이 학 석 사 학 위 논 문

Analysis for Doubly Repeated Omics Data from Crossover Design

교차설계 실험에서 이중 반복 측정된 오믹스 자료 분석

2017 년 2 월

서울대학교 대학원

통계학과

최 성 훈

Analysis for Doubly Repeated Omics Data from Crossover Design

지도교수 박 태 성

이 논문을 이학석사 학위논문으로 제출함

2017 년 2 월

서울대학교 대학원

통계학과

최 성 훈

최성훈의 이학석사 학위논문을 인준함

2017 년 2 월

위 원 장 김 용 대 (인)

부위원장 박 태 성 (인)

위 원 오 희 석 (인)

Abstract

Analysis for Doubly Repeated Omics Data from Crossover Design

Sunghoon Choi

Department of Statistics

The Graduate School

Seoul National University

Some crossover clinical trials produce doubly repeated omics data with two repeated factors. Linear mixed effect models (LMMs) are commonly applied to the data from the crossover design focusing on the analysis of repeatedly observed omics data themselves. Alternatively, the univariate analyses using the single summary measurements such as differences between time points and incremental area under curve (iAUC) are also widely used. In this study, we propose LMMs to simultaneously analyze

several summary measures by taking their correlations into account. We compare the performance of the proposed method with other existing methods for real doubly repeated omics data from a crossover study. We show that our method has less number of parameters but with equivalent power.

Key words: Linear mixed effect model (LMM), Crossover design, Repeated measurements

Student number: 2014-22362

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In crossover designs, especially, subjects serve as own control so that the variability between subjects becomes isolated, which allows the analysis be able to focus on treatment effects more precisely [1]. Since linear mixed effect models (LMMs) consider correlations into account among the measurements from the same subject, LMMs are commonly applied to the analysis of data from the crossover design [2]. Some crossover clinical trials have doubly repeated measurements in which there exist two repeated factors. For example, in our study of Oral Lipid Tolerance Test (OLTT) of Angelica keiskei, each subject intakes placebo, low dose, high dose of Angelica keiskei respectively in week 1, 2, and 3, in a randomly assigned crossover sequence. Then, every week the

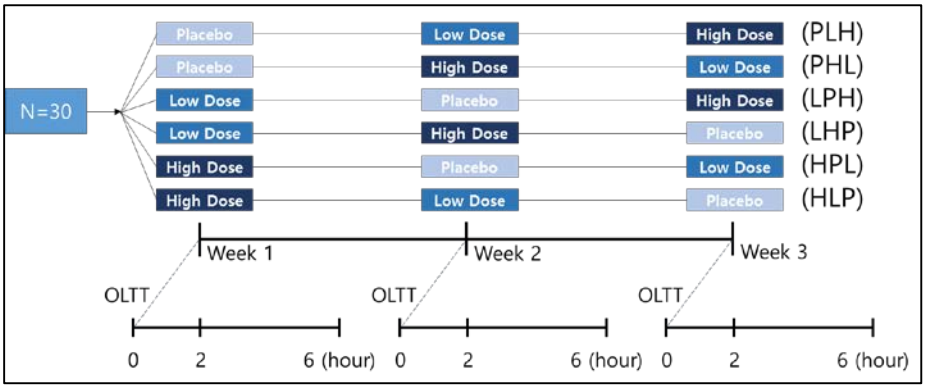

Figure 1. Three-level crossover design with two repeated factors

proteomics data are observed repeatedly at three time points: baseline (0 hour), after 2 hours, 6 hours, i.e. the repeated factors are week and time (Figure 1).

In doubly repeated measured crossover design, LMMs often focus on the raw repeated measures themselves [3, 4]. Alternatively, the univariate analyses using the single summary measurements such as differences of observations between time points and incremental area under curve (iAUC) of repeated measures are also widely used [5]. In this study, we propose LMMs in simultaneously analyzing several summary measures by taking their correlations into account appropriately. Through the analysis of OLTT of Angelica keiskei, we compare the performance of the proposed method with that of existing LMMs and univariate analysis using single summary measurements. We show that our method has less number of parameters but with equivalent power.

