Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Danni Liu Date

Grizzle Linear Model for Analysis of Clinical Trials Using Crossover Designs

By

Danni Liu

Master of Science in Public Health

Biostatistics

_____ Thesis Advisor's signature

Mary Kelley, PhD

Thesis Advisor

_____ Reader's signature

Rebecca Zhang

Reader

Grizzle Linear Model for Analysis of Clinical Trials Using Crossover Designs

By

Danni Liu

B.S., Georgia State University, 2013
MSPH, Rollins School of Public Health of Emory University, 2015

Thesis Committee Chair: Mary Kelley, PhD

An abstract of
A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics
2015

Abstract

Grizzle Linear Model for Analysis of Clinical Trials Using Crossover Designs

By Danni Liu

Background: The Melatonin and Metabolic Syndrome (MetSyn) study was a double-blind, placebo-controlled, crossover, phase II randomized clinical trial. Its main purpose was to investigate the effects of melatonin supplementation on treating patients with metabolic syndrome. This thesis analyzes results of this crossover trial with an emphasis on examining the univariate and multivariate effects of carry-over and treatment on the main components of metabolic syndrome.

Methods: In addition to fitting the traditional linear model (Grizzle) to detect the effects of carry-over and treatment on various measures, we tested linear combinations of the five main components of metabolic syndrome derived from a principal component analysis.

Results: Only one variable considered, a measure of sleep efficiency, showed a significant effect of carry-over ($F=6.29$, $p=0.02$). A significant treatment effect was detected for the average change of clinic SBP comparing melatonin with placebo ($F=6.86$, $p=0.01$). However, no treatment effects were found on any other measures. Interestingly, the first component retained from the principal component analysis reflected mainly changes of triglycerides and waist circumference, while the second represented mainly the changes in HDL and SBP average. Further, one of the two principal components retained from the principal component analysis exhibited a significant treatment effect ($F=4.70$, $p=0.04$).

Conclusions: Given that no significant carry-over effect was found among most of the measures of interest, the crossover design of this phase II clinical trial was considered appropriate. The significant treatment effect on clinic SBP suggested a promising benefit of melatonin supplementation in SBP improvement. Since the principal component analysis yielded two distinct principal components, it indicated the five major components of the metabolic syndrome may interact through different mechanisms and this should be considered in future studies. Thus the use of the data reduction method (principal component analysis) provided a multivariate solution to this crossover trial analysis that further showed the additional advantage of melatonin in improving HDL.

Keywords: crossover trials, Grizzle model, principal component analysis, metabolic syndrome, melatonin supplementation

Grizzle Linear Model for Analysis of Clinical Trials Using Crossover Designs

By

Danni Liu

B.S., Georgia State University, 2013
MSPH, Rollins School of Public Health of Emory University, 2015

Thesis Committee Chair: Mary Kelley, PhD

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Biostatistics
2015

Acknowledgements

I would like to extend my deepest gratitude to my thesis advisor, Dr. Mary Kelley. This work would not have been possible without your extensive knowledge and constant support. I would like to say thank you for giving me excellent guidance and constructive comments, and I am so deeply grateful to you for your kindness, understanding, and patience throughout the whole experience. I will always cherish the insights and wisdom that you shared.

My sincere appreciation also goes to Rebecca Zhang. Thank you for spending time reading my thesis. I am grateful to you for your insightful comments and helpful suggestions. Also, thank you for teaching such an amazing probability theory course that brings me into the sophisticated probability world. I would also like to thank Dr. Paul Terry for his critical comments and valuable input regarding this thesis.

A very special thanks goes out to my Biostatistics and RSPH family. Every classmate, staff and professor that I have met or worked with is very supportive and kind. I have always felt so welcome and enjoyed my time so much at Rollins. I know for sure that these little moments at Rollins Biostatistics in the last two years will eventually become precious memories that I will treasure forever.

Last, but not least, I would like to express my profound gratitude to my parents for the unconditional love, continuous encouragement and endless support you have given me through my entire life. Thank you for always believing in me and encouraging me to pursue my goals and dreams. Thank you for always inspiring me to do my best. Thank you for everything.

Table of Contents

Chapter 1

Introduction

The dataset used for this thesis, the Melatonin and Metabolic Syndrome (MetSyn) study, was a double-blind, placebo-controlled, crossover, phase II randomized clinical trial which was conducted between 2010 and 2012. The purpose of the study was to investigate the effects of melatonin supplementation on treating patients with metabolic syndrome. Registered at ClinicalTrials.Gov (NCT01038921), the trial was approved by Emory University institutional review board (14784), and the Investigational New Drug approval was given by the US Food and Drug Administration (105764).

In order to be eligible for the study, the subjects had to meet at least three of the five criteria for metabolic syndrome established by the National Heart, Lung, and Blood Institute and the American Heart Association. (Grundy, Brewer, Cleeman, Smith, & Lenfant, 2004). In addition, subjects who were smokers, had diabetes, or were taking calcium channel blockers were excluded. After the screening and run-in phases, a total of 39 women and men with metabolic syndrome were recruited from sources within Emory Healthcare. A placebo run-in phase was initiated before randomization for the purpose of compliance evaluation. An independent biostatistician then randomized these 39 subjects for determining whether each subject would receive 8.0mg of oral melatonin supplementation or placebo for the first 10 weeks. Subjects then underwent a 6-week washout period in order to make sure the potential carry-over effect was

ruled out before they crossed-over to the other treatment for another 10 weeks. The assigned treatment was taken once per day one hour before bedtime during each period. The dosage level of melatonin was chosen at 8mg based on prior studies which showed a 7-8mg dose level of melatonin would exhibit similar side effects compared to placebo. (Buscemi et al., 2006). The duration of study treatment administration was determined as ten weeks, according to previous human studies which demonstrated effects of melatonin on blood pressure, glucose and serum lipids. (Koziróg, et al., 2011). The treatment assignment was blind to the study investigators, participants and their personal health care providers, and laboratory staff. Only two subjects dropped out from the study during follow-up, but the reasons were not considered to be treatment-related.

Each of the five components of the metabolic syndrome were measured at the following five time points: 1) screening, 2) the beginning of the first ten-week period of the first treatment administration, 3) the end of the first ten weeks, 4) the beginning of the second ten-week period of the alternative treatment (i.e., the end of the six-week washout period), and 5) the end of the second ten weeks. The 24-hour ambulatory blood pressure was also measured at each time point except during screening. Other study variables included interview data, endogenous melatonin level, sleep duration and quality, and oxidative stress and inflammation biomarkers.

The primary outcome for the MetSyn study was to examine the mean change in each of the five metabolic syndrome components under melatonin supplementation treatment, in comparison to placebo. Researchers were also interested in a secondary outcome, which was the proportion of the patients who

were free from the metabolic syndrome. Preliminary analyses of this study have showed good tolerance of melatonin and high adherence among patients. (Goyal, et al., 2014). Further studies were recommended for assessing the efficacy of melatonin on treating metabolic syndrome as well as its downstream cardiovascular and metabolic complications. (Goyal, et al., 2014).

The purpose of this thesis was to analyze the results from this crossover trial with an emphasis on examining the effects of carry-over and treatment. The validity of the crossover design was inspected through the detection of the carry-over effects. In addition to the univariate effects, the multivariate solution was explored by considering the linear combinations among the outcome measures of interest. We aimed to provide a further understanding of the effects of melatonin through the combined statistical analyses of this crossover trial.

