Comparing the Statistical Power of Analysis of Covariance after Multiple Imputation and the Mixed Model in Testing the Treatment Effect for Pre-post Studies with Loss to Follow-up

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Wenna Xi, M.S.

Graduate Program in Public Health

The Ohio State University

2014

Master's Examination Committee:

Michael L. Pennell, Advisor

Rebecca R. Andridge

Abstract

Pre-post studies, where outcomes are measured both before and after an intervention, are common in biomedical research. When the outcomes at both pre- and post-test are completely observed, previous studies have shown that analysis of covariance (ANCOVA) is more powerful than the change score analysis in testing the treatment effect and therefore is usually recommended in analyzing pre-post studies. However, methods for analyzing pre-post studies with missing outcome values have not been compared. The goal of this study was to compare the power of two analysis methods in testing for a treatment effect when post-test values are missing: ANCOVA after multiple imputation (MI) and the mixed model. To do so, we analyzed data from a real study, the BePHIT study, and performed simulation studies. Four analysis methods were used to analyze the BePHIT and simulated data: ANCOVA after MI, ANCOVA using only complete cases (CC), the mixed model using all-available data, and the mixed model using complete cases. Simulation studies were conducted under various sample sizes, missingness rates, and missingness scenarios. In the analysis of the BePHIT data, ANCOVA after MI produced the smallest p-value for the test of a treatment effect. However, in the simulation studies, CC ANCOVA was generally the most powerful method. The simulation studies also showed that the power of ANCOVA after MI dropped the fastest when the percentage of missingness increased

and, for most scenarios, was the least powerful method when 50% of the post-test outcomes were missing.

Acknowledgments

This thesis would not have been completed without the help of many people. First, I would like to express my gratitude to my advisor, Dr. Michael Pennell, for the time he spent discussing the project and revising the writing of the thesis. Dr. Pennell came up with the idea of this project and guided me patiently through the whole research process. This project would not have been even started without him.

Besides my advisor, I would also like to thank my second reader, Dr. Rebecca Andridge. She also spent a lot of time discussing the project and shared her SAS codes for generating the data in simulations with me as well. This project would not look as good as it is without her suggestions and comments on multiple imputation and simulations.

My thanks also goes to Dr. Eloise Kaizar, the instructor of my missing data class. Her class notes were very helpful for me to understand the methods of missing data analysis and to write up the literature review part of the thesis.

Vita

2010 ..... B.S. Mathematics, Northwest Normal University

2012 ..... M.S. Biomathematics, University of Science and Technology of China

Fields of Study

Major Field: Public Health

Table of Contents

List of Tables

Table **Page**

List of Figures

Chapter 1: Introduction

In a pre-post study a treatment is evaluated by measuring responses both before and after the study for each participant in a treatment group and a control group. Pre-post study designs have been widely used in clinical trials, psychology, education, and sociology. For example, a pre-post design was used to investigate the effect of moderate-intensity exercise on self-rated sleep quality in older adults [1]. Another study examined the effectiveness of an SAT training program by comparing the change in students' SAT scores before and after the program to the change in the control group [2].

When there is complete follow-up, previous studies have shown that, in terms of testing the treatment effect, analysis of covariance (ANCOVA) is more powerful than a comparison of change scores [2–5]. However, in reality, missing data, in particular loss to follow-up, is very common in pre-post studies. With unbalanced sample sizes for pre- and post-test levels in each treatment group, regular ANCOVA or change score analysis cannot be conducted without dropping any subjects. Therefore the most straightforward method for handling missing values is to exclude all the subjects with missing data. This type of analysis is called the complete-case (CC) analysis. The CC analysis is usually not recommended, since it throws away some useful information collected for the study and does not follow the intent-to-treat principle for