Chapter 2

Materials and Notation

2.1 Materials

Our study was designed to use cross-over design experiment with 30 low-grade inflammation holders in order to exclude the effects of the individual in sample. 30 samples were divided into six groups with five members in each group and randomly assigned to six sequences of PLH (Placebo, Low dose, High dose), PHL, LPH, LHP, HLP, and HPL (Figure 1). They were assigned to intake Placebo, Low dose, High dose of Angelica keiskei juice in a sequence order through three weeks. During the experiment, they were assigned to intake the Netherlands

Organization for applied scientific research (TNO) fat challenge formula with Angelica keiskei juice solution to test the effect of dietary fat load on inflammatory response as well as to observe various omics data by analyzing blood collection of samples who took it before taking it (0 hour), after 2 hours, and 6 hours of ingestion.

There are 44 bio-markers including lipids (ox-LDL, TC, TG, HDL, LDL), MDA, and proteomic markers obtained by Luminex multiplex assay. We build models that evaluate the effect of Angelica keiskei intake on inflammatory response with omics technology to observe changes in the body inflammation by analyzing those bio-markers after loading the dietary fat challenge.

2.2 Notation

Let y_{ijk} be the observed value of *i*-th subject on *j*-th week and *k*-th time point, and $y'_{ij\cdot}$ be the summarized values of *i*-th subject on *j*-th week. β_p are fixed effect parameters. γ is the zero-mean normally distributed random subject effect parameter, $\gamma_i \sim N(0, \sigma_s^2)$. Also, ϵ_{ijk}, $\epsilon_{ij\cdot}$ are zero-mean normally distributed random errors, such that $\epsilon_{ijk} \sim N(0, \sigma_{e_1}^2)$, $\epsilon_{ij\cdot} \sim N(0, \sigma_{e_2}^2)$ independent with γ respectively.

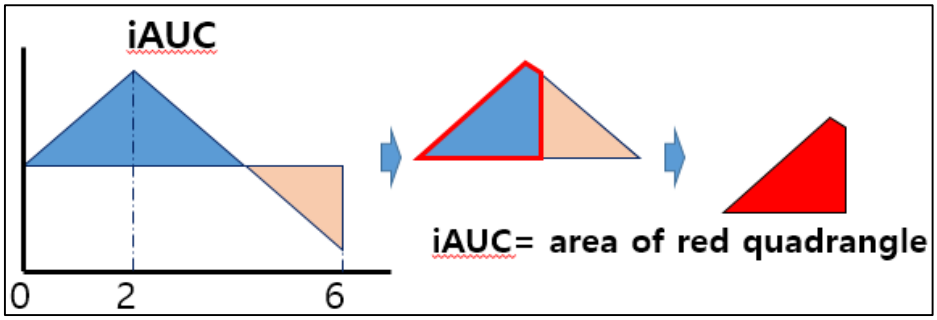

Figure 2. Incremental area under curve (iAUC)

In our study, we use the iAUC, as the summarized values [5, 6]. The iAUC includes all area below the curve and above the baseline [6]. The iAUC values of *i*-th subject on *j*-th week are denoted by $y_{ij\cdot}^{\text{iAUC}}$, and calculated by applying the trapezoid rule [7]. The process of the calculation is demonstrated in Figure 2, and the following formula is the result of the calculation using values of time points in this study design; 0, 2, and 6 hours after dosage. In addition, we use the differences between time points as another type of summarized values. $y_{ij\cdot}^{\Delta k_1 k_2}$ denotes the difference between k_1 and k_2 time points. Both summary measurements can be easily calculated as follows.

$$y_{ij\cdot}^{iAUC} = -5y_{ij1} + 3y_{ij2} + 2y_{ij3}$$
$$y_{ij\cdot}^{\Delta k_1 k_2} = y_{ijk_2} - y_{ijk_1}$$

Chapter 3

Method

Our aim of this study is to find bio-markers of which trends over time are affected by treatments. In other words, we want to find out which markers have significant *Treatment* effect or *Treatment* × *Hour* interaction effect. We consider three LMMs for doubly repeated measurements from crossover design using linear mixed effect model.