Chapter 2

Review of Literature

2.1 The Clinical Problem

2.1.1 Metabolic Syndrome

The metabolic syndrome is a collection of interconnected metabolic risk factors which may increase the risk of cardiovascular diseases and type 2 diabetes mellitus, and may also be associated with all-cause mortality. (Ford, 2005). In 2004, the American Heart Association/National Heart, Lung, and Blood Institute (AHA/NCLBI) established a definition for metabolic syndrome in adults which requires a minimum of three out of five criteria to be met: enlarged waist circumference (≥ 102 cm in men, ≥ 88 cm in women), elevated triglycerides (≥ 150 mg/dl), reduced HDL-cholesterol (< 40 mg/dl in men, < 50 mg/dl in women), elevated blood pressure (≥ 130/85 mmHg), and elevated fasting glucose (≥ 100 mg/dl). (Grundy, Brewer, Cleeman, Smith, & Lenfant, 2004). Although four other diagnostic criteria are also commonly used for metabolic syndrome, prevalence estimates are found to be similarly high and rising in western societies. (Hollman & Kristenson, 2008; Hillier, et al., 2006; do Carmo, et al., 2008). The estimated age-adjusted prevalence of metabolic syndrome in the U.S., for example, was approximately 34% according to the National Health and Nutrition Examination Survey (NHANES) 2003-2006. (Ford, Gile, & Mokdad, 2004).

The prevalence of metabolic syndrome was found to increase with age, and more dramatically, with an increasing BMI. (Ervin, 2009; Kassi, Pervanidou, Kaltsas, & Chrousos, 2011). The NHANES studies found that women had a larger prevalence compared with men. (Ford, Gile, & Mokdad, 2004). In addition, additional genetic and environmental factors may also play a role in the pathogenesis of metabolic syndrome, thus the susceptibility and age of onset of the disease may vary among different individuals with a similar risk profiles. (Ordovas, 2007). Lifestyle modifications, including diet, exercise and weight reduction, are the currently preferred approaches to treat the disorder. Also, pharmacological treatment may bring additional preventive benefits or improvement to those having difficulty in reducing risk factors through these preventive measures. (Dandona, Aljada, Chaudhuri, Mohanty, & Garg, 2005).

2.1.2 Melatonin Treatment

Generally being classified as a lipophilic hormone, melatonin (*N*-acetyl-5-methoxytryptamine), is in fact, an important secretory product of pineal gland and a major component of the circadian system. (Reiter, 2003; Koziróg, et al., 2011). Suppressed by light and controlled by the suprachiasmatic nucleus (SCN), the concentration of melatonin remains at a low level during daytime and starts to rise at night during which it peaks around 2-4am. (Brzezinski, 1997). Melatonin acts as a highly effective oxidant, and has been recently found to be a free radical scavenger. (Tan, et al., 2003; Tan, et al., 2009). Several studies have shown that melatonin improves antioxidative enzyme activity and reduces oxidative damage. (Rodriguez, et al., 2004; Samantaray, et al., 2008; De Filipis, et al., 2008). It is thought that melatonin also plays roles in inhibiting oxidative

stress, lowering blood pressure, and potentially reducing inflammatory processes. (Kędziora-Kornatowska, et al., 2009; Simko & Pechanova, 2009; Jung, et al., 2010).

Reiter in 2003 reviewed several clinical aspects of melatonin with respect to its effects on the immune system, as well as its antioxidant effects. He evaluated the physiological and pharmacological benefits of melatonin administration in humans and animals and demonstrated non-toxicity of the molecule. (Reiter, 2003). Studies have reported that the dosage level and time of administration is critical for the efficacy of exogenous melatonin administration. (Dollins, 1993). In addition, Vural, van Munster and de Rooij conducted a systematic review on melatonin dose variation in 2014, and suggested that a lower supplementation level may be associated with better outcome. (Vural, van Munster, & de Rooij, 2014).

2.1.3 Melatonin Efficacy in Treating Metabolic Syndrome

The effects of melatonin on improving signs and symptoms of metabolic syndrome, as well as the overall syndrome itself, have been shown in various animal studies. (Puchalski, Green, & Rasmussen, 2003; Hoyos, et al., 2000). Both animal and human studies have documented melatonin's hypotensive effect, antioxidative potential, and improvement in lipid profile and reduction of pharmaceutical agent toxicity. (Koziróg, et al., 2011; Reiter, Tan, Sainz, & Mayo, 2002). Koziróg et al. conducted a study in 2011 to assess melatonin efficacy in treating metabolic syndrome and found that melatonin significantly reduced the

patients' BMI, systolic and diastolic blood pressure (SBP, DBP) and thiobarbituric acid reactive substrates after one-month and two-month administration. They also reported that melatonin significantly lowered SBP comparing the two visits at the end of the first and second month and further proposed that prolonged administration of melatonin may yield a better outcome. (Koziróg, et al., 2011). An eight-week, double-blind, randomized, placebo-controlled, parallel-group clinical trial done in 2014 examined the metabolic effects of melatonin among second-generation antipsychotics (SGA) patients, a group of people who tend to have metabolic disturbances. The results supported the therapeutic role of melatonin in preventing patients from experiencing SGA metabolic effects. (Romo-Nava, et al., 2014).

In a review paper by Nduhirabandi, du Toit, and Lochner et al, the dosage level of melatonin was proposed to be a particular challenge since both low and high doses seemed to be effective in experimental animals. The authors suggested that the therapeutic characteristics of melatonin need further investigation considering its potential benefits on preventing or reversing harmful effects of obesity and related metabolic disorder. (Nduhirabandi, du Toit, & Lochner, 2012). In addition, Srinivasan et al. recommended more studies to further prove the efficacy of melatonin with dosage level increases as well as longer duration of melatonin treatment. (Srinivasan, et al., 2013).

2.2 The Statistical Problem

2.2.1 Crossover Trials

As an alternative to the standard parallel or independent designs for clinical trials, the crossover study design has some unique characteristics that make it beneficial to clinical disciplines as well as the pharmaceutical industry. In crossover designs, each treatment is administered to every study subject at different times using a randomly assigned sequence. Each subject thus receives all treatments, and therefore can serve as his or her own control. Senn in 1993 defined the crossover design, as a design "in which subjects are given sequences of treatments with the object of studying differences between individual treatments (or sub-sequences of treatments)". (Senn, 1993).

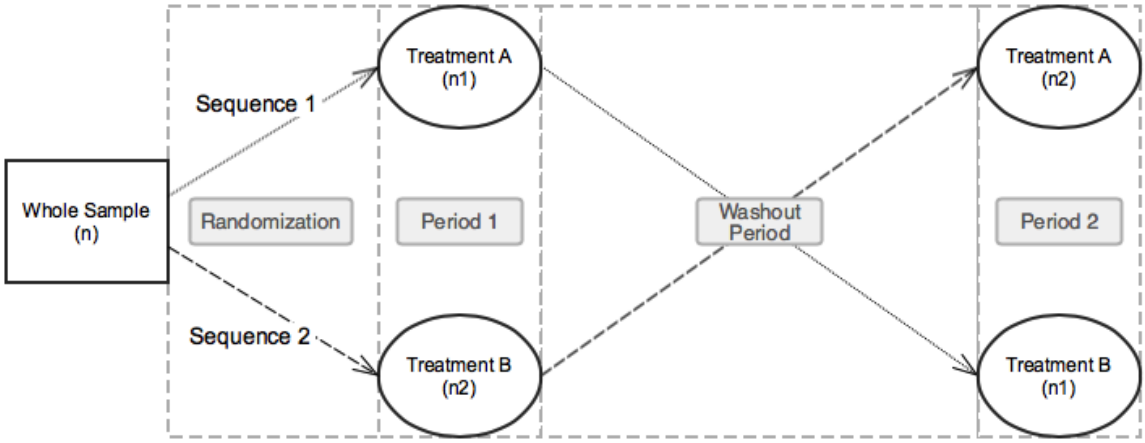
Crossover designs have been found especially common in neurology and psychiatry research, as well as of trials of pain treatment. (Wellek & Blettner, 2012). In addition, researchers propose that crossover trials are appropriate in chronic stable disease research due to the unique design features. (Feng & Ding, 2004). In that case, the objective of a crossover trial should be to potentially alleviate symptoms, rather than treating the cause of the disease. (Feng & Ding, 2004).