clinical trials [6, 7]. Nowadays, one popular way to handle missing data is multiple imputation (MI) [6]. For instance, in pre-post studies, the missing follow-ups can be simulated based on a model of post-test levels predicted by the pre-test levels and some other related variables collected in the study. Missing data are simulated several times to generate several complete data sets. Regular analysis methods can then be carried out on each complete data set and the results are combined to account for the uncertainty introduced by the imputations [8]. Another approach often used for data with repeated measures is the mixed model, where all pre- and post-test values are treated as outcomes and are regressed over treatment and timepoint indicators. Mixed models allow for missing follow-ups, since the pre-test data can still be used to fit the model even if the corresponding post-test data are missing. Both MI and mixed models assume data to be missing at random (MAR), which means that the missingness depends only on observed variables not the data that would have been observed.

Many papers have been published which discuss the analysis methods for pre-post studies when data are completely observed [2–5, 9]. However, to our knowledge, there have been no papers comparing methods for handling missing post-test values. This study was proposed to fill in this gap. The goal of this study was to compare the power of the mixed model and ANCOVA after MI for pre-post studies when missing post-test is present. All other variables, including the pre-test values, are assumed to be fully observed. The two methods were first compared in the context of a real data example. The data came from the Behavior and Exercise for Physical Health Intervention (BePHIT) study. The change in time of a one-mile walk during the study was compared between the treatment group and the control group. Simulation

studies motivated by the BePHIT data were also conducted using various sample sizes, missingness rates, and missingness scenarios (i.e., MCAR and different MAR scenarios).

The rest of the thesis is organized as follows: Chapter 2 is a review of pre-post studies and their analyses. A discussion of missing data and statistical methods for handling loss to follow-up in pre-post studies is also presented. Chapter 3 includes the analysis of the BePHIT data. Chapter 4 includes the details of the simulation studies with a discussion of some important results. Chapter 5 is a discussion of the results of this study and some future research directions. Appendix A contains the results of additional simulation studies and Appendix B contains the SAS code used in the BePHIT analyses and the simulation studies.

Chapter 2: Pre-post Studies, Missing Data, and Their Analyses

2.1 Pre-post Studies

A pre-post study is a randomized controlled study where outcome values are measured both before and after the study. As opposed to treatment-control studies where the outcome variable is only measured once, pre-post studies allow investigators to account for the level of the outcome variable before the treatment is applied. Different from a one-group pre-post design, a treatment-control pre-post study controls for secular trends [10, 11]. In the BePHIT study, for instance, besides the intervention, the improvement of women's one-mile walk time may also be caused by some other factors, such as a national walking campaign, affecting the women during the same intervention period. A one-group pre-post study fails to consider these factors; however a treatment-control pre-post study accounts for secular trends by comparing the results from the treatment group to a control group observed over the same period of time.

If we let R be the randomization process, T be the treatment process, and $(Y_{\text{pre,t}}, Y_{\text{post,t}})$ and $(Y_{\text{pre,c}}, Y_{\text{post,c}})$ be the pre- and post-test measure of a treated and control participant, respectively, then a pre-post study design can be illustrated

by the following:

$$\begin{array}{ccccc} R & \longrightarrow & Y_{\text{pre,t}} & \stackrel{T}{\longrightarrow} & Y_{\text{post,t}} \\ R & \longrightarrow & Y_{\text{pre,c}} & \longrightarrow & Y_{\text{post,c}} \end{array}$$

If subjects are not randomly assigned to the treatment or the control group, the study is said to have a quasi-experimental pre-post design. For non-experimental data, the pre-post design can rule out the possibility that Y causes T and greatly reduce the possibility that some other variable causes both T and Y by testing the relationship between T and Y_{post} adjusting for Y_{pre} [2].

2.2 Analysis of Pre-post Studies

Many analysis approaches for pre-post studies have been discussed [2,3,5,10,12,13]. Probably the two most common analysis methods are the change score analysis and Analysis of Covariance (ANCOVA) [2]. We discuss these two methods and their statistical power in the following.

