TCP Transl Clin Pharmacol
2014;22(2):78-82
http://dx.doi.org/10.12793/tcp.2014.22.2.78

On comparison of SAS codes with GLM and MIXED for the crossover studies with QT interval data

Kyungmee Choi[1]*, Taegon Hong[2,3] and Jongtae Lee[2,3]

[1]Divison of Mathematics, College of Science and Technology, Hongik University at Sejong, Sejong 339-701, Korea, [2]Department of Clinical Pharmacology and Therapeutics, Seoul St. Mary's Hospital, The Catholic University of Korea, Seoul 137-701, Korea, [3]PIPET (Pharmacometrics Institute for Practical Education & Training), Seoul 137-701, Korea

*Correspondence: K. M. Choi; Tel: +82-44-860-2237, Fax: +82-44-860-2648, E-mail: kmchoi@hongik.ac.kr

ORIGINAL ARTICLE

Received 6 Nov 2014
Revised 27 Nov 2014
Accepted 29 Nov 2014

Keywords
QT interval,
PROC GLM,
PROC MIXED,
Crossover studies,
repeated measures

pISSN: 2289-0882
eISSN: 2383-5427

The structural complexity of crossover studies for bioequivalence test confuses analysts and leaves them a hard choice among various programs. Our study reviews PROC GLM and PROC MIXED in SAS and compares widely used SAS codes for crossover studies. PROC MIXED based on REML is more recommended since it provides best linear unbiased estimator of the random between-subject effects and its variance. Our study also considers the covariance structure within subject over period which most PK/PD studies and crossover studies ignore. The QT interval data after the administration of moxifloxacin for a fixed time point are analyzed for the comparison of representative SAS codes for crossover studies.

Introduction

The confusion about the crossover designs for bioequivalence test[1-4] starts from understanding the complex treatment combination with confounded alias effects and continues further to choosing a program for analysis. Above all, since each subject is randomly sampled from a population and becomes a blocking factor, crossover designs are mixed-effects models. They are not full factorial designs and each main effect is confounded with a two-way interaction.[5] On the other hand, the crossover design has within-subjects effects since there are measurements over all periods for each subject. Repeated measurements in a subject also cause correlations within subjects. Due to the mixture of these complex statistical concepts, analysts confront a hard choice among various programs.

This paper is to review and compare widely used SAS codes for crossover studies, and will apply them to QT interval data collected after the administration of moxifloxacin. First of all, we review the experimental design terminologies related to crossover studies. Second part of the paper compares five SAS codes for crossover studies and selects one code. For the comparison, our study will consider the covariance structure within subject over period which most PK/PD studies and crossover studies ignore. In the final part, we will fit models in SAS code with the QT interval data for a fixed time point and make a final choice.

This paper reviews two SAS procedures to deal with the complexity of crossover studies: GLM and MIXED. PROC GLM uses least squares or method of moments to fit general linear models.[6] On the other hand, PROC MIXED uses Restricted (or residual) Maximum Likelihood (REML).[7] PROC MIXED is recommended to avoid pitfalls of PROC GLM.[5] This paper uses PROC GLM only to verify the expected mean squares in the analysis of variance and PROC MIXED to test and estimate the effects in the crossover design model.

Methods

Data description

We collected QT interval data from 33 healthy male Korean subjects in order to compare the equality of three treatments: A, B, and C (A: placebo, 240 ml water only; B: moxifloxacin 400 mg; C: moxifloxacin 800 mg). A William's square design, one of the crossover studies, with three periods was used to compare the three treatments by controlling carryover effect. There are 6 sequences and each sequence consists of three periods such as ABC, ACB, BAC, BCA, CAB, and CBA. The subjects were

randomly assigned to one of the 6 sequences. Electrocardiograms (ECGs) were measured just before and 1,2,3,4,6,8,12,16, 24 hours after administration of placebo or moxifloxacin. ECGs were taken for two consecutive days: the first day to record diurnal change (baseline data) and the next day to evaluate placebo or drug effect. Fridericia's formula was used to adjust QT interval (QTcF) for heart rate because increased heart rate usually shortens the QT interval. Change from baseline in QTcF (ΔQTcF) at each time point was analyzed for the comparison of three treatments using SAS® version 9.2 (SAS Institute Inc, Cary, USA). Each subject's baseline value (QTcFb) was also used as a covariate.