3.1 Linear Mixed Effect Model (LMM-Y)

First of all, we consider *Subject* as random intercept and build a model with two repeated effects using the Kronecker product as $R = \Sigma \otimes \Psi$, where Σ is usually unstructured and Ψ is structured such as compound

symmetry or 1st order autocorrelation. [4, 8]. We denote this method by LMM-Y and is given by as follows:

$$y_{ijk} = \beta_0 + \beta_1 Treatment_{ij} + \beta_2 Week_j + \beta_3 Hour + \beta_4 Baseline_{ij} + \beta_5 Sequence_{i_k} + \beta_6 (Treatment \times Hour)_{ijk} + \gamma_i + \epsilon_{ijk}$$

Since we have three levels of treatment and three time points, the unstructured (UN) covariance matrix of the first repeated factor (*Treatment*), Σ, and the compound symmetry (CS) correlation matrix of the second repeated factor (*Hour*), Ψ can be represented as follows.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{pmatrix}$$

$$\Psi = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

Then, the covariance structure of repeated measures $R = \Sigma \otimes \Psi$ is 9×9 matrix.

3.2 LMM for Univariate Summary Measurements (LMM-US)

Secondly, we transform the observed data into some summary measurements, iAUC[5] and differences. Since we summarize a vector $(y_{ij1}, y_{ij2}, y_{ij3})'$ into a value $y'_{ij\cdot}$, *Hour* which is one of the repeated factors is omitted. Considering the random subject intercept and compound symmetry structure for another repeated factor, *Treatment*, the model is given by as follows.

$$y'_{ij\cdot} = \beta_0 + \beta_1 Treatment_{ij} + \beta_2 Week_j + \beta_3 Baseline_{ij} + \beta_4 Sequence_i + \gamma_i + \epsilon_{ij\cdot}$$

Since we have three levels of treatment, the covariance matrix of repeated measures, *R*, is 3×3 matrix as follows.

$$R = \begin{pmatrix} \sigma^2+\sigma_1 & \sigma_1 & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & \sigma_1 \\ \sigma_1 & \sigma_1 & \sigma^2+\sigma_1 \end{pmatrix}$$

We denote this method as LMM-US.

3.3 LMM for Multivariate Summary Measurements (LMM-US)

Lastly, we propose a new combined method to analyze l types of summary measurements jointly by combining the multiple summary measures into a response vector. We use LMM to consider the correlation among different types of summary measurements. In this case, iAUC can explain the total amount of observation through three time points, and differences may account for changes in the body after intakes. The proposed method is expected to show performance better than or equal to that of LMM-Y, by transforming a response vector into more optimized summary values, while using less number of parameters.

We build a random subject intercept model with Kronecker product covariance structure, $\mathrm{R} = \Sigma \otimes \Psi$, where Σ is unstructured and Ψ is compound symmetry for doubly repeated factors.

$$\begin{aligned}\begin{pmatrix} y^1_{\mathrm{ij}\cdot} \\ y^2_{\mathrm{ij}\cdot} \end{pmatrix} = \begin{pmatrix} \beta^1_0 \\ \beta^2_0 \end{pmatrix} + \begin{pmatrix} \beta^1_1 \\ \beta^2_1 \end{pmatrix} Treatment_{ij} + \begin{pmatrix} \beta^1_2 \\ \beta^2_2 \end{pmatrix} Week_j \\ + \begin{pmatrix} \beta^1_3 \\ \beta^2_3 \end{pmatrix} Baseline_{ij} + \begin{pmatrix} \beta^1_4 \\ \beta^2_4 \end{pmatrix} Sequence_i + \begin{pmatrix} \gamma^1_i \\ \gamma^2_i \end{pmatrix} \\ + \begin{pmatrix} \epsilon^1_{\mathrm{ij}\cdot} \\ \epsilon^2_{\mathrm{ij}\cdot} \end{pmatrix} \qquad \text{if } l = 2\end{aligned}$$

$$\begin{pmatrix} y_{\mathrm{ij}\cdot}^{1} \\ y_{\mathrm{ij}\cdot}^{2} \\ y_{ij.}^{3} \end{pmatrix} = \begin{pmatrix} \beta_0^1 \\ \beta_0^2 \\ \beta_0^3 \end{pmatrix} + \begin{pmatrix} \beta_1^1 \\ \beta_1^2 \\ \beta_1^3 \end{pmatrix} Treatment_{ij} + \begin{pmatrix} \beta_2^1 \\ \beta_2^2 \\ \beta_2^3 \end{pmatrix} Week_j$$

