PharmaSUG China 2019 - Paper CC-26

# Implement MCP-Mod using SAS and R

Xuan Yang, Boehringer Ingelheim

## ABSTRACT

MCP-Mod (Multiple Comparison Procedure - Modelling) is an efficient methodology for the design and analysis of phase II dose-finding trial. It starts with a set of potential models for the description of the dose-response relationship in the data, then all candidate models are tested if significantly different from a flat dose-response curve. The best of those models will be fitted to the data and target dose is estimated.

The paper focuses on implementing MCP-Mod using both R and SAS to generate flexible outputs based on study needs. Main reporting tasks are done by SAS; R statements will be submitted using PROC IML. The practical implementation of the MCP-Mod approach is accomplished by R functionalities provided in the "DoseFinding" package.

## INTRODUCTION

MCP-Mod (Multiple Comparison Procedure - Modelling) is a popular and efficient approach in phase II dose-response trials. It combines hypothesis testing and modeling in a structured manner with the purpose of finding the right dose(s). "DoseFinding" package for R provides tools for the analysis in dose-response studies. However, currently there is no specific procedure for MCP-Mod developed by SAS. Therefore, this paper will discuss

1) Main procedures and functions used to perform MCP-Mod analysis in R and SAS
2) How to call R within SAS to get analysis results
3) Reporting and result visualization

## DATA

| | STUDYID | USUBJID | SUBJID | TRTFL | PPROTFL | TRT01PNDC | TRT01PN | PARAM | PARAMN | PARAMCD | PARAMNDC | AVAL | ABLFL | BASE | CHG | ANL01FL | AVISITNDC | AVISITN |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | test-0001 | test-0001-101 | 101 | Y | Y | Bl 10mg | 50100 | Biomarker [ ] | 1 | BIOM | Biomarker | 133.9568 | Y | 133.9568 | 0 | Y | Baseline | 200 |
| 2 | test-0001 | test-0001-101 | 101 | Y | Y | Bl 10mg | 50100 | Biomarker [ ] | 1 | BIOM | Biomarker | 124.007 | | 133.9568 | -9.94975 | Y | Week 4 | 2901 |
| 3 | test-0001 | test-0001-101 | 101 | Y | Y | Bl 10mg | 50100 | Biomarker [ ] | 1 | BIOM | Biomarker | 93.33614 | | 133.9568 | -40.6206 | Y | Week 8 | 5701 |
| 4 | test-0001 | test-0001-101 | 101 | Y | Y | Bl 10mg | 50100 | Biomarker [ ] | 1 | BIOM | Biomarker | 120.3118 | | 133.9568 | -13.645 | Y | Week 12 | 8501 |
| 5 | test-0001 | test-0001-102 | 102 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 104.9758 | Y | 104.9758 | 0 | Y | Baseline | 200 |
| 6 | test-0001 | test-0001-102 | 102 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 112.7904 | | 104.9758 | 7.814523 | Y | Week 4 | 2901 |
| 7 | test-0001 | test-0001-102 | 102 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 78.22367 | | 104.9758 | -26.7522 | Y | Week 8 | 5701 |
| 8 | test-0001 | test-0001-102 | 102 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 64.75045 | | 104.9758 | -40.2254 | Y | Week 12 | 8501 |
| 9 | test-0001 | test-0001-103 | 103 | Y | Y | Bl 15mg | 50150 | Biomarker [ ] | 1 | BIOM | Biomarker | 123.9894 | Y | 123.9894 | 0 | Y | Baseline | 200 |
| 10 | test-0001 | test-0001-103 | 103 | Y | Y | Bl 15mg | 50150 | Biomarker [ ] | 1 | BIOM | Biomarker | 113.9195 | | 123.9894 | -10.0699 | Y | Week 4 | 2901 |
| 11 | test-0001 | test-0001-103 | 103 | Y | Y | Bl 15mg | 50150 | Biomarker [ ] | 1 | BIOM | Biomarker | 144.3519 | | 123.9894 | 20.36243 | Y | Week 8 | 5701 |
| 12 | test-0001 | test-0001-103 | 103 | Y | Y | Bl 15mg | 50150 | Biomarker [ ] | 1 | BIOM | Biomarker | 117.1068 | | 123.9894 | -6.8826 | Y | Week 12 | 8501 |
| 13 | test-0001 | test-0001-104 | 104 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 120.9763 | Y | 120.9763 | 0 | Y | Baseline | 200 |
| 14 | test-0001 | test-0001-104 | 104 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 107.7634 | | 120.9763 | -13.213 | Y | Week 4 | 2901 |
| 15 | test-0001 | test-0001-104 | 104 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 92.30972 | | 120.9763 | -28.6666 | Y | Week 8 | 5701 |
| 16 | test-0001 | test-0001-104 | 104 | Y | Y | Placebo | 2000 | Biomarker [ ] | 1 | BIOM | Biomarker | 103.9194 | | 120.9763 | -17.0569 | Y | Week 12 | 8501 |