Experimental design effects

Confounding effects

Cheng et al.[5] reviewed and analyzed the confounding structure in crossover studies as follows. Each main effect is confounded with a two-factor interaction.[5] Since a combination of SEQUENCE and PERIOD subsequently determines a TREATMENT, the main effect of TREATMENT cannot be separated from the two-factor interaction of SEQUENCE*PERIOD. The main effect, its alias two-factor interaction, or both can cause the same significant result. Random assignment of subjects to SEQUENCEs, however, might allow analysts to ignore SEQUENCE effect, SEQUENCE*TREATMENT interaction, and SEQUENCE*PERIOD interaction.[5] A significant SEQUENCE effect is, therefore, likely to result from a PERIOD*TREATMENT interaction not from a difference between SEQUENCEs.[5] Similarly, a significant PERIOD effect can be interpreted as its own effect instead of a SEQUENCE*TREATMENT interaction, and a significant TREATMENT effect can be interpreted as its own effect instead of a SEQUENCE*PERIOD interaction.[5]

Fixed effect vs. Random effect

There are two types of effects in the experimental design: fixed and random. If an experimenter chooses k treatments and wishes to test the hypothesis about the treatment means, then this variable is a fixed effect.[8,9] In this case, the conclusion is applied only to the chosen k factor levels.[8] Alternatively, if the k treatments are randomly sampled from the population and the conclusion is extended to the population, then this variable is a random effect.[8,9] For example, since knowledge about particular subjects randomly selected from the population is not of our interest[8] and the subjects can be replaced by other subjects from the population, subject is a random effect. Since the analysts in the crossover studies expect subjects to affect only the variation but not the mean, the subject is a random effect while sequence, period, and treatment are fixed effects.

Between-subject effect vs. Within-subject effect

Between-subject effects are those whose levels remain constant within subject, whereas within-subject effects change those whose levels change within subject.[7] That is, "the levels of the between-subject variables include different subjects and the levels of the within-subject variables contain the same subject".[10] In the crossover design with QT interval data, gender, sequence, and treatment are fixed between-subject effects and subject is a random between-subject effect while period and time are fixed within-subject effects. Since subjects are measured more than once over time, a within-subject variable is also called as a repeated measured variable and the values could be correlated. [10] In NONMEM, between-subject variability is often called inter-individual variability.[11]

SAS codes for crossover studies

PROC GLM with random

As reviewed by Brunelle,[5,12] in PROC GLM is as in program1

```
/*Program 1 [12]*/
PROC GLM data=QT;
CLASS seq subject period trt ;
MODEL QTcF = QTcFb seq subject(seq) period trt;
RANDOM subject(seq) /test;
TEST H=seq E=subject(seq) ;
LSMEANS seq trt;
run;
```

CLASS specifies all factors in the model and MODEL equates the response variable QTcF to a linear combination of all fixed and random effects. RANDOM specifies random effects and all random interactions, and 'test' option returns the theoretical expected mean squares helping the analysts determine appropriate error terms in hypothesis tests on each effect of the model. TEST performs actual hypothesis tests for each effect in the model with appropriate error terms specified by H and E.[13] LSMEANS calculates the ML means of factors of interest, which is actually the within-group means adjusted for the other factors and especially important for unbalanced data.[13] The pdiff option performs the multiple comparisons based on one of the methods such as Bonferroni, Duncan, LSD, TUKEY, and others to check in which pair of groups the actual difference occur. The default error mean square is the overall residual or MSE.[13]

PROC MIXED with Random

There are several ways of expressing the crossover design model. A common linear model which has been used for the crossover design is as follows:[2]

$$y_{ijk} = \mu + \gamma_i + \xi_{i(k)} + \pi_j + \tau_{d(i,j)} + \epsilon_{ijk}$$

where, for γ_i is the sequence effect, π_i is the period effect, and τ_i is the treatment effect. The 103 subscript d(i,j) is either A or B or C determined by sequence i and period j.