$$+ \begin{pmatrix} \beta_3^1 \\ \beta_3^2 \\ \beta_3^3 \end{pmatrix} Baseline_{ij} + \begin{pmatrix} \beta_4^1 \\ \beta_4^2 \\ \beta_4^3 \end{pmatrix} Sequence_i + \begin{pmatrix} \gamma_i^1 \\ \gamma_i^2 \\ \gamma_i^3 \end{pmatrix}$$

$$+ \begin{pmatrix} \epsilon_{\mathrm{ij}\cdot}^{1} \\ \epsilon_{\mathrm{ij}\cdot}^{2} \\ \epsilon_{\mathrm{ij}\cdot}^{3} \end{pmatrix} \qquad \text{if } l = 3$$

Let $\boldsymbol{\gamma}_{\mathrm{i}} = (\gamma_i^1,\ \gamma_i^2)$ *and* $\boldsymbol{\epsilon}_{ij\cdot} = (\epsilon_{ij\cdot}^1,\ \epsilon_{ij\cdot}^2)$, then

$$\mathrm{E}\left[\begin{pmatrix} \boldsymbol{\gamma}_{\mathrm{i}} \\ \boldsymbol{\epsilon}_{ij\cdot} \end{pmatrix}\right] = \mathbf{0}_4$$

$$\mathrm{Var}\left(\begin{pmatrix} \boldsymbol{\gamma}_i \\ \boldsymbol{\epsilon}_{ij\cdot} \end{pmatrix}\right) = \begin{pmatrix} G & \mathbf{0} \\ \mathbf{0} & R \end{pmatrix}$$

Since we have l types of summary measurements and three levels of treatment, Σ, the unstructured (UN) covariance matrix of the first repeated factor (*Summary measurements type*) and Ψ, the compound symmetry (CS) correlation matrix of the second repeated factor (*Treatment*) can be represented as follows.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \qquad \text{if } l = 2$$

$$= \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{pmatrix} \text{ if } l=3$$

$$\Psi = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix}$$

Then, the covariance structure of repeated measures is $R = \Sigma \otimes \Psi$, which is 6×9 matrix in case of $l = 2$ or 9×9 matrix in case of $l = 3$.

Chapter 4

Result

We compare three models by applying to OLTT data of Angelica keiskei, which have two repeated factors, week and hour. Among 44 bio-markers, in order to find significant bio-markers or trends of bio-marker which are affected by three levels of treatments, we start with modeling three methods for each bio-marker, respectively. Then, we find bio-markers which have significant *Treatment* effects or *Treatment* $\times$ *Hour* interaction effects in LMM-Y. Since *Hour* effects are summarized into iAUC or differences in LMM-US or LMM-MS, it suffices to find bio-markers with only significant treatment effects.

Tables 1 to 7 are the results of F-test for fixed effects in each model. The F-test statistic is calculated using Wald type χ^2 statistic. The p-values were found by computing the degrees of freedom. NumDF is the

numerator degree of freedom and DenDF, the denominator degree of freedoms, is calculated by Kenward-Rodger procedure[9] which is recommended for calculating DenDF in crossover design[10, 11]. At the 5% significance level, the list of significant effects is in bold letters along with their p-values.

4.1 LMM-Y

There exist one bio-marker having a significant treatment effect, and two bio-markers for interaction $Treatment \times Hour$ effects in the first model. Table 1 is the summary of type 3 tests of those significant bio-markers.