This is a 12-week, placebo-controlled, parallel-group mock study used for demonstration purpose. In this data, 84 subjects are randomized into 5 treatment groups (placebo, 10mg, 15mg, 20mg, 30mg, each subject receives one active treatment or placebo), with visits at week 0 (baseline), 4, 8, and 12 to get test results of certain target biomarker (PARAMCD = "BIOM").

## MAIN PROCEDURES AND FUNCTIONS

### BEFORE MCPMOD

Before MCPMod approach in R, a model has to be fitted to data with ANOVA-type parametrization to the data to obtain estimates and covariance matrix. In this example repeated measurement analysis (MMRM) is used in SAS with PROC MIXED, point estimates and covariance matrix could be found in "LSMEANS=" statement, which then will be passed to R to perform MCPMod analysis:

```
ODS OUTPUT LSMEANS=_lsmeans Diffs=_diffs tests3=_tests solutionf=_solutf;
proc mixed data=_inds method=reml CL order=internal;
        class trt01pn(REF= LAST) avisitn usubjid;
        model chg = avisitn trt01pn base base*avisitn avisitn*trt01pn / solution residual ddfm=kr CL HTYPE=3 ;
        repeated avisitn/type=un subject=usubjid r;
```

```
        lsmeans  trt01pn*avisitn/CL alpha=0.05 COV PDIFF=all;
run;
```

This step can also be done in R and not necessarily in SAS. How to export data from SAS to R and vice versa see next section cALL R WITHIN SAS.

## TRIAL DESIGN STAGE

At the study planning stage, or to start with MCPMod analysis, a set of candidate models need to be pre-specified about the dose-relationship. In the “DoseFinding” package, possible models are "linlog", "linear", "quadratic", "emax", "exponential", "sigEmax", "betaMod" and "logistic". The function “guesst” will pass the parameters of each model to function “Mods”, which will then produce a list that contains all information about candidate models. “plot” function can plot all candidate models as shown in Figure 1.

```
### ---> 1.TRIAL DESIGN STAGE
### Pre-specified information for defining candidate models
### use guesst function: needs pairs (d, p) as input parameters
pboeff <- -12
acteff <- -15
quad   <- guesst(d = 14, p = 0.9, "quadratic")
exp    <- guesst(d = 14, p = 0.9, "exponential", Maxd = 3)
emax   <- guesst(d = 14, p = 0.9, "emax")
sigemax<- guesst(d = c(10 , 16), p = c(0.3, 0.9), "sigEmax")
logis  <- guesst(d = c(10 , 16), p = c(0.3, 0.9), model="logistic")

### Mods(): define the dose-response models
my.model <- Mods(linear      = NULL,
                 linlog      = NULL,
                 quadratic   = quad,
                 exponential = exp,
                 emax        = emax,
                 sigEmax     = sigemax,
                 logistic    = logis,
                 doses       = dose,
                 placEff     = pboeff,
                 maxEff      = acteff)

### use plot function to get defined MCPMod shapes
plot(my.model)
```