Since sequence, period, and treatment are fixedeffects and subject is a random effect, a mixed 105 effect model is appropriate to fit crossover studies using three different PROC MIXED as in programs106 2,[12] 3,[5] and 4[2];

```
/*Program 2 [12]*/
PROC MIXED data=QT;
CLASS seq subject period trt;
MODEL QTcF = QTcFb seq period trt ;
RANDOM subject(seq) / subject=subject type=cs;
run ;
```

```
/*Program 3 [5]*/
PROC MIXED data= QT;
CLASS seq subject period trt;
MODEL QTcF = QTcFb seq period trt / solution;
RANDOM subject(seq)/VCorr;
LSMEANS trt/alpha=0.1 cl diff adjust=tukey;
run;
```

```
/*Program 4 [2]*/
PROC MIXED data= QT ;
CLASS seq subject period trt;
MODEL QTcF = QTcFb seq period trt/ ddfm= kenwardrodger;
RANDOM subject(seq) / subject=subject;
LSMEANS trt/pdiff cl alpha=0.1;
run;
```

It primarily uses Restricted (or residual) Maximum Likelihood (REML) whilePROC GLM uses method of moments estimators.[6,7] REML in PROC MIXED estimates variance components based on the residuals which are free from fixed effects.[14,15] PROC MIXED provides the best linear unbiased predictors (BLUPs) of the random effects based on REML.[16] On the other hand, the MLE underestimates thevariance parameters. As an easy example, MLE of σ^2, $\sigma^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$, is well known to underestimate the variance σ^2.

PROC MIXED canbetter dealwith missing repeated measures at random, while PROC GLM ignores data with missing repeated measures.[13] The SUBJECT=option enables PROC MIXED to process the model by subjects, which typically takes less time and memory.[17]

Repeated Measures Analysis

Since the QT interval data for a fixed time point have the design structure of crossover studies, 3 measurements on a subject are repeated over time. As measurements are made of the same characteristic on the same observational unit but on more than one occasion or the data structure is considered as one with profiles, REPEATED option is also appropriate.[5,18] Since measurements on the same experimental unit, subject, could be possibly correlated, the model assumption of independent errors may be violated.[5,18]

The covariance structures in repeated measures analysis can be

$$y_i = X_i\beta + Z_i\eta_i + \varepsilon_i,\ i = 1,2,\cdots,N,$$

$$\eta_i \sim N(0,\Omega_i),\ \varepsilon_i \sim N(0,\Sigma_i),$$

defined in the matrix form as in Pinheiro and Bates:[16]
where there areN subjects with n_i repeated responsesin eachith subject. β is the p-dimensional vector of fixed effects while η_i is the q-dimensional vector of random effects and Ω_i to be the correlation between random effects. X_i (n_i x p) is the fixed-effect explanatory regressor matrix while Z_i (n_i x q) is the between-subjects random effect regressor matrix. ε_i is the n_i–dimensional

$$y_1, y_2, \cdots, y_N \sim N(X_i\beta, V_i),$$

errors η_i and ε_i are assumed to be independent. Then,
where $V_i = Z_i\Omega_i Z_i^T + \Sigma_i$, the variance-covariance structure of the ith subject. In general, Σ_i and Ω_i are assumed to be the same for all subjects. Typically Σ_i varies among unstructured, ARMA(1,1), equally correlated, uncorrelated, or compound symmetry, while the software package NONMEM[19] for population pharmacokinetic modeling assumes that $\Sigma_i = \sigma^2 I$.

The PRINTE statement in PROC GLM carries out Mauchly [20] and Anderson's sphericity tests.[21] The Anderson's test checks whether or not the within-subjects variance-covariance matrix has a Type H covariance structure.[22] If the sphericity test is not significant, then the matrix is of Type H and the standard univariate test for the within-subjects effects is appropriate. Otherwise, the multivariate MANOVA tests like Wilks' Lambda [10] are necessary for the within-subjects effects. The assumption of sphericity is not required to test the between-subjects effect. The sphericity test, however, is not used in this paper since the within-subject correlation structure is of our interest.

In PROC MIXED as in program 5 [17] all fixed effects, both between- and within-subjects, must be included in the MODEL statement.

```
/*Program 5 [17]*/
PROC MIXED data = QT;
CLASS seq subject period trt;
MODEL QTcF = QTcFb seq period trt;
REPEATED / subject=subject type=cs R;
RUN;
```

The subject=subject option makes the entire variance-covariance matrix to be N block diagonals of $n_i \times n_i$ blocks.There

are various options for the $n_i \times n_i$ variance-covariance matrix: unstructured, AR, ARMA, equi-correlation, uncorrelated, compound symmetry, hetero compound symmetry, and more. The compound symmetry is often recommended because of its simplicity[17]since it has less number of parameters to estimate.The variance-covariance with compound symmetry is as follows:

$$\begin{pmatrix} \sigma_e^2+\sigma_b^2 & \sigma_b^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_e^2+\sigma_b^2 & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_e^2+\sigma_b^2 \end{pmatrix}.$$