2.2.1 Change Score Analysis

The idea of the change score analysis is to first obtain the difference in outcome values before and after the experiment, and then regress the difference on the treatment assignment using the following model:

$$Y_i - X_i = \beta_0 + \beta_1 T_i + \varepsilon_i, \tag{2.1}$$

where Y_i is the post-test outcome level for subject i, X_i is the pre-test outcome level for subject i, T_i is the indicator variable for treatment assignment (defined as below), and ε_i is the error term for subject i. Note that $Y_i - X_i$ is the change score for subject i during the experiment. For person i, indicator variable T_i is defined as

$$T_i = \begin{cases} 1, & \text{if subject } i \text{ is in the treatment group;} \\ 0, & \text{if subject } i \text{ is in the control group.} \end{cases}$$

Finally, the change score analysis assumes $\varepsilon_i \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$.

In the model above, β_1 quantifies the effect of treatment assignment on change in outcome level from pre to post. Therefore, to test the treatment effect on the change scores, the null and alternative hypotheses are:

$$\begin{aligned} H_0 : &\quad \beta_1 = 0, \\ H_a : &\quad \beta_1 \neq 0. \end{aligned}$$

Note that since T_i is a binary variable, the change score test is equivalent to a two-sample t-test comparing the mean of $Y_i - X_i$ between treatment and control groups.

2.2.2 ANCOVA

Unlike the change score method, in ANCOVA, post-test value (Y_i) is treated as the outcome variable and pre-test value (X_i) is treated as a predictor. The ANCOVA model can be expressed as:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 X_i + \varepsilon_i, \tag{2.2}$$

where T_i and ε_i are as defined in the change-score model. ANCOVA assumes $\varepsilon_i \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$ and that pre-test values are measured without error [5]. No measurement error indicates that the pre-test values are the actual values for the subjects. This assumption holds for variables, such as weight and height, that can be measured precisely. However, it is often violated for self-reported measurements, and educational or psychological tests.

In the model above, β_1 is the effect of treatment assignment on the post-test scores adjusting for the pre-test scores. Therefore, the null and alternative hypotheses

$$\begin{aligned} H_0 : &\quad \beta_1 = 0, \\ H_a : &\quad \beta_1 \neq 0 \end{aligned}$$

are now testing the effect of treatment on the post-test scores controlling the pre-test scores.

2.2.3 Change Score Analysis vs. ANCOVA

Assumptions

Before comparing the two methods, it is worth mentioning that ANCOVA model (2.2) can also be written as

$$Y_i - X_i = \beta_0 + \beta_1 T_i + \beta_2^* X_i + \varepsilon_i, \tag{2.3}$$

where $\beta_2^* = \beta_2 - 1$ from (2.2) [2, 5]. Thus, ANCOVA can be viewed as an extension of the change score model (2.1) to include pre-test level X_i as a predictor. If we consider the change score model (2.1) as a reduced model of (2.3), this implies another assumption for the change score analysis: pre-test score, X_i, has no effect on change score, $Y_i - X_i$.

Also, in the change score model (2.1), we can move X_i to the right hand side and treat it as a covariate,

$$Y_i = \beta_0 + \beta_1 T_i + X_i + \varepsilon_i. \tag{2.4}$$

Comparing to the original ANCOVA model (2.2), the change score analysis additionally assumes that a one unit increase in the pre-test leads to one unit increase in the post-test, controlling for the treatment assignment. However, unlike the ANCOVA model, (2.4) does not assume zero measurement error for X_i.

Power

Statistical power is the probability of rejecting the null hypothesis when the null is false. In the language of probability, it can be written as

$$\text{Power} = P(\text{reject the null}|\text{null is false}).$$

The power of a test depends on the smallest difference the test could detect. The smaller the detectable difference is, the higher the chance to reject the null, and therefore the larger the statistical power of the test is. Hence, instead of comparing the power of the two methods, we can compare their minimum detectable differences.