Table 1. Tests of fixed effect of LMM-Y (α=0.05)

Bio-marker	Effect	NumDF	DenDF	FValue	ProbF
HDL	**Treatment**	2	39.1	4.1	**0.0242**
	Hour	1	41.6	1.25	0.2706
	Treatment ×Hour	2	39.7	2.23	0.1207
	Sequence	5	36.8	0.94	0.4668

	Week	2	64.5	0.79	0.4598
	Baseline	1	47	1965.07	<.0001
IL-9	Treatment	2	34.4	0.9	0.4168
	Hour	1	43.3	2.75	0.1048
	Treatment ×Hour	2	23.7	3.67	**0.0409**
	Sequence	5	18.2	0.69	0.6382
	Week	2	45.8	0.48	0.6213
	Baseline	1	30.8	177.69	<.0001
IL1- α	Treatment	2	31.6	2.14	0.1348
	Hour	1	30.4	2.65	0.114
	Treatment × Hour	2	34.7	4.71	**0.0155**
	Sequence	5	19.1	1.51	0.2339
	Week	2	48.1	1.63	0.2057
	Baseline	1	49.4	1.87	0.1779

4.2 LMM-US

In the LMM-US, we use three types of summary measurement, respectively, iAUC, differences.

i. LMM-US-1

Using the iAUC ($y_{ij\cdot}^{iAUC}$) as the summarized response, we can find two bio-markers (HDL and IL-alpha) that have significant treatment effects.

Table 2. Tests of fixed effect of LMM-US-1 (α=0.05)

Bio-marker	Effect	NumDF	DenDF	FValue	ProbF
HDL	**Treatment**	2	55.5	3.68	**0.0316**
	Sequence	5	23	0.69	0.6351
	Week	2	56	0.77	0.4668
	Baseline	1	30.9	2.01	0.1664
IL1-α	**Treatment**	2	44.9	3.4	**0.0423**
	Sequence	5	15.7	1.52	0.2394
	Week	2	44.8	1.19	0.3141
	Baseline	1	70.2	196.87	<.0001

ii. LMM-US-2

Using the difference between observations at 0 hour (baseline) and 2 hour after dosage ($y_{ij\cdot}^{\Delta 02}$) as the summarized response, there is no significant bio-marker detected for the treatment effect at the 5% significant level.

iii. LMM-US-3

In the LMM-US-3, we consider modeling the difference between baseline and 6 hour after treatment ($y_{ij\cdot}^{\Delta 06}$) as the summarized response. We find two bio-markers that have significant treatment effects.

Table 3. Tests of fixed effect of LMM-US-3 (α=0.05)

Bio-marker	Effect	NumDF	DenDF	FValue	ProbF
IL1-RA	**Treatment**	2	55.8	5.11	**0.0092**
	Sequence	5	23.4	0.27	0.9242
	Week	2	56.5	1.13	0.3314
	Baseline	1	34	2.7	0.1096
	Treatment	2	55.9	3.49	**0.0374**
	Sequence	5	23.1	2.17	0.0929
	Week	2	55.9	1.98	0.1473
	Baseline	1	25.4	14.63	0.0008

4.3 LMM-MS

We use several summary measurements jointly in one model. Multivariate summary measurement LMM. iAUC is calculated in the form of weighted average, so that it can reflect the overall amount of expressions. And the difference between time points is expected to account for the trend over time.

We consider the following three LMM-MS:

i. LMM-MS-1 for $y_{ij\cdot}^{iAUC}$, $y_{ij\cdot}^{\Delta 02}$

ii. LMM-MS-2 for $y_{ij\cdot}^{iAUC}$, $y_{ij\cdot}^{\Delta 06}$

iii. LMM-MS-3 for $y_{ij\cdot}^{iAUC}$, $y_{ij\cdot}^{\Delta 02}$, $y_{ij\cdot}^{\Delta 06}$

i. LMM-MS-1

We model multivariate summary measurement LMM with iAUC ($y_{ij\cdot}^{iAUC}$) and difference between 0 and 2hour ($y_{ij\cdot}^{\Delta 02}$) as summarized response variables. LMM-MS-1 finds three bio-markers MDC, IL-1α, and MIP-1α that have significant treatment effects.