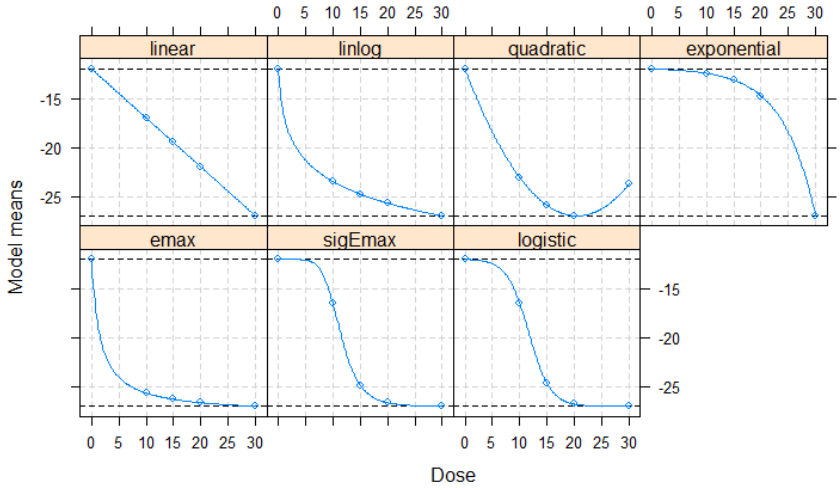



**Figure 1 Pre-specified candidate models**

## MCT STEP

Once the set of candidate models has been pre-specified, Mod step fits the dose response models and estimates target doses based on the model defined in the MCP step.

### 1) Optimal contrasts

Firstly, the optimal contrast coefficients is used to identify the best matching model, which can be computed with “optContr” function. It requires input of matrix, here in the example it uses the covariance matrix exported from SAS, using repeated measurement analysis (MMRM) by PROC MIXED.

```
### optContr(): calculation of optimal contrasts corresponding to the candidate models
contMatnew <- optContr(my.model, S=cov)
```

### 2) Multiple contrast test

Model-specific contrast tests with null hypotheses of a flat dose-response curve is performed by “MCTtest” function.

```
### MCTtest(): Use MCTtest to test the  null hypothesis, if a non flat curve exists
##If at least one dose-response model is statistically significant, rejecting the null hypothesis of
##a flat dose-response curve is indicating a benefit of the drug over placebo.
set.seed(2019, "L'Ecuyer") #mvtnorm() includes random process. Get stable results for p-value and critical value
my.test <- MCTtest( dose=dose, resp=estimate, data = est, my.model, S = cov,
                    type = "general",
                    placAdj = FALSE, df = Inf,
                    critV = TRUE, pval = TRUE, alpha = 0.025,
                    alternative = c("one.sided"), na.action = na.fail,
                    mvtcontrol = mvtnorm.control(), contMat = contMatnew)
```

Below shows the output of MCT in R:

```
Multiple Contrast Test

Contrasts:
   linear linlog quadratic exponential   emax sigEmax logistic
0   0.701  0.855     0.885       0.422  0.877   0.769    0.768
10  0.115 -0.080    -0.098       0.192 -0.140   0.188    0.189
15 -0.010 -0.141    -0.241       0.153 -0.171  -0.181   -0.173
20 -0.111 -0.172    -0.287       0.101 -0.178  -0.247   -0.254
30 -0.695 -0.462    -0.259      -0.867 -0.388  -0.529   -0.530

Contrast Correlation:
            linear linlog quadratic exponential  emax sigEmax logistic
linear       1.000  0.905     0.778       0.877 0.849   0.937    0.938
linlog       0.905  1.000     0.960       0.616 0.993   0.923    0.922
quadratic    0.778  0.960     1.000       0.388 0.976   0.885    0.884
exponential  0.877  0.616     0.388       1.000 0.527   0.684    0.685
emax         0.849  0.993     0.976       0.527 1.000   0.884    0.883
sigEmax      0.937  0.923     0.885       0.684 0.884   1.000    1.000
logistic     0.938  0.922     0.884       0.685 0.883   1.000    1.000

Multiple Contrast Test:
            t-Stat   adj-p
linear       2.705 0.00934
logistic     2.558 0.01393
sigEmax      2.543 0.01442
linlog       2.387 0.02106
exponential  2.316 0.02488
emax         2.215 0.03259
quadratic    2.085 0.04283

Critical value: 2.322 (alpha = 0.025, one-sided)
```