Assume

1. $Var(X_i) = Var(Y_i) = \sigma^2$ regardless of the experiment group, and

2. number of subjects in each group is the same.

Under the assumption of normally distributed errors, it can be shown that the detectable difference for the change score analysis at type-I and type-II error rates of α and β is

$$\Delta_{CS} = \sqrt{\frac{4\sigma^2(1-\rho)(Z_{1-\alpha/2} + Z_{1-\beta})^2}{n}},$$

and the detectable difference for ANCOVA is:

$$\Delta_A = \sqrt{\frac{2\sigma^2(1-\rho^2)(Z_{1-\alpha/2} + Z_{1-\beta})^2}{n}},$$

where $\rho = Corr(X_i, Y_i)$, Z_x is the x quantile of the standard normal distribution, and n is the number of subjects in each experiment group. Therefore we have

$$\frac{\Delta_{CS}}{\Delta_A} = \sqrt{\frac{2}{1+\rho}}. \tag{2.5}$$

Assuming $0 \leq \rho \leq 1$,

$$1 \leq \frac{\Delta_{CS}}{\Delta_A} \leq \sqrt{2}.$$

Thus the detectable difference for ANCOVA is always less than or equal to the detectable difference for the change score analysis. Also, from (2.5) we can see that the discrepancy in power increases as the correlation between pre- and post-test decreases.

In general, based on the power comparison above, ANCOVA should be chosen over the change score analysis for pre-post studies. The power of ANCOVA and the change score analysis is the same when there is a perfect correlation between pre- and post-tests ($\rho = 1$). As the correlation decreases, ANCOVA becomes more powerful than the change score analysis. However, there is one exception to this recommendation. ANCOVA assumes that the pre-test values are obtained without measurement error. Since the change score analysis does not require this assumption, the change score analysis should be used when we know that the error for the pre-test measurement cannot be ignored.

2.3 Missing Data

Both the change score analysis and ANCOVA require complete follow-up of subjects. Unfortunately, missing post-test is very common in pre-post studies. Participants may drop out because they moved, are unsatisfied with their performance in the study, etc. Thus, it is often necessary to analyze incomplete data in pre-post studies.

The change score analysis and ANCOVA for completely observed data have two parallel methods for data with missing responses: the mixed model and ANCOVA following multiple imputation. However, to our knowledge, power analyses of these

two methods have not been considered. Before going into the details of a power comparison of these two methods, this section will provide a review of missing data and their analyses.

This section consists of two parts. In Section 2.3.1, notations and terminologies of missing data are introduced. In Section 2.3.2, discussions of complete-case analysis, mixed model, and multiple imputation (MI) are presented. Also, the regression method for MI available in the software package SAS (SAS Inc., Cary, NC), which will be used in the following chapters, is summarized.

2.3.1 Missing Data Mechanisms

To better illustrate the missing data mechanisms in mathematical form, we need to first define the following terms:

Y	Complete data matrix
Y_{obs}	Observed data
Y_{mis}	Missing data
M	Missing data matrix
ϕ	All parameters from the data generating process
$f(M, Y, \phi)$	Distribution of the complete data.

Suppose we are going to observe k variables for n individuals. The data can be written in a matrix form:

$$Y_{i\times j} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1k} \\ y_{21} & y_{22} & \cdots & y_{2k} \\ \vdots & & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nk} \end{pmatrix},$$

where y_{ij} is the value of the j-th variable for the i-th subject. This matrix Y is called a complete data matrix. Matrix Y consists of both observed data and missing data. If we use Y_{obs} to denote all observed data in Y, and use Y_{mis} to denote all missing data in Y, then Y can be separated into two parts:

$$Y = [Y_{\text{obs}}, Y_{\text{mis}}].$$

Note that Y_{obs} and Y_{mis} are not matrices. They were introduced only to simplify the notations and explanations.