One primary advantage of the crossover trial design is that precision can be increased resulting from reducing the variance of the estimated difference in treatments. (Piatadosi, 1997; Kuehl, 1999). Since each patient is his/her own control, problems with between-patient variation and some confounding factors (e.g. age, sex, and ethnicity) can be eliminated. Also, a crossover trial needs fewer

patients compared to independent group designs because each study subject will receive all treatments – this may also improve patient recruitment and compliance. (Piatadosi, 1997). In addition, a much smaller sample size is needed to reach the same statistical power to test for treatment difference compared to a conventional parallel design, hence the designs are much more cost-effective. (Kuehl, 1999).

However, this type of design also has some potential disadvantages. A treatment administered in one period may have the possibility of influencing the patient's response in the following period. The effects of a previous treatment being carried over to the next treatment period are called carry-over effects. In order to rule out the remaining effects from a completed treatment, a washout phase must be placed between the two treatment periods and it has to last long enough to ensure that patient's response or disease status returns to baseline. Additionally, the treatment by period interaction is also a potential issue with crossover trial designs, in which case the effects of the treatment vary in different treatment periods. Unfortunately, in a simple two-treatment two-period (AB/BA) design (shown in Figure 1), this treatment by period interaction is indistinguishable from the carry-over effect. Moreover, as each subject will receive all treatments of interest, a longer time of commitment to the study in addition to an increased chance of suffering side effects may both contribute to a higher possibility of dropping out. (Piatadosi, 1997). As a result, the analysis will be affected by drop-outs because the data from those partially completed periods cannot be used directly. Thus use of this design may result in more data loss compared to some conventional designs. (Piatadosi, 1997).

Figure 1. Illustration of AB/BA crossover design



2.2.2 Illustration of Comparisons

A simple AB/BA crossover trial can be analyzed by using three two-sample t tests. In a very straightforward manner, the effects of period, treatment, and treatment-period interaction can be tested separately through one of the three tests. By calculating the mean of the responses in each period within each sequence group, we can obtain the differences of the mean responses in the two periods for each sequence, as well as the average of the sequence-specific responses in the two periods (illustrated in Table 1).

Table 1. Illustration of using two-sample t tests to analyze a simple placebo-controlled crossover trial

	Period 1	Period 2	Difference	Average
Sequence 1: Treatment – Placebo	$\overline{y_{11}}$	$\overline{y_{12}}$	$\overline{y_{11}} - \overline{y_{12}}$ [a]	$\frac{\overline{y_{11}}+\overline{y_{12}}}{2}$ [c]
Sequence 2: Placebo – Treatment	$\overline{y_{21}}$	$\overline{y_{22}}$	$\overline{y_{21}} - \overline{y_{22}}$ [b]	$\frac{\overline{y_{21}}+\overline{y_{22}}}{2}$ [d]

The outcome measure is denoted by y_{ij}, where $i = 1,2; j = 1.2$. The sequence is denoted by i, and the period is denoted by j. The carry-over effect can be tested by a simple two-sample t test for the averages. A two-sample t test for the differences can be used for testing for treatment effect.

The treatment effect can be tested by a two-sample t test to compare the mean differences between the two periods in the two sequence groups (a vs. b). (Altman, 1990). Furthermore, the treatment-period interaction, or the effect of carry-over, can be tested from a two-sample t test comparing the average of mean responses in the two periods between the two sequence groups (c vs. d). (Altman, 1990). In the absence of carry-over effects, the subject's mean response to the two treatments should be the same regardless of in which sequence the treatments were assigned to the subject. Rejection of this test then allows for the desired interpretation of the treatment effect.

2.2.3 The Grizzle Model

Analyzing a crossover trial is more complex compared to parallel-groups designs. Using a simple pre/post comparison or paired t test without examining the carry-over effects beforehand will result in errors and provide false estimates. Grizzle (1965) developed a model that has the ability to test the effects of treatment, period and carry-over simultaneously for a simple AB/BA crossover trial:

$$y_{ijk} = \mu + b_{ij} + \pi_k + \phi_l + \lambda_l + \epsilon_{ijk}$$

$$i = 1,2; j = 1,2, \dots, n_i; k = 1,2; l = 1,2;$$

where μ is the overall mean, b_{ij} is the random effect of the j^{th} subject within the i^{th} sequence, π_k is the k^{th} period effect, ϕ_l is the direct effect of the l^{th} treatment, λ_l is

the carry-over effect of the l^{th} treatment, and ϵ_{ijk} is the random error. (Grizzle, 1965). In order to perform the hypothesis testing for direct treatment effects, carry-over effects and period effects, Grizzle asserted that b_{ij} and ϵ_{ijk} must be assumed to be independent and normally distributed with mean 0 and variance σ_{ij}^2 and σ_e^2, respectively. Therefore, observations from different subjects are considered independent.

In an AB/BA trial, the hypothesis of absent carry-over effects (Ho: $\lambda_1 = \lambda_2$) can be tested by using the sum of the observations in the two periods on the same subject:

$$y_{ij1} + y_{ij2} = 2(\mu + b_{ij}) + (\pi_1 + \pi_2) + (\phi_1 + \phi_2) + \lambda_l + \epsilon_{ij1} + \epsilon_{ij2}$$

where λ_l reflects the carry-over effect of the l^{th} treatment following the *AB* treatment sequence. (Feng & Ding, 2004). If no carry-over effect is present, then the hypothesis of no direct treatment effects (Ho: $\phi_1 = \phi_2$) can be tested by using the difference between the observations in the two periods on the same subject, denoted as follows:

$$y_{ij1} - y_{ij2} = (\pi_1 - \pi_2) + (-1)^{i+1}(\phi_1 - \phi_2) + (\epsilon_{ij1} - \epsilon_{ij2})$$

where $(-1)^{i+1} = 1$ if the subject follows an AB treatment sequence and $(-1)^{i+1} = -1$ if follows a BA treatment sequence. (Feng & Ding, 2004).

However, if the carry-over effects cannot be ignored or eliminated, only data from the first period can be used and the analysis will be the same as the one for a conventional parallel design. If there are significant clinically important carry-over effects, it is possible that the subject's disease status or response is

affected permanently, in which case the crossover design is less efficient than a conventional independent-groups trial. (Piatadosi, 1997).

Nonetheless, some researchers consider these designs still pertinent under the circumstance that even if the carry-over effects are not completely eliminated, they are small enough compared to the treatment effects. (Brown, 1980). Others think if sequence, period and carry-over effects are all negligible in comparison to the direct treatment effects, the designs remain appropriate. (Jones & Kenward, 1989).

2.2.4 Principal Component Analysis

The use of principal component analysis (PCA) commonly serves for two purposes: data reduction and interpretation. (Johnson & Wichern, 2007). Given a set of correlated variables, a different set of uncorrelated variables called principal components, can be generated through some linear combinations of the originals to explain the information contained in the original set. The principal components then reduce the number of variables under consideration, but still represent most of the information. (Bartholomew, Steele, Moustaki, & Galbraith, 2008; Johnson & Wichern, 2007). In this way, the dataset can be reduced to a probably smaller number of uncorrelated variables and become easier for interpretation.