Table 4. Tests of fixed effect of LMM-MS-1 (α=0.05)

Bio-marker	Effect	NumDF	DenDF	FValue	ProbF
IL1-α	**Treatment**	2	56.4	3.68	**0.0316**
	Sequence	5	18.7	0.63	0.6817
	Week	2	57.2	0.07	0.9354
	Baseline	1	26.5	22.85	<.0001
MDC	**Treatment**	2	53.3	5.51	**0.0067**
	Sequence	5	24.5	1.87	0.1364
	Week	2	53.2	0.85	0.4339
	Baseline	1	92	122.32	<.0001
MIP-1α	**Treatment**	2	54.9	3.23	**0.0473**
	Sequence	5	17.7	2.19	0.1009
	Week	2	54.5	0.87	0.4227
	Baseline	1	20.9	7.84	0.0108

ii. LMM-MS-2

We fit the proposed method with iAUC ($y_{ij\cdot}^{iAUC}$) and difference of observation between baseline and 6 hours after dosage ($y_{ij\cdot}^{\Delta 06}$) as the

response variables. Two bio-markers HDL and IL-1RA have significant treatment effects.

Table 5. Tests of fixed effect of LMM-MS-2 (α=0.05)

Bio-marker	Effect	NumDF	DenDF	FValue	ProbF
HDL	**Treatment**	2	75	7.22	**0.0014**
	Sequence	5	39.5	0.42	0.8339
	Week	2	75.5	1.5	0.2291
	Baseline	1	54	3.1	0.084
IL1-RA	**Treatment**	2	57.5	3.46	**0.0382**
	Sequence	5	27.9	3.05	0.0256
	Week	2	57.5	1.85	0.1668
	Baseline	1	30.5	22.35	<.0001

iii. LMM-MS-3

In LMM-MS-3 we consider three summary measures jointly, iAUC ($y_{ij\cdot}^{iAUC}$), difference between 0 and 2 hour (($y_{ij\cdot}^{\Delta 02}$), and difference between 0 and 6 hour ($y_{ij\cdot}^{\Delta 06}$) with $l = 3$. We identify three significant bio-markers HDL, Flt3 ligand, and MIP-1α.

Table 6. Tests of fixed effect of LMM-MS-3 (α=0.05)

Bio-marker	Effect	NumDF	DenDF	FValue	ProbF
HDL	**Treatment**	2	131	6.08	**0.003**
	Sequence	5	63.8	1.01	0.42
	Week	2	132	1.26	0.2866
	Baseline	1	86.3	2.08	0.1524
Flt3 ligand	**Treatment**	2	102	3.22	**0.0439**
	Sequence	5	47.5	1.19	0.3303
	Week	2	103	1.37	0.2577
	Baseline	1	148	366.95	<.0001
MIP-1α	**Treatment**	2	107	5.27	**0.0066**
	Sequence	5	39.8	3.84	**0.0062**
	Week	2	106	2.62	0.0772
	Baseline	1	46.8	17.56	0.0001

The results are summarized in Table 7 which show the list of significant bio-markers for each model. $\Delta 02$ indicates the difference between 0 and 2 hour, $\Delta 06$ means the difference between 0 and 6 hour.

LMM-MS-1 have three bio-markers showing significant treatment effects with same number as LMM-Y but the list of significant bio-

markers is not the same. This means that proposed method can find the significant bio-markers which might not be detected in LMM-Y and vice versa. HDL has significant treatment effect in LMM-Y and other various models. However, LMM-MS-1 did not detect the significance of HDL (p-value= 0.1236). In testing MDC, LMM-Y did not find significance of MDC whereas, our LMM-MS-1 identified the significant effect. We got the p-value of the test $H_0 : \beta_1 = 0$ of 0.4284 and p-value of the test $H_0 : \beta_6 = 0$ of 0.3778 by using LMM-Y. MDC and MIP-1α, which are significant with LMM-MS-1, but not with LMM-Y, are known as inflammation-related bio-markers. This illustration shows that LMM-MS can successfully find bio-markers that cannot be detected by LMM-Y, which demonstrates its higher power than that of the other methods.

We compare three models: LMM-Y, LMM-US, LMM-MS. Since we are interested in two effects of LMM-Y, treatment and interaction, we will compare -$\log_{10}$p-values of four tests with pairwise plots (Figure 3), including p-value of treatment × hour effect in LMM-Y. We found deflation in pairwise plot of LMM-MS-1 and other models, indicating that LMM-MS-1 tends to have smaller p-values.

Table 7. List of significant bio-markers of each models.