The unstructured variance-covariance is as follow:

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}.$$

For comparison of models with different covariance structures, the likelihood-based information criteria of Akaike (AIC)[23] and Schwarz (BIC)[24] criteria can be used, which are as follows:

$$\text{AIC} = -2 \log \text{Likelihood} + 2p,$$
$$\text{BIC} = -2 \log \text{Likelihood} + p \log(N).$$

Here p is the number of parameters in the model and N is the number of observations. We select the model with the smallest BIC or AIC.

Results

Four programs 1, 2, 3, and 5 with both type=un and type=cs were run, ignoring program 4[10] since programs 2 and 4 are almost the same. In order to analyze the crossover designs, PROC MIXED is often recommended to avoid pitfalls of PROC GLM.[5] From the expected mean squares in output from test option in program 1,[12] the sequence effect should be tested based on the subject(seqence)

The GLM Procedure

Source	Type III Expected Mean Square
seq	Var(Error) + 2.4035 Var(subject(seq)) + Q(seq)
subject(seq)	Var(Error) + 2.4524 Var(subject(seq))
period	Var(Error) + Q(period)
trt	Var(Error) + Q(trt)

effect instead of MSE (error). Other fixed effects can be tested based on MSE (error). The programs 2 and program 3 gave the same AIC and BIC. The mixed effects model with compound symmetry variance-covariance matrix turned out to be the best among four programs in Table 1.

Table 2 presents the type 3 tests of fixed effects and only the treatment effect is significant at the significance level 0.10. Table 3 presents the multiple comparison results adjusted by Tukey. A and B, A and C, B and C are significantly different at the significance level 0.10 and each difference is estimated by lsmeans.

For a subject inapartial SAS output from program 3 shows that σ_e^2=47.9056 and the covariance between two unequal period is σ_b^2=73.9055, the variance at a particular period is $\sigma_e^2+\sigma_b^2$=121.1811, and the correlation between periods is $\sigma_b^2/\sigma_e^2+\sigma_b^2$ =0.6067.

Table 1. AIC and BIC for programs

Program	AIC	BIC
2	179.7	181.2
3	177.7	178.7
5 with cs	177.7	178.7
5 with un	181.9	184.8

AIC, information criteria of Akaike; BIC, information criteria of Schwar

Table 2. Type 3 tests of fixed effects

Effect	Numertor DF	Denomenator DF	F	*P*-value
QTcFb	1	16	3.79	0.0694
seq	5	5	1.08	0.4674
period	2	16	0.78	0.4766
Trt	2	16	33.18	<0.0001

DF, degree of freedom

Partial SAS output from Program 3

Estimated R Correlation Matrix for subject 1

Row	Col1	Col2	Col3
1	1.0000	0.6067	0.6067
2	0.6067	1.0000	0.6067
3	0.6067	0.6067	1.0000

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
CS	subject	73.9055
Residual		47.9056

Discussion

The complex structure of crossover studies for bioequivalence test of drugs still leaves many tasks to analysts. The expected mean square table in PROC GLM, however, provides the analysts a first good guide to a correct model frame. PROC MIXED based on REML in SAS, is then recommended since it provides the BLUP of the random between-subject effects and its vari-

Table 3. Differences of least squares means

TRT	TRT	Estimate	SE	DF	t	Adjusted *p*	Upper CI	Lower CI
A	B	-15.0223	3.0916	16	-4.86	0.0005	-20.4199	-9.6247
A	C	-25.0443	3.0927	16	-8.10	<0.0001	-30.4438	-19.6448
B	C	-10.0220	3.0761	16	-3.26	0.0129	-15.3925	-4.6514

SE, standard error; DF, degree of freedom; CI, confidence interval

ance. PROC MIXED with REPEATED is further recommended since it considers the within-subject correlations with the various variance-covariance structures which most PKPD studies and crossover studies have overlooked.