If the original set of the random variables is denoted as a random vector $\boldsymbol{X} = [X_1, X_2, \dots, X_p]'$, and the principal components for those variables denoted as $\boldsymbol{Y} = [Y_1, Y_2, \dots, Y_p]'$, then the linear combinations with real weights a_{ij} can be written as

$$Y_1 = \boldsymbol{a}_1'\boldsymbol{X} = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p$$

$$Y_2 = \boldsymbol{a}_2'\boldsymbol{X} = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p$$

$$\vdots$$

$$Y_p = \boldsymbol{a}_p'\boldsymbol{X} = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p$$

The main concept of PCA is to ensure that the total variance among those uncorrelated components $(Y_1, Y_2, \ldots, Y_p)$ remains equal to the total variance among the original variables $(X_1, X_2, \ldots, X_p)$, i.e., $\sum_{j=1}^{p} Var_{(Y_j)} = \sum_{j=1}^{p} Var_{(X_j)}$. (Bartholomew, Steele, Moustaki, & Galbraith, 2008; Johnson & Wichern, 2007). By deriving the principal components in a decreasing order based on their accountability of explaining the total variance, the first principal component will thus have a maximum variance and account for the largest proportion of the total variance. The second principal component then follows, explaining the maximum variance that the first fails to account for, and so forth. (Johnson & Wichern, 2007).

If the data are standardized with a mean of zero and a unit variance for each variable of interest, rather than using the covariance matrix from the original data, the correlation matrix can be used for the analysis. (Johnson & Wichern, 2007). Further, the variance of the component j will be λ_j, which can be obtained from the eigenvalues of the components $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p)$. (Bartholomew, Steele, Moustaki, & Galbraith, 2008). Therefore, the sum of the eigenvalues will equal to the total variance of the standardized variables p, and the proportion of the total variance explained by the k^{th} component then can be calculated as

$$\begin{pmatrix} proportion\ of\ total\ variance \\ explaned\ by\ the\ k^{th}\ component \end{pmatrix} = \frac{\lambda_k}{p}.$$

A commonly used criterion for determining the number of components to retain is the Kaiser criterion. (Kaiser, 1960). It selects the components to retain for further analysis as long as the eigenvalues are greater than 1, excluding all negative values.

The principal component loadings derived from the correlation matrix represent the correlation between the variable of interest and the calculated component. Thus these loadings are used for interpretation of the retained components.

Once the number of components is decided upon, a subject-specific score for each component can be calculated for use in further analysis. The component score for the *j*th component, for example, is equal to

$$y_j = a_{1j}x_1 + a_{2j}x_2 + \cdots + a_{pj}x_p$$

where *x*'s are all standardized to have a mean of zero and a unit variance and the variance for y_j is λ_j. (Johnson & Wichern, 2007). However, if the data are not standardized beforehand, the component scores can be standardized instead to have a variance of 1, and the formula for the standardized *j*th component score will thus become

$$\widetilde{y_j} = \widetilde{a_{1j}}x_1 + \widetilde{a_{2j}}x_2 + \cdots + \widetilde{a_{pj}}x_p$$

where $\widetilde{y_j} = y_j/\sqrt{\lambda_j}$ and $\widetilde{a_{ij}} = a_{ij}/\sqrt{\lambda_j}$ for the *j*th component on the subject *i*. (Johnson & Wichern, 2007). Individual's component scores thus provide an index of the individual's value for each of the retained components. Those

calculated values may then be used as potential predictor variables in the future analyses.

The implementation of PCA sometimes will reflect an unnoticeable link between the dependent and independent variables if they are correlated and insignificant results are shown in the previous analyses. (Bartholomew, Steele, Moustaki, & Galbraith, 2008). Therefore, it is beneficial for further investigations on some potential relationships as well as for a better interpretation of the results.

Chapter 3

Results

By fitting the data using the Grizzle model, the effects of carry-over and treatment were examined for every component of metabolic syndrome, including waist circumference, triglycerides, HDL cholesterol, fasting plasma glucose (FPG), systolic blood pressure (SBP), and diastolic blood pressure (DBP); as well as for some extra measures, including LDL cholesterol, total cholesterol, HsCRP, HgbA1C and a Pittsburgh Sleep Quality Index (PSQI) global score.

As shown in Table 2, the mean responses from pre-treatment and post-treatment were calculated for each measure in the two treatment groups. At a descriptive level, both waist circumference and HgbA1C level had slightly lower pre- and post-treatment levels on average in the placebo group compared to the melatonin group. The pre/post average levels for LDL, HDL and total cholesterol, on the contrary, were lower comparing melatonin vs. placebo groups. In addition, triglycerides, HsCRP, SBP, DBP, and PSQI global score all shared a common pattern: higher pre-treatment average but lower post-treatment average in the melatonin group. The post mean level of HsCRP in the placebo group, in particular, was much higher compared to the pre mean level in the same group, as well as than both of the pre and post means for the melatonin group. Instead, the fasting plasma glucose level had an opposite pattern – in comparison with the placebo group, the average was lower in pre-melatonin group, but the average in

post-melatonin group was higher. Moreover, the average levels of HgbA1C were almost constant before and after treatment in both groups (F=0.04, p=0.84).

Overall, only PSQI global score showed a significant carry-over effect (F=6.29, p=0.02), however there was no treatment effect. Given there were no other carry-over effects detected, the treatment effects can be interpreted as valid. The only significant effect of treatment detected was for the average level of SBP comparing melatonin with placebo (F=6.86, p=0.01). As opposed to an increasing average SBP within the placebo group (pre vs. post: 122.34±17.29 vs. 126.96±15.10), the melatonin group showed a decreasing average SBP after the 10-week supplementation (pre vs. post: 126.77±16.29 vs. 123.45±15.24). However, no treatment effects were significant for the other measures.

The correlation-based principal component analysis was used to determine if a linear combination of the 5 main components of the metabolic syndrome might combine in some way that differed from the univariate effects. Considering SBP and DBP were highly correlated, DBP was excluded from the PCA analysis thus the following variables were used: waist circumference, triglycerides, HDL cholesterol, fasting plasma glucose, and SBP average clinic value. Since we were interested in comparing the mean post level vs. pre level for each treatment, the principal components were derived on the difference scores from pre- to post-treatment for each variable of interest. By using the Kaiser's eigenvalue-one criterion (setting the critical eigenvalue at 1) together with the principal components extraction, the results showed that the variances explained by the first two principal components (y_1 and y_2) were 30.16% and 23.70%, respectively. Therefore, approximate 54% of the total variance among the five

components could be explained by these two retained principal components. From the factor pattern shown in Table 3, the linear combinations of the five components can be written as below using the factor loadings:

$$y_1 = 0.73 * \Delta_{waist} + 0.72 * \Delta_{trig} - 0.31 * \Delta_{HDL} + 0.60 * \Delta_{FPG} + 0.02 * \Delta_{\overline{SBP}}$$

$$y_2 = -0.18 * \Delta_{waist} + 0.05 * \Delta_{trig} + 0.76 * \Delta_{HDL} + 0.58 * \Delta_{FPG} - 0.48 * \Delta_{\overline{SBP}}$$

($\Delta = post - pre$; $waist$: waist circumference; $trig$: triglycerides; HDL: HDL cholesterol; FPG: fasting plasma glucose; $\overline{SBP}$: systolic blood pressure clinic average)

While the loadings for change in FPG were nearly evenly distributed on the two components, it is clear that the first component reflects mainly changes of triglycerides and waist circumference, while the second component represents mainly the changes in HDL and SBP average.

Another Grizzle model was then fitted using the above two principal components. Carry-over effects were not detected for either of the components, making the treatment effects interpretable (Table 4). However, a significant treatment effect was found on the second principal component (F=4.70, p=0.04), indicating that the changes in SBP and HDL were more in the positive range for the treatment group. Examination of the raw data indicated that a difference of 1 standard deviation towards the positive (score=1.0) was associated with a decrease in SBP of 7.6 units, and a corresponding increase in HDL of 9.1 units.