	Effect	Bio-markers		
LMM-Y	Treatment	HDL		
	Treatment × Hour	IL-9	IL-1α	

	[Model] Summary type	Bio-markers		
LMM-US	[1] iAUC	HDL	IL-1α	
	[2] Δ02			
	[3] Δ06	HDL	IL-1RA	
LMM-MS	[1] iAUC & Δ02	MDC	IL-1α	MIP-1α
	[2] iAUC & Δ06	HDL	IL-1RA	
	[3] iAUC & Δ02 & Δ06	HDL	Flt3 ligand	MIP-1α

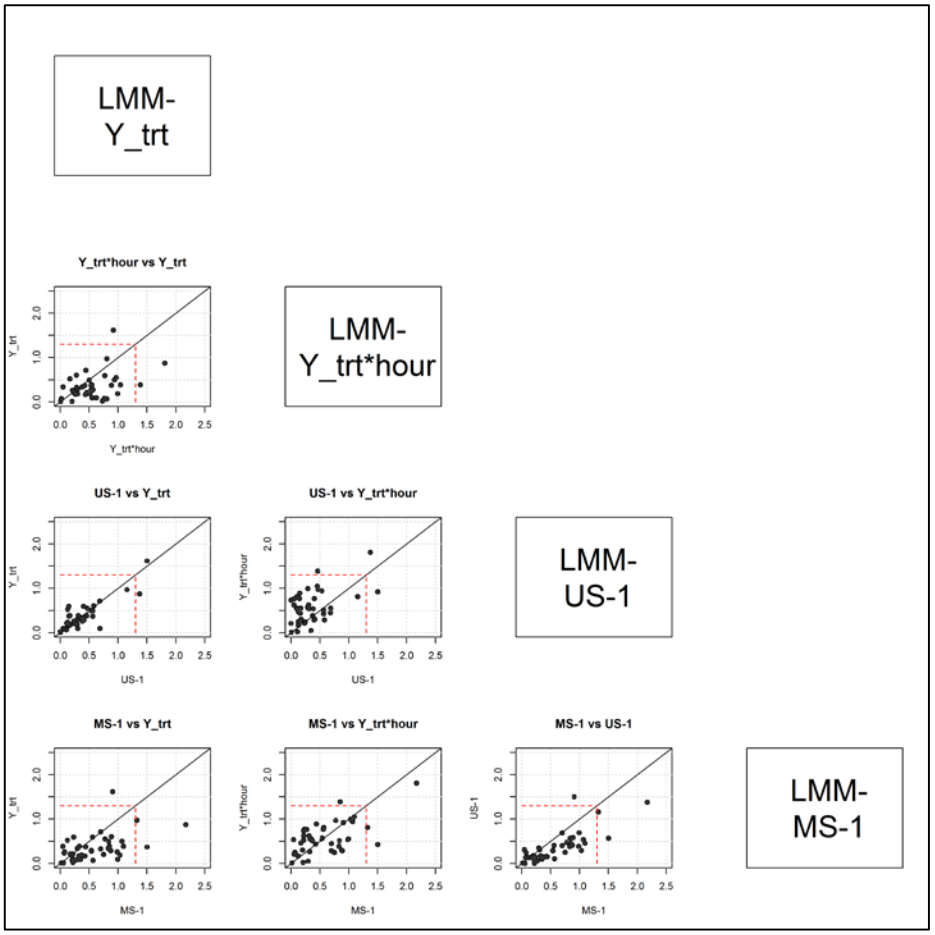

Figure 3. Pairwise plot of log p-value

Chapter 5

Discussion

We found that every LMM has different significant bio-markers. Thus, we should be more cautious when using LMMs for detecting significant bio-markers. The proposed LMM-MS-1 identified same number of bio-markers as LMM-Y. Both LMM-Y and LMM-MS-1 find significance of IL-1α. However, HDL and IL-9 are significant with LMM-Y, but not with LMM-MS-1. On the other hand, MDC and MIP-1α are significant with LMM-MS-1, but not with LMM-Y. Those two bio-markers, MDC and MIP-1α, are known to be related to inflammation [12-14]. Furthermore, LMM-MS-1 has smaller p-value than other models as shown in pairwise plots (Figure 3). Although a further simulation study is required, our proposed method seems to be a good

supplementary method to analyze repeated measurement with two repeated factors.