## MOD STEP

### 1) Fit the dose-response curve and estimate target dose

Models for each shape could be fitted by “fitMod” function in R:

```
## model parameter estimation needs to be done for each model separately
fit.linear <- fitMod(dose=dose, resp=estimate, data = est,  S = cov, model='linear', type="general")
fit.linlog <- fitMod(dose=dose, resp=estimate, data = est,  S = cov, model='linlog', type="general")
```

“plot” function can visualize the fitting in R:

```
plot(fit.linear,CI=TRUE,plotData='meansCI',level=0.95)
```

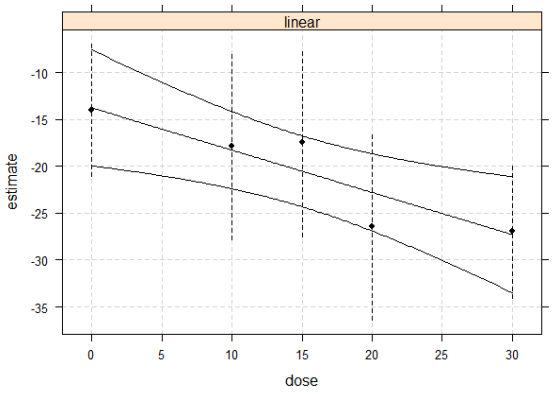



**Figure 2 Fitted model with "plot" function**

In addition, after the MCT if any model is significantly different from a flat dose-response curve, the best fit model could be selected and target dose could be calculated. The “MCPMod” function simultaneously conducts both steps. It includes the multiple contrast test as performed by the “MCTtest” function and fits all models with positive test decisions to the data. Furthermore, the TD/ED is estimated for all significant models by defining “Delta”, or “p”. The example requires that the target dose should be where the placebo-corrected biomarker change from baseline is 10 units, therefore “Delta = 10”. In the “selModel” option, a criterion for the model selection can be selected:

- model with the minimum p-value/maximum test statistic (“maxT”), or
- model with smallest AIC (“AIC”), or
- weighted average over significant models (“aveAIC”)

```
### MCPMod(): tests for a dose-response effect
set.seed(2019, "L'Ecuyer") #Need to set seed again to keep same results with MCTtest()
resultAIC <- MCPMod(dose=dose, resp=estimate, data = est, my.model, S=cov,
                    type="general", critV=T, alpha=0.025, alternative = c("one.sided"), Delta=10, selModel = c("AIC"))
```

Output:

```
> resultAIC$selMod
[1] "linear"
> resultAIC$doseEst
     linear      linlog exponential     sigEmax    logistic
   22.17510    18.91036    23.58260    19.57068    18.90719
attr(,"addPar")
[1] 10
```

## 2) Get predictions

Function “predict” will compute model averaged predictions. To calculate confidence interval, “se.fit” argument is marked as “TRUE” to get standard errors.

```
## Estimate ls-means + SE for a grid of doses
doseVec<-seq(0,30,by=0.1)
predlinear <- predict(fit.linear , se.fit=TRUE, doseSeq=doseVec, predType="ls-means")
```

Although R produce very nice figures (see Figure 2 Fitted model with "plot" function), sometimes further modification is needed and in this example the predictions are composed into a dataframe then passed to SAS for GTL plotting.

```
pred.linear <- data.frame(cbind(doseVec, predlinear$fit, predlinear$se.fit))
names(pred.linear) <- c("dose","predict", "predict_se")
```

Dataframe pred.linear to pass to SAS:

| dose | predict | predict_se |
|---|---|---|
| 0.0 | -13.73911 | 3.160127 |
| 0.1 | -13.78420 | 3.146956 |
| 0.2 | -13.82930 | 3.133817 |
| 0.3 | -13.87439 | 3.120713 |
| 0.4 | -13.91949 | 3.107642 |
| 0.5 | -13.96458 | 3.094606 |
| 0.6 | -14.00968 | 3.081605 |
| 0.7 | -14.05478 | 3.068640 |
| 0.8 | -14.09987 | 3.055710 |
| 0.9 | -14.14497 | 3.042817 |
| 1.0 | -14.19006 | 3.029961 |

## CALL R WITHIN SAS

Statements „ExportDataSetToR" and „ImportDataSetFromR" in PROC IML will pass data between SAS and R. Both calls convert variable format types automatically so that each program can use them. Note that within SUBMIT/R and ENDSUBMIT statements should be pure R code. Specifically, if PROC IML is within a MACRO, "SUBMIT/R", "ENDSUBMIT", and R code should be saved into a file first and then use "%include" to read the code.