Table 2. Results from the Grizzle model

Measure	Melatonin		Placebo		Carry-over Effect		Treatment Effect	
	Pre Mean (±SD)	Post Mean (±SD)	Pre Mean (±SD)	Post Mean (±SD)	Test Statistics	P-value	Test Statistics	P-value
Waist circumference, cm	108.82 (±12.18)	108.06 (±11.32)	106.84 (±10.53)	107.99 (±10.58)	1.07	0.31	2.17	0.15
Triglycerides, mg/dL	207.19 (±271.86)	142.57 (±74.13)	175.27 (±172.14)	172.37 (±216.92)	1.94	0.17	2.00	0.17
HDL cholesterol, mg/dL	40.84 (±7.60)	40.95 (±8.35)	42.51 (±10.09)	41.51 (±8.33)	0.04	0.85	0.29	0.59
Fasting plasma glucose, mg/dL	103.41 (±16.46)	102.89 (±17.05)	104.58 (±13.74)	101.94 (±10.23)	3.77	0.06	1.17	0.29
SBP average clinic value, mmHg	126.77 (±16.29)	123.45 (±15.24)	122.34 (±17.29)	126.96 (±15.10)	0.83	0.37	6.86	0.01
DBP average clinic value, mmHg	76.60 (±11.31)	74.96 (±10.81)	75.31 (±10.33)	76.36 (±10.01)	0.19	0.66	2.27	0.14
Extra Measures								
LDL cholesterol, mg/dL	112.64 (±32.25)	108.46 (±33.22)	112.66 (±31.24)	114.59 (±35.27)	1.15	0.29	0.34	0.56
Total cholesterol, mg/dL	185.41 (±57.02)	174.53 (±41.34)	189.45 (±63.12)	186.95 (±58.84)	0.10	0.75	0.69	0.41
HsCRP, mg/L	3.59 (±4.04)	3.72 (±4.54)	3.58 (±3.85)	15.47 (±71.53)	1.23	0.27	1.23	0.27
HgbA1C, %	5.89 (±0.38)	5.87 (±0.43)	5.89 (±0.45)	5.85 (±0.34)	0.91	0.35	0.04	0.84
PSQI global score	6.60 (±3.51)	5.46 (±3.26)	6.34 (±3.47)	6.11 (±3.08)	6.29	0.02	1.55	0.22

DBP=diastolic blood pressure; SBP=systolic blood pressure; SD=standard deviation.

Table 3. Factor pattern from principal component analysis on a correlation matrix

	Component 1	**Component 2**
Change of waist circumference, cm	0.73	-0.18
Change of triglycerides, mg/dL	0.72	0.05
Change of HDL cholesterol, mg/dL	-0.31	0.76
Change of fasting plasma glucose, mg/dL	0.60	0.58
Change of SBP average clinic value, mmHg	0.02	-0.48

Table 4. Results from Grizzle model using factors from principal component analysis

	Melatonin post - pre Mean (±SD)	**Placebo** post - pre Mean (±SD)	**Carry-over Effect** Test Statistics	p-value	**Treatment Effect** Test Statistics	p-value
Component 1	-0.13 (±1.23)	0.12 (±0.72)	3.41	0.07	1.00	0.32
Component 2	0.29 (±0.81)	-0.27 (±1.09)	0.04	0.84	4.70	0.04

Chapter 4

Conclusions

Given that no significant carry-over effect was found among most of the measures of interest (except for the PSQI global score), the crossover design of this phase II clinical trial was appropriate. This further demonstrated that the results were not biased from the sequence of the treatment. Since a significant carry-over effect was detected on the PSQI global score, the data from the second study period should be discarded, and only the data from the first period should be analyzed in the way that it was a traditional parallel independent two-group trial. Nonetheless, the significant treatment effect on clinic SBP suggests a promising benefit of melatonin supplementation in treating patients with metabolic syndrome for the purpose of SBP improvement.

The implementation of principal component analysis provided us a deeper understanding of the five main components of metabolic syndrome. Since neither of the principal components was associated with a significant carry-over effect, this once again confirmed the validity of the crossover design for this study. In addition, based on the loadings of the two retained components, the first principal component may be better at explaining the positive relationships among changes in triglycerides and waist circumference, than each variable separately. The second component, on the other hand, can be viewed as being associated with an increase in HDL with a corresponding decrease in SBP. Also, the second component was associated with a significant treatment effect, which

corresponded with the conclusion above that the treatment effect of melatonin on clinic SBP was significant. Considering the possible underlying correlations among all of the five main components, the principal component analysis provided a multivariate solution to this crossover trial analysis by taking the linear combinations among the components into consideration. The potential advantage of melatonin in improving HDL became evident when accounting for the linear combinations of the components. Although the Grizzle model for this type of design considers each outcome variable independently, our results suggest that it may also be helpful to consider the possible linear combinations of the outcome measures when making conclusions or designing future studies.

Overall, this double-blind, placebo-controlled, crossover, phase II randomized clinical trial suggested a modest role of melatonin supplementation in the improvement of metabolic syndrome. Further investigations, such as large-scale parallel controlled clinical trials, may be beneficial to examine the efficacy of melatonin supplementation in more depth, for example, determining optimal dosage for treating metabolic syndrome.

Bibliography

Altman, D.G. (1990). *Practical statistics for medical research*. London/Florida: Chapman & Hall/CRC.

Bartholomew, D.J., Steele, F., Moustaki, I., & Galbraith, J. (2008). *Analysis of multivariate social science data*. London/Florida: Chapman and Hall/CRC.

Brown, B.W., Jr. (1980). The cross-over experiment for clinical trials. *Biometrics, 36*: 69-79.

Brzezinski, A. (1997). Melatonin in humans. *N Engl J Med, 336*(3): 186-195.

Buscemi, N., Vandermeer, B., Hooton, N., Pandya, R., Tjosvold, L., Hartling, L., … Baker, G. (2006). Efficacy and safety of exogenous melatonin for secondary sleep disorders and sleep disorders accompanying sleep restriction: meta-analysis. *BMJ, 332*: 385-393.

Dandona, P., Aljada, A., Chaudhuri, A., Mohanty, P., & Garg, R. (2005). Metabolic syndrome: a comprehensive perspective based on interactions between obesity, diabetes, and inflammation. *Circulation, 111*(11): 1448-1454.

De Filipis, D., Ivonne, T., Esposito, G., Steardo, L., Amold, G.H., Paul, A.P., … De Winter Benedicte.Y. (2008). Melatonin reverses lipopolysaccharide-induced gastro-intestinal mobility disturbances through inhibition of oxidative stress. *J Pineal Res, 44*: 45–51.

do Carmo, I., Dos Santos, O., Camolas, J., Vieira, J., Carreira, M., Medina, L., … Galvão-Teles, A. (2008). Overweight and obesity in Portugal: national prevalence in 2003-2005. *Obes Rev, 9*: 11-19.

Dollins, A.B., Lynch, H.J., Wurtman, R.J., Dneg, M.H., Kischka, K.U., Glearson, R.E., & Lieberman, H.R. (1993). Effect of pharmacological daytime doses of melatonin on human mood and performance. *Psycho-pharmacology (Berl), 112*(4): 490-496.

Ervin, R.B. (2009). Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003-2006. *Natl Health Stat Report, 13*: 1-7.