For each complete data matrix, there is a corresponding missing data matrix. A missing data matrix is of the same dimension as its corresponding complete data matrix. It is introduced to better illustrate the structure of the missing data. The missing data matrix $M_{i\times j}$ contains elements m_{ij} defined as follows:

$$m_{ij} = \begin{cases} 1, & \text{if } y_{ij} \text{ is missing;} \\ 0, & \text{if } y_{ij} \text{ is observed.} \end{cases} \tag{2.6}$$

Furthermore, we use ϕ to denote all parameters in the data generating process, which includes parameters used to generate Y and M. Finally, we define $f(M, Y, \phi)$ as the joint distribution of the complete data.

When the data collection procedure is finished, we have full information about M and Y_{obs}. Since we are missing Y_{mis}, we will never know the values in Y_{mis}.

Missing Completely at Random (MCAR)

Missing completely at random (MCAR) means the missing data does not depend on anything in the data collecting process. To be more precise, MCAR is when missingness does not depend on anything in the Y matrix. The formal definition can be given as

$$f(M|Y, \phi) = f(M|\phi)$$

for all Y and ϕ. This also means M and Y are independent. One example of MCAR is missing blood pressure due to a broken sphygmomanometer [14].

Missing at Random (MAR)

We say a study has data missing at random (MAR), when missingness can only depend on observed values, but not those values that are missing (or would have been

observed). In probability sense, it means

$$f(M|Y,\phi) = f(M|Y_{\text{obs}}, Y_{\text{mis}}, \phi) = f(M|Y_{\text{obs}}, \phi)$$

for all Y_{mis} and ϕ. MCAR is a special case of MAR, since $f(M|\phi) = f(M|Y_{\text{obs}}, \phi)$ under the MCAR assumption. For example, if older people are more likely to receive blood pressure measurements, then the missingness of blood pressure measures is MAR since it depends on the age of the participants, which is observed in the study [14].

Not Missing at Random (NMAR)

When MAR does not hold, we say that data are not missing at random (NMAR). When data are NMAR, missingness depends on both observed and missing values. In this case, the distribution of missing data matrix $f(M|Y_{\text{obs}}, Y_{\text{mis}}, \phi)$ cannot be simplified. For example, people may be more likely to receive blood pressure measurements when they think their blood pressure is high. In this case, the missingness is NMAR since it is related to the values of both observed and missing blood pressures.

2.3.2 Statistical Methods for Missing Data

Commonly used methods for handling missing data include: complete-case (CC) analysis, weighting adjustments, imputation, and model-based methods [6, 15, 16]. The CC analysis is unbiased under MCAR, while the rest are valid under MAR. Models for NMAR data have been proposed which involve either *a priori* restrictions on the parameter space for ϕ or using an informative Bayesian prior distribution on ϕ [17].

There are three natural approaches to handling loss to follow-up in a pre-post study. First, one can remove the data for subjects with post-test missing and carry

out a complete-case (CC) analysis. The change score model can be generalized to a mixed model, in which unequal sample sizes for pre- and post-test levels is allowed. ANCOVA could be applied following imputation of the missing data. These three approaches are summarized below.

Complete-Case (CC) Analysis

The most straightforward way to deal with missing data is to delete all subjects with any missing observations, no matter if data are partially collected or not. The dataset obtained after deleting all subjects with missing values is called a complete-case dataset and the analysis of these data is called a complete-case (CC) analysis. The CC analysis is simple to conduct and unbiased under MCAR. In certain situations, it is also unbiased under MAR. For instance, when estimating the regression model of Y on $X_1, X_2, \ldots, X_n$ from data with Y being incompletely observed, the estimation conditions on the values of the X's. Thus, the CC analysis is unbiased if the missingness only depends on the X's but not the Y [6, 16]. As a result, the CC ANCOVA model is unbiased under MAR for pre-post studies with loss to follow-up.

