 3.5.0)
##   backports      1.1.2   2017-12-13 [1] CRAN (R 3.5.0)
##   base64enc      0.1-3   2015-07-28 [1] CRAN (R 3.5.0)
##   BE           * 0.1.1   2018-07-19 [1] CRAN (R 3.5.1)
##   bindr          0.1.1   2018-03-13 [1] CRAN (R 3.5.0)
##   bindrcpp       0.2.2   2018-03-29 [1] CRAN (R 3.5.0)
##   bookdown       0.7     2018-02-18 [1] CRAN (R 3.5.0)
##   broom          0.5.0   2018-07-17 [1] CRAN (R 3.5.1)
##   callr          3.0.0   2018-08-24 [1] CRAN (R 3.5.1)
##   cellranger     1.1.0   2016-07-27 [1] CRAN (R 3.5.0)
##   cli            1.0.1   2018-09-25 [1] CRAN (R 3.5.1)
##   colorspace     1.3-2   2016-12-14 [1] CRAN (R 3.5.0)
##   crayon         1.3.4   2017-09-16 [1] CRAN (R 3.5.0)
##   debugme        1.1.0   2017-10-22 [1] CRAN (R 3.5.0)
##   desc           1.2.0   2018-05-01 [1] CRAN (R 3.5.0)
##   devtools       2.0.0   2018-10-19 [1] CRAN (R 3.5.1)
##   digest         0.6.18  2018-10-10 [1] CRAN (R 3.5.1)
##   dplyr        * 0.7.7   2018-10-16 [1] CRAN (R 3.5.1)
##   evaluate       0.12    2018-10-09 [1] CRAN (R 3.5.1)
##   forcats      * 0.3.0   2018-02-19 [1] CRAN (R 3.5.0)
##   fs             1.2.6   2018-08-23 [1] CRAN (R 3.5.1)
##   ggplot2      * 3.0.0   2018-07-03 [1] CRAN (R 3.5.1)
##   glue           1.3.0   2018-07-17 [1] CRAN (R 3.5.1)
##   gtable         0.2.0   2016-02-26 [1] CRAN (R 3.5.1)
##   haven          1.1.2   2018-06-27 [1] CRAN (R 3.5.0)
##   hms            0.4.2   2018-03-10 [1] CRAN (R 3.5.0)
##   htmltools      0.3.6   2017-04-28 [1] CRAN (R 3.5.0)
##   httr           1.3.1   2017-08-20 [1] CRAN (R 3.5.0)
##   jsonlite       1.5     2017-06-01 [1] CRAN (R 3.5.0)
##   knitr        * 1.20    2018-02-20 [1] CRAN (R 3.5.0)
##   lattice        0.20-35 2017-03-25 [2] CRAN (R 3.5.1)
##   lazyeval       0.2.1   2017-10-29 [1] CRAN (R 3.5.0)
##   lubridate      1.7.4   2018-04-11 [1] CRAN (R 3.5.0)
##   magrittr       1.5     2014-11-22 [1] CRAN (R 3.5.0)
##   memoise        1.1.0   2017-04-21 [1] CRAN (R 3.5.0)
```

```
##  modelr       0.1.2   2018-05-11 [1] CRAN (R 3.5.0)
##  munsell      0.5.0   2018-06-12 [1] CRAN (R 3.5.0)
##  nlme         3.1-137 2018-04-07 [2] CRAN (R 3.5.1)
##  pillar       1.3.0   2018-07-14 [1] CRAN (R 3.5.1)
##  pkgbuild     1.0.2   2018-10-16 [1] CRAN (R 3.5.1)
##  pkgconfig    2.0.2   2018-08-16 [1] CRAN (R 3.5.1)
##  pkgload      1.0.1   2018-10-11 [1] CRAN (R 3.5.1)
##  plyr         1.8.4   2016-06-08 [1] CRAN (R 3.5.0)
##  prettyunits  1.0.2   2015-07-13 [1] CRAN (R 3.5.0)
##  processx     3.2.0   2018-08-16 [1] CRAN (R 3.5.1)
##  ps           1.2.0   2018-10-16 [1] CRAN (R 3.5.1)
##  purrr      * 0.2.5   2018-05-29 [1] CRAN (R 3.5.0)
##  R.methodsS3  1.7.1   2016-02-16 [1] CRAN (R 3.5.0)
##  R.oo         1.22.0  2018-04-22 [1] CRAN (R 3.5.0)
##  R6           2.3.0   2018-10-04 [1] CRAN (R 3.5.1)
##  Rcpp         0.12.19 2018-10-01 [1] CRAN (R 3.5.1)
##  readr      * 1.1.1   2017-05-16 [1] CRAN (R 3.5.0)
##  readxl       1.1.0   2018-04-20 [1] CRAN (R 3.5.0)
##  remotes      2.0.1   2018-10-19 [1] CRAN (R 3.5.1)
##  rlang        0.3.0   2018-10-22 [1] CRAN (R 3.5.1)
##  rmarkdown    1.10    2018-06-11 [1] CRAN (R 3.5.0)
##  rprojroot    1.3-2   2018-01-03 [1] CRAN (R 3.5.0)
##  rstudioapi   0.8     2018-10-02 [1] CRAN (R 3.5.1)
##  rtf        * 0.4-13  2018-05-17 [1] CRAN (R 3.5.0)
##  rvest        0.3.2   2016-06-17 [1] CRAN (R 3.5.0)
##  scales       1.0.0   2018-08-09 [1] CRAN (R 3.5.1)
##  sessioninfo  1.1.0   2018-09-25 [1] CRAN (R 3.5.1)
##  stringi      1.2.4   2018-07-20 [1] CRAN (R 3.5.1)
##  stringr    * 1.3.1   2018-05-10 [1] CRAN (R 3.5.0)
##  testthat     2.0.1   2018-10-13 [1] CRAN (R 3.5.1)
##  tibble     * 1.4.2   2018-01-22 [1] CRAN (R 3.5.0)
##  tidyr      * 0.8.1   2018-05-18 [1] CRAN (R 3.5.0)
##  tidyselect   0.2.5   2018-10-11 [1] CRAN (R 3.5.1)
##  tidyverse  * 1.2.1   2017-11-14 [1] CRAN (R 3.5.0)
##  usethis      1.4.0   2018-08-14 [1] CRAN (R 3.5.1)
##  withr        2.1.2   2018-03-15 [1] CRAN (R 3.5.0)
##  xfun         0.3     2018-07-06 [1] CRAN (R 3.5.1)
##  xml2         1.2.0   2018-01-24 [1] CRAN (R 3.5.0)
##  yaml         2.2.0   2018-07-25 [1] CRAN (R 3.5.1)
##
## [1] /Users/Sungpil/Rlib
## [2] /usr/local/Cellar/r/3.5.1/lib/R/library
```

References

Bae, Kyun-Seop. 2018. *BE: Bioequivalence Study Data Analysis.* https://CRAN.R-project.org/package=BE.

Chow, Shein-Chung, and Jen-pei Liu. 2008. *Design and Analysis of Bioavailability and Bioequivalence Studies (Chapman & Hall/Crc Biostatistics Series).* Chapman; Hall/CRC.

Drug Evaluation, Center for, Food Research (CDER), U.S. Department of Health Drug Administration, and Human Services. 2001. *Guidance for Industry Statistical Approaches to Establishing Bioequivalence.*

https://www.fda.gov/downloads/drugs/guidances/ucm070244.pdf.

Hauschke, Dieter, Volker Steinijans, and Iris Pigeot. 2007. *Bioequivalence Studies in Drug Development: Methods and Applications.* Wiley.