Feng, W.W., & Ding, D. (2004, May 27). SAS@ application in 2*2 crossover clinical trial. Retrieved from

http://www.lexjansen.com/pharmasug/2004/statisticspharmacokinetics/sp02.pdf

Ford, E.S. (2005). Risks for all-cause mortality, cardiovascular disease, and diabetes associated with the metabolic syndrome: a summary of the evidence. *Diabetes Care, 28*(7): 1769-1778.

Ford, E.S., Giles, W.H., & Mokdad, A.H. (2004). Increasing prevalence of the metabolic syndrome among U.S. adults. *Diabetes Care, 27*: 2444-2449.

Goyal, A., Terry, P.D., Superak, H.M., Nell-Dybdahl, C.L., Chowdhury, R., Phillips, L.S., & Kutner, M.H. (2014). Melatonin supplementation to treat the metabolic syndrome: a randomized controlled trial. *Diabetic & Metabolic Syndrome, 6*: 124.

Grizzle, J.E. (1965). The two-period change-over design and its use in clinical trials. *Biometrics, 21*(2): 467-480.

Grundy, S.M., Brewer, J.H.B., Cleeman, J.I., Smith, J.S.C., & Lenfant, C. (2004). Definition of the metabolic syndrome: report of the National Heart, Lung, and Blood Institute/American Heart Association conference on scientific issues relted to the definition. *Circulation*, *109*: 433-438.

Hillier, T.A., Fagot-Campagna, A., Eschwege, E., Vol, S., Cailleau, M., & Balkau, B. (2006). Weight change and changes in the metabolic syndrome as the French population moves towards overweight the S.E.S.I.R. cohort. *Epidemiol, 35*: 190-196.

Hollman, G., & Kristenson, M. (2006). The prevalence of the metabolic syndrome and its risk factors in a middle-ages Swedish population – mainly a function of overweight? *Eur J Cardiovasc Nurs, 7*: 21-26.

Hoyos, M., Guerrero, J.M., Perez-Cano, R., Olivan, J., Fabiani, F., Garcia-Perganeda, A. & Osuna, C. (2000). Serum cholesterol and lipid peroxidation are decreased by melatonin in diet-induced hypercholesterolemic rats. *J Pineal Res, 28*: 150–155.

Johnson, R.A., & Wichern, D.W. (2007). *Applied multivariate statistical analysis*. Upper Saddle River, NJ: Pearson.

Jone, B., & Kenward, M.G. (1989). *Design and analysis of cross-over trials*. London: Chapman and Hall.

Jung, K.H., Hong, S.W., Zheng, H.M., Lee, H.S., Lee, H., Lee, D.H., … Hong, S.S. (2010). Melatonin ameliorates cerulein-induced pancreatitis by the modulation of nuclear erythroid 2-related factor 2 and nuclear factor kappa B in rats. *J Pineal Res, 48*: 239–250.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*: 141-151.

Kassi, E., Pervanidou, P., Kaltsas G., & Chrousos, G. (2011). Metabolic syndrome: definitions and controversies. *BMC Medicine, 9*: 48.

Kędziora-Kornatowska, K., Szewczyk-Golec, K., Koza-kiewicz, M., Pawluk, H., Czuczejko, J., Kornatowski, T., … Kedziora, J. (2009). Melatonin improves oxidative stress parameters measured in the blood of elderly type 2 diabetic patients. *J Pineal Res, 46*: 333–337.

Koziróg, M., Poliwczak, A.R., Duchnowicz, P., Koter-Michalak, M., Sikora, J., & Broncel, M. (2011). Melatonin treatment improves blood pressure, lipid profile, and parameters of oxidative stress in patients with metabolic syndrome. *J Pineal Res, 50*(3): 261-266.

Kuehl, R.O. (1999). *Design of experiments: statistical principles of research design and analysis*. Pacific Grove, CA: Duxbury Press.

Nduhirabandi, F., du Toit, E.F., & Lochner, A. (2012). Melatonin and the metabolic syndrome: a tool for effective therapy in obesity-associated abnormalities? *Acta Ohysiol, 205*(2): 209-223.

Ordovas, J.M. (2007). Genetic links between diabetes mellitus and coronary atherosclerosis. *Current Atherosclerosis Reports, 9*(3): 204-210.

Piatadosi, S. (1997). *Clinical trials: a methodologic perspective*. New York, NY: Wiley-Interscience.

Puchalski, S.S., Green, J.N., & Rasmussen, D.D. (2003). Melatonin effect on rat body weight regulation in response to high-fat diet at middle age. *Endocrine, 21*: 163–167.

Reiter, R.J. (2003). Melatonin: clinical relevance. *Best Pract & Res Clin Endocrinol Metab, 17*(2): 273-285.

Reiter, R.J., Tan, D.X., Sainz, R.M., & Mayo, J.C. (2002). Melatonin: reducing the toxicity and increasing the efficiency of drugs. *J Pharm Pharmacol, 54*: 1299–1321.

Rodriguez, C., Mayo, J.C., Sainz, R.M., Antolin, I., Herrera, F., Martin, V., & Reiter, R.J. (2004). Regulation of antioxidant enzymes: a significant role for melatonin. *J Pineal Res, 36*: 1–9.

Romo-Nava, F., Alvarez-Icaza González, D., Fresán-Orellana, A., Saracco Alvarez, R., Becerra-Palars, C., Moreno, J., … Buijs, R.M. (2014). Melatonin

attenuates antipsychotic metabolic effects: an eight-week randomized, double-blind, parallel-group, placebo-controlled clinical trial. *Bipolar Disord, 16*: 410–421.

Samantaray, S., Sribnick, E.A., Das, A., Knaryan, V.H., Matzelle, D.D., Yallapragada, A.V., … Banik, N.L. (2008). Melatonin attenuates calpain up regulation, axonal damage and neuronal death in spinal cord injury in rats. *J Pineal Res, 44*: 348–357.

Senn, S. (1993). *Cross-Over trials in clinical research*. Chichester: John Wiley & Sons.

Simko, F., & Pechanova, O. (2009). Potential roles of melatonin and chronotherapy among the new trends in hypertension treatment. *J Pineal Res, 47*: 127–133.

Srinivasan, V., Ohta, Y., Espino, J., Pariente, J.A., Rodríguez, A.B., Mohamed, M., & Zakaria, R. (2013). Metabolic syndrome, its pathophysiology and the role of melatonin. *Recent Pat Endocr Metab Immune Drug Discov, 7*(1): 11-25.

Tan, D.X., Machester, L.C., Hardeland, R., Lopez-Burillo, S., Mayo, J.C., Sainz, R.M., & Reiter, R.J. (2003). Melatonin: a hormone, a tissue factor, an autocoid, a paracoid, and an antioxidant vitamin. *J Pineal Res, 34*(1): 75-78.

Tan, D.X., Reiter, R.J., Machester, L.C., Yan, M.T., El-Sawi, M., Sainz, R.M., … Hardeland, R. (2002). Chemical and physical properties and potential mechanisms: melatonin as a broad spectrum antioxidant and free radical scavenger. *Curr Top Med Chem, 2*(2): 181-197.

Vural, E.M.S., van Munster, B.C., & de Rooij, S.E. (2014). Optimal dosages for melatonin supplementation therapy in older adults: a systematic review of current literature. *Drugs Aging, 31*: 441-451.

Wellek, S., & Blettner, M. (2012). On the proper use of the crossover design in clinical trials. *Dtsch Arztebl Int, 109*(15): 276-281.