In the example, SAS datasets _est and _cov (point estimates and covariance matrix from PROC MIXED) is passed to R named as "est" and "cov". After analysis in R, results (test statistics, contrast, critical value, target dose, predictions, and model coefficients) are transferred back to SAS, ready for restructure and reporting.

```
PROC IML;
 *EXPORT DATASET TO R;
 CALL EXPORTDATASETTOR("work._est", "est");
 CALL EXPORTDATASETTOR("work._cov", "cov");
 SUBMIT / R;
       %INCLUDE "c:\demo\mcpmod-rcode.sas";
 ENDSUBMIT;
 *IMPORT DATAFRAME FROM R;
 CALL IMPORTDATASETFROMR("work.mcpm_contrast","contrast");
 CALL IMPORTDATASETFROMR("work.mcpm_ctest",   "ctest");
 CALL IMPORTDATASETFROMR("work.mcpm_critval", "critval" );
 CALL IMPORTDATASETFROMR("work.mcpm_target",  "predDose");
 CALL IMPORTDATASETFROMR("work.pred_linear",  "pred.linear"  );
 CALL IMPORTDATASETFROMR("work.coeff_linear", "coeff.linear" );
QUIT;
```

## REPORTING AND RESULT VISULIZATION

### 1) TABLES

SAS can organize the data in a more structured way.

| | estimates | linear* | logistic* | sigEmax* | linlog* | exponential | emax | quadratic |
|---|---|---|---|---|---|---|---|---|
| MMRM estimates: | | | | | | | | |
| Placebo | -14.02 | | | | | | | |
| Bl 10mg | -17.80 | | | | | | | |
| Bl 15mg | -17.44 | | | | | | | |
| Bl 20mg | -26.37 | | | | | | | |
| Bl 30mg | -26.94 | | | | | | | |
| Contrast: | | | | | | | | |
| Placebo | | 0.7005 | 0.7677 | 0.7690 | 0.8549 | 0.4221 | 0.8768 | 0.8849 |
| Bl 10mg | | 0.1154 | 0.1889 | 0.1880 | -0.0797 | 0.1918 | -0.1399 | -0.0985 |
| Bl 15mg | | -0.0096 | -0.1733 | -0.1812 | -0.1410 | 0.1525 | -0.1706 | -0.2407 |
| Bl 20mg | | -0.1110 | -0.2538 | -0.2473 | -0.1723 | 0.1006 | -0.1779 | -0.2868 |
| Bl 30mg | | -0.6954 | -0.5296 | -0.5285 | -0.4618 | -0.8670 | -0.3884 | -0.2589 |
| Multiple Contrast Test: | | | | | | | | |
| t-Stat | | 2.7054 | 2.5584 | 2.5432 | 2.3866 | 2.3160 | 2.2147 | 2.0847 |
| adj. p-value | | 0.0094 | 0.0133 | 0.0143 | 0.0210 | 0.0248 | 0.0321 | 0.0437 |
| Critical value: 2.318 (alpha = 0.025, one-sided) | | | | | | | | |