The correlation within subject in our study was 0.6067, which does not seem to be ignorable. Thus both the crossover design and the correlation within subject in the model helped even if fitting the mixed effects model with variance-covariance structure involves more complex mathematics and calculations. Kenwardroger option is also recommended for crossover designs even if we dropped it in some previous programs in order to keep the original codes. We can include the interaction, treatment*period, separately in the Williams design. Since the whole QT interval data were actually collected as a series over time, the model can fit again over the time variable. An additional time dimension will incorporate huge variance-covariance matrix which requires rigorous mathematics and computing work.

Acknowledgments

This study was supported by a grant of the Korea Healthcare Technology R&D Project, Ministry of Health & Welfare, Republic of Korea. (HI07C00010000)

Conflict of Interest

None of the authors have any conflicts of interest to disclose.

References

1. Rosner B. Fundamentals of Biostatistics. 7th ed. Cengage Learning, 2000;666-673.
2. Patterson S, Jones B. In: Keidling N (ed) Bioequivalence and Statistics in Clinical Pharmacology. Chapman and Hall/CRC, Boca Raton, 2006;39-77.
3. Wallenstein S, Fisher AC. The analysis of two-period repeated measurements crossover design with application to clinical trials. Biometrics 1977;33:261-269.
4. Woods JR, Williams JG, Tavel M. The two-period crossover design in medical research. Ann Intern Med 1989;110:560-566.
5. Cheng J, Olbricht G, Gunaratna N, Kendall R, Lipka A, Paul S, et al. Mixed Models, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.136.7749&rep=rep1&type=pdf, Accessed 29 October 2014.
6. Jennrich RI, Schluchter MD. Unbalanced repeated-measures models with structured covariance matrices. Biometrics 1986;42:805-820.
7. Wolfinger RM, Chang M. Comparing the SAS® GLM and MIXED Procedures for Repeated Measures, http://www.ats.ucla.edu/stat/sas/library/mixedglm.pdf, Accessed 29 October 2014.
8. Montgomery DC. Design and Analysis of Experiment, John Wiley & Sons, 1984;71.
9. Park SH. Design of Experiments, Min-Young-SA, 2006;58
10. Meyers LS, Gamst G, Guarino AJ. Applied Multivariate Research, Design and Interpretation, Sage, 2006;279-314.
11. Owen JS, Fiedler-Kelly J. Introduction to Population Pharmacokinetic/Pharmacodynamic Analysis with Nonlinear Mixed Effects Models, Wiley, New York, 2014;9-29.
12. Brunelle R. Review various methods to perform the analysis of a 2 treatment, 2 period crossover study http://www.math.iupui.edu/~indyasa/crosover.pdf, Accessed 29 October 2014.
13. SAS (2014)., http://support.sas.com/onlinedoc/913/getDoc/en/statug.hlp/glm_sect34.htm. Accessed 29 October 2014.
14. Saarinen F. Using mixed models in a cross-over study with repeated measurements within periods, http://www2.math.su.se/matstat/reports/serieb/2004/rep22/report.pdf, Accessed 29 October 2014.
15. OehLert GW. A few words about REML. http://users.stat.umn.edu/~corbett/classes/5303/REML.pdf, Accessed 29 October 2014.
16. Pinheiro CJ, Bates DM. Mixed-Effects Models in S and S-PLUS, Springer, 2004;57-82.
17. Kiernan K, Tao J, Gibbs P. Tips and strategies for mixed modeling with SAS/Stat procedures, SAS global forum 2012, http://support.sas.com/resources/papers/proceedings12/332-2012.pdf, Accessed 29 October 2014.
18. Frison L, Pocock SJ. Repeated measures in clinical trials: Analysis using mean summary statistics and its implications for design. Stat Med 1992;11:1685-1704.
19. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters III. Monoexponential model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm 1983;11:303-319.
20. Mauchly JW. Significance test for sphericity of a normal n-variate distribution. Annals of Mathematical Statistics 1940;11:204-209.
21. Anderson TW. An Introduction to Multivariate Statistical Analysis, New York: John Wiley & Sons, Inc.,1984;440-442.
22. Huynh H, Feldt LS. Conditions under Which Mean Square Ratios in Repeated Measurements Designs Have Exact F-Distribution. Journal of the American Statistical Association 1970;65:1582-1589.
23. Sakamoto Y, Ishiguro M, Kitagawa G. Akaike information criterion statistics. Psychometrika 1986;54:539-541.
24. Schwarz G. Estimating the dimension of a model. Ann Statistics 1978;6:461-464.