Appendix

SAS code

```
*************************************************************************;
*
* Name: Danni Liu
* Data: Thesis Data (PSQI & Lab values)
* Purpose: Cleaning the data and Running Analyses
* Created: 01/23/2015
* Edited:  04/05/2015
*
*************************************************************************;

libname raw "H:\Thesis\Analysis";

* Importing the raw lab data (one observation for each subject);
PROC IMPORT OUT= RAW.lab
            DATATABLE= "F3: Lab_Values"
            DBMS=ACCESSCS REPLACE;
              DATABASE="H:\Thesis\Analysis\Melatonin_Metsyndrome
access_HbA1c renamed.mdb";
              SCANMEMO=YES;
              USEDATE=NO;
              SCANTIME=YES;
RUN;

* Subsetting the data by visit and renaming the variables;
/* Visit 1*/
data v1;
      set raw.lab (keep=PID FPG_V1 SBP_1L_V1    SBP_2L_V1   SBP_1R_V1
      SBP_2R_V1   DBP_1L_V1   DBP_2L_V1   DBP_1R_V1
                                DBP_2R_V1 HbA1c_V1    Total_Chol_V1
      HDL_V1      LDL_V1      Trig_V1     Waist_V1    hsCRP_V1);
      rename FPG_V1=FPG SBP_1L_V1=SBP_1L  SBP_2L_V1=SBP_2L
SBP_1R_V1=SBP_1R SBP_2R_V1=SBP_2R
               DBP_1L_V1=DBP_1L     DBP_2L_V1=DBP_2L DBP_1R_V1=DBP_1R
DBP_2R_V1=DBP_2R
               HbA1c_V1=HbA1c Total_Chol_V1=Total_Chol HDL_V1=HDL
LDL_V1=LDL
               Trig_V1=Trig Waist_V1=Waist hsCRP_V1=hsCRP;
      visit=1;
run;
/* Visit 3 */
data v3;
      set raw.lab (keep=PID SBP_1L_V3     SBP_2L_V3   SBP_1R_V3
      SBP_2R_V3   DBP_1L_V3   DBP_2L_V3   DBP_1R_V3
                                DBP_2R_V3 FPG_V3      HbA1c_V3
      Total_Chol_V3     HDL_V3      LDL_V3      Trig_V3     Waist_V3
      Weight_V3   hsCRP_V3);
      rename SBP_1L_V3=SBP_1L SBP_2L_V3=SBP_2L  SBP_1R_V3=SBP_1R
      SBP_2R_V3=SBP_2R  DBP_1L_V3=DBP_1L  DBP_2L_V3=DBP_2L
```

```
                DBP_1R_V3=DBP_1R DBP_2R_V3=DBP_2R       FPG_V3=FPG
        HbA1c_V3=HbA1c    Total_Chol_V3=Total_Chol       HDL_V3=HDL
        LDL_V3=LDL
                Trig_V3=Trig    Waist_V3=Waist    Weight_V3=Weight
        hsCRP_V3=hsCRP;
        visit=3;

run;
/* Visit 5 */
data v5;
        set raw.lab (keep=PID SBP_1L_V5     SBP_2L_V5    SBP_1R_V5
        SBP_2R_V5    DBP_1L_V5    DBP_2L_V5    DBP_1R_V5    DBP_2R_V5
                                    FPG_V5     HbA1c_V5     Total_Chol_V5
        HDL_V5       LDL_V5       Trig_V5      Waist_V5     Weight_V5
        hsCRP_V5);
        rename SBP_1L_V5=SBP_1L SBP_2L_V5=SBP_2L  SBP_1R_V5=SBP_1R
        SBP_2R_V5=SBP_2R  DBP_1L_V5=DBP_1L  DBP_2L_V5=DBP_2L
                DBP_1R_V5=DBP_1R     DBP_2R_V5=DBP_2R  FPG_V5=FPG
        HbA1c_V5=HbA1c    Total_Chol_V5=Total_Chol       HDL_V5=HDL
        LDL_V5=LDL
                Trig_V5=Trig    Waist_V5=Waist    Weight_V5=Weight
        hsCRP_V5=hsCRP;
        visit=5;
run;
/* Visit 7 */
data v7;
        set raw.lab (keep=PID SBP_1L_V7     SBP_2L_V7    SBP_1R_V7
        SBP_2R_V7    DBP_1L_V7    DBP_2L_V7    DBP_1R_V7    DBP_2R_V7
                                    FPG_V7     HbA1c_V7     Total_Chol_V7
        HDL_V7       LDL_V7       Trig_V7      Waist_V7     Weight_V7
        hsCRP_V7);
        rename  SBP_1L_V7=SBP_1L SBP_2L_V7=SBP_2L SBP_1R_V7=SBP_1R
        SBP_2R_V7=SBP_2R  DBP_1L_V7=DBP_1L  DBP_2L_V7=DBP_2L
                    DBP_1R_V7=DBP_1R DBP_2R_V7=DBP_2R FPG_V7=FPG
HbA1c_V7=HbA1c    Total_Chol_V7=Total_Chol       HDL_V7=HDL  LDL_V7=LDL
                    Trig_V7=Trig      Waist_V7=Waist    Weight_V7=Weight
        hsCRP_V7=hsCRP;
        visit=7;
run;

* Merging the subsets together and labeling the variables (four
observations per subject);
data lab_comb;
        set V1 V3 V5 V7;
        label SBP_1L="SBP_1L" SBP_2L="SBP_2L" SBP_1R="SBP_1R"
SBP_2R="SBP_2R"
                DBP_1L="DBP_1L" DBP_2L="DBP_2L" DBP_1R="DBP_1R"
DBP_2R="DBP_2R"
                FPG="FPG" HbA1c="HbA1c" Total_Chol="Total_Chol" HDL="HDL"
LDL="LDL"
                Trig="Trig" Waist="Waist" Weight="Weight" hsCRP="hsCRP"
                visit="visit" PID="PID";
        SBP_avg=mean(SBP_1L, SBP_2L, SBP_1R, SBP_2R); /* calculating SBP
clinic average value */
        DBP_avg=mean(DBP_1L, DBP_2L, DBP_1R, DBP_2R); /* calculating DBP
clinic average value */
        label SBP_avg="SBP average"
```

```
              DBP_avg="DBP average";
run;

* Sorting the dataset by subject's ID and visit number;
proc sort data=lab_comb;
      by PID visit;
run;

*proc contents data=lab_comb;
*run;

* Deleting those observations whose all lab values were missing;
data lab_comb_del;
      set   lab_comb;

      array vars(17) SBP_1L SBP_2L SBP_1R SBP_2R DBP_1L DBP_2L DBP_1R
DBP_2R
                           FPG HbA1c Total_Chol HDL LDL Trig Waist
Weight hsCRP;

      numMissing = cmiss(of vars[*]);
      if numMissing = 17 then delete; /* deleting those all-blank
observations */
run;

* Creating a temporary randomization dataset from the raw PSQI data;
data randomz (keep= PID visit_number rand med period
                    rename=(visit_number=visit));
      set raw.PSQI;
run;

* Checking for randomization assignment in the lab data;
data lab_comb_new;
      merge randomz  lab_comb_del;
      by PID visit ;

      /* specifying period */
      if visit= 1 then period = 1;
            else if visit =3 then period = 1;
            else if visit =5 then period = 2;
            else if visit =7 then period = 2;

      /* treatment sequence assignment */
      if rand=1 then do;
            if period =1  then med= 1;
            else if period = 2 then med = 0;
      end;
      else if rand=2 then do;
            if period =1 then med= 0;
            else if period =2  then med = 1;
      end;

      drop SBP_1L  SBP_2L  SBP_1R  SBP_2R DBP_1L  DBP_2L  DBP_1R
DBP_2R; * Dropping raw SBP and DBP variables- using average values of
the original four for each instead;
run;
```

```
* Outputing final randomization information;
data randomz_update;
      set lab_comb_new;
      keep PID visit rand med period;
run;

/* Randomization for Period 1 */
data randomz_p1 ;
      set randomz_update;