**Example table 1 Summary of multiple contrast test from MCPMod and MMRM estimates ("*" means**

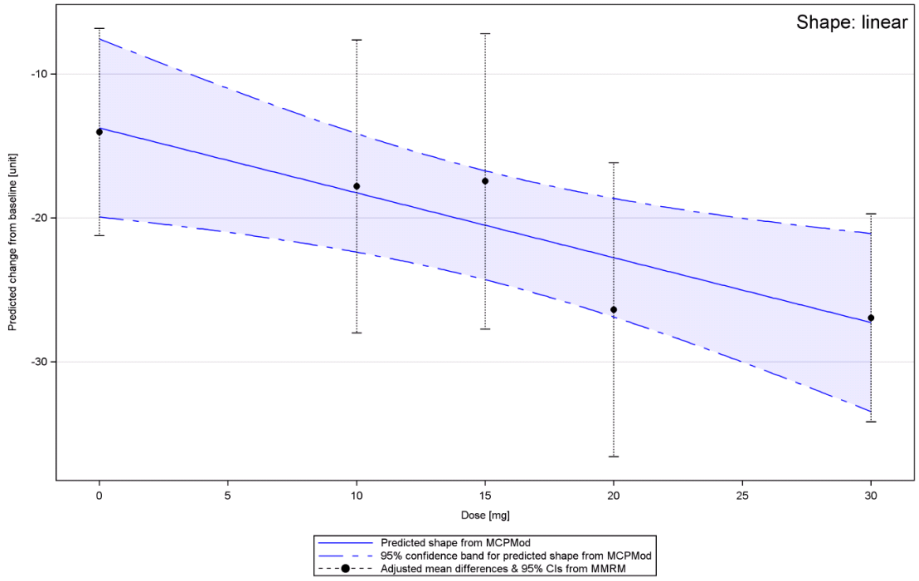
**significant models)**

| Model | E0 | E1 | Emax | Delta | B1 | B2 | ED50 | H |
|---|---|---|---|---|---|---|---|---|
| linear | -13.74 | . | . | -0.45 | . | . | . | . |
| linlog | -16.22 | . | . | -2.40 | . | . | . | . |
| quadratic | -13.79 | . | . | . | -0.4337 | -.000576 | . | . |
| exponential | -14.10 | -20.77 | . | 60.00 | . | . | . | . |
| emax | -13.24 | . | -33.01 | . | . | . | 45.00 | . |
| sigEmax | -15.13 | . | -12.14 | . | . | . | 16.77 | 10.00 |
| logistic | -15.22 | . | -11.79 | 1.26 | . | . | 16.74 | . |

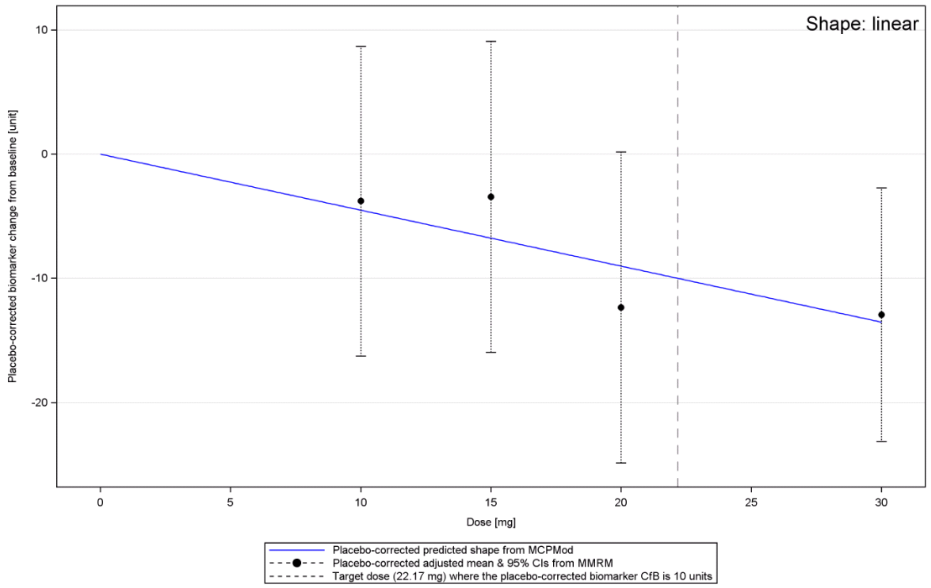
**Example table 2 Summary of model coefficients**

## 2) FIGURES

Below Example figure 1 Dose response of biomarker change from baseline can also be done by R with “plot” function as shown in Figure 2 Fitted model with "plot" function