We used only iAUC and difference between two time points, but it might not be an optimal type of summary measurements. When fluctuated patterns exist for some markers, the difference between two time points does not summarize these patterns well, because it only considers linear change over time. In a future study, we consider extending the proposed method with other type of summary measurements such as the curvature or the 2^{nd} degree coefficient of quadratic approximations.

Bibliography

1. Wellek, S. and M. Blettner, *On the proper use of the crossover design in clinical trials*. Dtsch Arztebl Int, 2012. **109**(15): p. 276-281.
2. Jones, B. and M.G. Kenward, *Design and analysis of cross-over trials*. 2014: CRC Press.
3. Zhou, L., *Application of linear mixed-effects models to crossover designs*. 2012.
4. Tao, J., K. Kiernan, and P. Gibbs, *Advanced Techniques for Fitting Mixed Models Using SAS/STAT® Software*. Paper SAS, 1919. **2015**.
5. Wopereis, S., et al., *Assessment of inflammatory resilience in healthy subjects using dietary lipid and glucose challenges*. BMC medical genomics, 2013. **6**(1): p. 1.
6. Wolever, T.M., *Effect of blood sampling schedule and method of calculating the area under the curve on validity and precision of glycaemic index values*. British Journal of Nutrition, 2004. **91**(02): p. 295-300.
7. Shiang, K. *The SAS calculations of areas under the curve (AUC) for multiple metabolic readings*. in *SAS Conference Proceedings Western Users of SAS (WUSS)*. 2004.
8. Moser, E.B. *Repeated measures modeling with PROC MIXED*. in *Proceedings of the 29th SAS Users Group International Conference*. 2004.
9. Kenward, M.G. and J.H. Roger, *Small sample inference for fixed effects from restricted maximum likelihood*. Biometrics, 1997: p. 983-997.
10. Cheng, J., et al., *Mixed models*. 2005, Unknown.
11. Choi, K., T. Hong, and J. Lee, *On comparison of SAS codes with GLM and MIXED for the crossover studies with QT interval data*. Translational and Clinical Pharmacology, 2014. **22**(2): p. 78-82.
12. Mantovani, A., et al., *Macrophage-derived chemokine (MDC)*. Journal of leukocyte biology, 2000. **68**(3): p. 400-404.
13. Matsukawa, A., et al., *Pivotal role of the CC chemokine, macrophage-derived chemokine, in the innate immune response*. The Journal of Immunology, 2000. **164**(10): p. 5362-5368.
14. Cook, D.N., *The role of MIP-1 alpha in inflammation and hematopoiesis*. Journal of leukocyte biology, 1996. **59**(1): p. 61-66.

초 록

일부 교차 설계 임상 시험은 두 개의 반복 요인에 대해 이중으로 반복된 오믹스 자료를 생산한다. 일반적으로 교차설계 분석에는 반복 측정된 오믹스 자료 그 자체에 초점을 맞추어 크로네커 곱을 이용한 공분산 구조를 이용하여 분석하는 선형 혼합 모형이 이용된다. 한 편, 두 번째 반복요인인 시간에 걸쳐 반복 측정된 벡터를 iAUC (incremental Area Under Curve), 또는 변화량 등의 요약 값으로 변환하여 반응변수로 적합하는 일변량 분석도 널리 사용된다. 이에 필자는 서로 다른 둘 이상의 요약 값 사이의 상관성을 크로네커 곱을 이용한 공분산 구조를 통해 고려하며 요약 값들을 동시에 적합하는 다변량 선형 혼합모형을 제안한다. 제안하는 분석 방법의 성능을 시뮬레이션을 통해 확인하였으며 실제 이중 반복된 교차설계 연구에 적용하여 기존의 두 방법들과 결과를 비교하였다. 또한 제안하는 방법이 더 적은 수의 모수를 사용하며 동등한 검정력을 보이는 것을 확인하였다.

주요어: 선형혼합모형, 교차 설계, 반복 측정

학 번: 2014-22362