      where period=1 and visit=1;
      drop visit;
run;

/* Randomization for Period 2 */
data randomz_p2 ;
      set randomz_update;
      where period=2 and visit =5;
      drop visit;
run;

/* Subsetting the updated lab data based on visit number */
data labv1 (keep=PID LDL HDL Total_Chol Trig hsCRP HbA1c visit SBP_avg
DBP_avg FPG Waist Weight
                  rename=(LDL=LDL1 HDL=HDL1 Total_Chol=Total_Chol1
Trig=Trig1 hsCRP=hsCRP1 HbA1c=HbA1c1
                              SBP_avg=SBP_avg1 DBP_avg=DBP_avg1
FPG=FPG1  Waist=Waist1 Weight=Weight1))

       labv3 (keep=PID LDL HDL Total_Chol Trig hsCRP HbA1c visit
SBP_avg DBP_avg FPG Waist Weight
                  rename=(LDL=LDL2 HDL=HDL2 Total_Chol=Total_Chol2
Trig=Trig2 hsCRP=hsCRP2 HbA1c=HbA1c2
                              SBP_avg=SBP_avg2 DBP_avg=DBP_avg2
FPG=FPG2  Waist=Waist2 Weight=Weight2))

       labv5 (keep=PID LDL HDL Total_Chol Trig hsCRP HbA1c visit
SBP_avg DBP_avg FPG Waist Weight
                  rename=(LDL=LDL1 HDL=HDL1 Total_Chol=Total_Chol1
Trig=Trig1 hsCRP=hsCRP1 HbA1c=HbA1c1
                            SBP_avg=SBP_avg1 DBP_avg=DBP_avg1 FPG=FPG1
Waist=Waist1 Weight=Weight1))


       labv7 (keep=PID LDL HDL Total_Chol Trig hsCRP HbA1c visit
SBP_avg DBP_avg FPG Waist Weight
                  rename=(LDL=LDL2 HDL=HDL2 Total_Chol=Total_Chol2
Trig=Trig2 hsCRP=hsCRP2 HbA1c=HbA1c2
                              SBP_avg=SBP_avg2 DBP_avg=DBP_avg2
FPG=FPG2  Waist=Waist2 Weight=Weight2));

      set lab_comb;
      if visit=1 then output labv1;
      else if visit=3 then output labv3;
      else if visit=5 then output labv5;
      else if visit=7 then output labv7;

run;
```

```
* Creating period-specific data after randomization and generating the
"change" variables (pre-post) for each lab measure;
/* Period 1 */
data labp1;
      merge randomz_p1(in=inp1) labv1(in=inv1) labv3(in=inv3);
      by PID;
      if inp1 and inv1 and inv3;
      drop visit;
      chLDL=LDL2-LDL1;
      chHDL=HDL2-HDL1;
      chTotal_Chol=Total_Chol2-Total_Chol1;
      chTrig=Trig2-Trig1;
      chhsCRP=hsCRP2-hsCRP1;
      chhbA1c=HbA1c2-HbA1c1;
      chSBP_avg=SBP_avg2-SBP_avg1;
      chDBP_avg=DBP_avg2-DBP_avg1;
      chFPG=FPG2-FPG1;
      chWaist=Waist2-Waist1;
      chWeight=Weight2-Weight1;
run;

/* Period 2 */
data labp2;
      merge  randomz_p2(in=inp2) labv5(in=inv5) labv7(in=inv7);
      by PID;
      if inp2 and inv5 and inv7;
      drop visit;
      chLDL=LDL2-LDL1;
      chHDL=HDL2-HDL1;
      chTotal_Chol=Total_Chol2-Total_Chol1;
      chTrig=Trig2-Trig1;
      chhsCRP=hsCRP2-hsCRP1;
      chhbA1c=HbA1c2-HbA1c1;
      chSBP_avg=SBP_avg2-SBP_avg1;
      chDBP_avg=DBP_avg2-DBP_avg1;
      chFPG=FPG2-FPG1;
      chWaist=Waist2-Waist1;
      chWeight=Weight2-Weight1;
run;

* Creating a permanent dataset by concatenating the data from the two
periods for each subject (two observations per subject);
data raw.lab_final;
      set labp1 labp2;
run;

* Sorting the dataset by subject's ID and period;
proc sort data=raw.lab_final;
      by PID period;
run;

* Descriptive statistics comparing pre vs. post in each treatment
group;
proc means data= raw.lab_final  N mean std maxdec=2;
      var Waist1 Trig1  HDL1 FPG1 SBP_avg1 DBP_avg1
            LDL1 Total_Chol1 hsCRP1  HbA1c1;
```

```
      var Waist2 Trig2  HDL2 FPG2 SBP_avg2 DBP_avg2
            LDL2 Total_Chol2 hsCRP2  HbA1c2;
      class med;
run;

* Pittsburgh Sleep Quality Index (PSQI);
* Importing the cleaned PSQI dataset (two observations per subject);
data PSQI_raw;
      set   raw.PSQI_tmp_cc;
run;

* Descriptive statistics for PSQI variables (pre vs. post in each
treatment group);
proc means data= PSQI_raw  N mean std maxdec=2;
      var PSQI1 PSQI2;
      class med;
run;



 ************************Fitting the Grizzle model*********************;

* Creating a temporary dataset that only contains the variables of
interest for fitting the Grizzle model;
data lab_final_measure;
      set raw.lab_final;
      keep PID period rand med chLDL chHDL chTotal_Chol chTrig chhsCRP
chhbA1c
            chSBP_avg chDBP_avg chFPG chWaist chWeight;
run;

* Fitting the Grizzle model with Proc Mixed statement (repeated
measures);
%macro Grizzle (dat,ch_var);
proc mixed data=&dat; /* use the temporary lab_final_measure dataset */
      class rand PID period med;
      model &ch_var=rand period med;
            repeated/type=cs sub=PID(rand)r;
run;
%mend;

options mprint mlogic symbolgen;
%Grizzle (lab_final_measure,chWaist);
%Grizzle (lab_final_measure,chTrig);
%Grizzle (lab_final_measure,chHDL);
%Grizzle (lab_final_measure,chFPG);
%Grizzle (lab_final_measure,chSBP_avg);
%Grizzle (lab_final_measure,chDBP_avg);
%Grizzle (lab_final_measure,chLDL);
%Grizzle (lab_final_measure,chTotal_Chol);
%Grizzle (lab_final_measure,chhsCRP);
%Grizzle (lab_final_measure,chhbA1c);
%Grizzle (PSQI_raw,chpsqi);
options nomprint nomlogic nosymbolgen;
```

```
**********Principal Component Analysis on Main MetS Components********;

proc factor data=lab_final_measure rotate=varimax;
      var chtrig chhdl chfpg chsbp_avg chwaist;
run;

proc factor data=lab_final_measure rotate=varimax score out=work.labfac
nfactors=2;
      var chtrig chhdl chfpg chsbp_avg chwaist;
run;

* Creating a permanent dataset with component loadings;
data raw.labfac;
      set work.labfac;
      grp2=.;
      if factor2 ge 1 then grp2=1;
      else if factor2 lt 1 then grp2=0;
run;

proc means data=raw.labfac N mean std maxdec=2;
      var chSBP_avg chHDL chFPG;
      class grp2;
run;

* Descriptive statistics of the two principal components;
proc means data=raw.labfac N mean std maxdec=2;
      var factor1 factor2;
      class med;
run;

* Fitting the Grizzle model with principal components;
proc mixed covtest;
      class rand pid period med;
      model factor1=rand period med/s;
            repeated /type=cs sub=pid(rand);
run;

proc mixed covtest;
      class rand pid period med;
      model factor2=rand period med/s;
            repeated /type=cs sub=pid(rand);
run;
```