**Example figure 1 Dose response of biomarker change from baseline**




**Example figure 2 Placebo-corrected of biomarker change from baseline over dose, marked with target dose**

## 3) CONSOLE INFORMATION FROM R

PROC IML can also pass the output in R console into SAS output window automatically, which can be a good tracking source to check the analysis process and object details in R. To organize the console output, blow R code can be adapted. Equal and arrow signs could properly separate each output topic and make SAS output more reader-friendly:

```
##########################################
## PART 2: Output statistics, for log in SAS##
##########################################
# 1. Output read-in dose, estimates, and matrix
cat(" \n===> Input estimates and covariance matrix\n  ")
est
cat(" ")
cov

# 2. Output MCTtest results
cat(" \n===> MCT test\n  ")
my.test

# 3. Output fit statistics for each model
cat(" \n===> Model fitting\n  ")
summary(fit.linear)
cat("\nGeneralized residual sum of squares:",fit.linear$gRSS)
cat("\n=============================================================\n  ")
```

Then SAS output will be like:

```
===> Input estimates and covariance matrix
     estimate dose
1 -14.01663    0
2 -17.79892   10
3 -17.44470   15
4 -26.37129   20
5 -26.93801   30
             COL1        COL2        COL3        COL4        COL5
[1,] 13.08529255  0.01569928  0.08109754 -0.05291055 -0.03503621
[2,]  0.01569928 26.16322323  0.09724026 -0.06344256 -0.04201027
[3,]  0.08109754  0.09724026 26.64671147 -0.32772426 -0.21701180
[4,] -0.05291055 -0.06344256 -0.32772426 26.35821643  0.14158523
[5,] -0.03503621 -0.04201027 -0.21701180  0.14158523 13.16595411

===> MCT test
  Multiple Contrast Test

Contrasts:
   linear linlog quadratic exponential   emax sigEmax logistic
0   0.701  0.855     0.885       0.422  0.877   0.769    0.768
10  0.115 -0.080    -0.098       0.192 -0.140   0.188    0.189
15 -0.010 -0.141    -0.241       0.153 -0.171  -0.181   -0.173
20 -0.111 -0.172    -0.287       0.101 -0.178  -0.247   -0.254
30 -0.695 -0.462    -0.259      -0.867 -0.388  -0.529   -0.530

Contrast Correlation:
            linear linlog quadratic exponential  emax sigEmax logistic
linear       1.000  0.905     0.778       0.877 0.849   0.937    0.938
linlog       0.905  1.000     0.960       0.616 0.993   0.923    0.922
quadratic    0.778  0.960     1.000       0.388 0.976   0.885    0.884
exponential  0.877  0.616     0.388       1.000 0.527   0.684    0.685
emax         0.849  0.993     0.976       0.527 1.000   0.884    0.883
sigEmax      0.937  0.923     0.885       0.684 0.884   1.000    1.000
logistic     0.938  0.922     0.884       0.685 0.883   1.000    1.000

Multiple Contrast Test:
            t-Stat   adj-p
linear       2.705 0.00939
logistic     2.558 0.01332
sigEmax      2.543 0.01434
linlog       2.387 0.02099
exponential  2.316 0.02476
emax         2.215 0.03211
quadratic    2.085 0.04368

Critical value: 2.318 (alpha = 0.025, one-sided)
```

## CONCLUSION

The paper shows an example of implementing MCP-Mod combine SAS and R to generate composed and structured outputs. First, MMRM is done by SAS and pass estimates and covariance matrix to R with PROC IML, then R performs MCPMod analysis using “DoseFinding” package. Important statistics are passed back to SAS again for reporting. With the combination use of SAS and R, results are organized in a tidy and elegant way.

## REFERENCES


Bjoern Bornkamp, Jose Pinheiro and Frank Bretz (2018). DoseFinding: Planning and Analyzing Dose Finding Experiments. R package version 0.9-16. https://CRAN.R-project.org/package=DoseFinding

Pinheiro, J., Bornkamp, B., Glimm, E. and Bretz, F. (2013). Model-based dose finding under model uncertainty using general parametric models. Statistics in Medicine, 33(10), pp.1646-1661.


## ACKNOWLEDGMENTS


I would like to thank my colleagues Jasmin Link, Barbara Doll, Gabriele Biegert, and Qiqi Deng in providing statistical and programming support and expertise.


## CONTACT INFORMATION

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


Name: Xuan Yang
E-mail: xuan.yang@boehringer-ingelheim.com