Another disadvantage of the CC analysis is that, information is thrown away by deleting subjects. With fewer data points, estimation based on a CC analysis may result in larger variances than methods for incomplete data. Also, since CC analysis does not include all subjects randomized in the final analyses, it does not adhere to the intent-to-treat principle of clinical trials [7]. It is for these reasons that CC analyses are usually not employed in medical research.

Mixed Models

The mixed model is a regression model that includes both population-level and subject-level effects. It assumes responses are MAR. For a pre-post study with one treatment group and one control group, a mixed model can be written as:

$$Y_{ij} = \beta_0 + \beta_1 Trt_i + \beta_2 Post_j + \beta_3 (Trt_i \times Post_j) + b_{i1} + b_{i2} Post_j + \varepsilon_{ij}. \qquad (2.7)$$

Here, Y_{ij} is the response of subject i at time point j ($j = 1, 2$) and Trt_i and $Post_j$ are indicators of treatment group and time, respectively. They are defined as

$$Trt_i = \begin{cases} 0, & \text{if subject } i \text{ is in the control group,} \\ 1, & \text{if subject } i \text{ is in the treatment group.} \end{cases}$$

and

$$Post_j = \begin{cases} 0, & \text{if } Y_{ij} \text{ is the pre-test outcome } (j = 1), \\ 1, & \text{if } Y_{ij} \text{ is the post-test outcome } (j = 2). \end{cases}$$

The last three terms $b_{i1} + b_{i2} Post_j + \varepsilon_{ij}$ together is the error for subject i at time j, in which b_{i1} is the effect of subject i at pre-test, b_{i2} is the effect of subject i at post-test, and ε_{ij} is the random error for subject i at time j.

The mixed model has the following assumptions:

1. $\begin{pmatrix} b_{i1} \\ b_{i2} \end{pmatrix} \overset{iid}{\sim} BVN\left(\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \sigma_{12}^2 \\ \sigma_{12}^2 & \sigma_2^2 \end{pmatrix}\right)$,
2. $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$ for all i and j,
3. ε_{ij}'s are independent from b_{i1}'s and b_{i2}'s.

Under these assumptions,

$$\begin{aligned} Var(Y_{i1}) &= Var(b_{i1} + \varepsilon_{ij}) = \sigma_1^2 + \sigma_\varepsilon^2, \\ Var(Y_{i2}) &= Var(b_{i1} + b_{i2} + \varepsilon_{ij}) = \sigma_1^2 + \sigma_2^2 + 2\sigma_{12}^2 + \sigma_\varepsilon^2, \\ Cov(Y_{i1}, Y_{i2}) &= Var(b_{i1}) + Cov(b_{i1}, b_{i2}) = \sigma_1^2 + \sigma_{12}^2. \end{aligned}$$

The variance-covariance matrix is said to be unstructured in this model. If we furthermore assume $\sigma_2^2 = 0$, then $\sigma_{12}^2 = 0$ and we have

$$\begin{aligned} Var(Y_{i1}) = Var(Y_{i2}) &= \sigma_1^2 + \sigma_\varepsilon^2, \\ Cov(Y_{i1}, Y_{i2}) &= \sigma_1^2. \end{aligned}$$

Hence, $\sigma_2^2 = 0$ is a sufficient condition for a compound symmetry (CS) variance-covariance structure for the mixed model (2.7).

Note that in (2.7), the post-test outcome is

$$Y_{i2} = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)Trt_i + b_{i1} + b_{i2} + \varepsilon_{i2},$$

and the pre-test outcome is

$$Y_{i1} = \beta_0 + \beta_1 Trt_i + b_{i1} + \varepsilon_{i1}.$$

By subtracting the two we get

$$Y_{i2} - Y_{i1} = \beta_2 + \beta_3 Trt_i + b_{i2} + \varepsilon_{i2} - \varepsilon_{i1}.$$

If we let $\tilde{\varepsilon}_i = b_{i2} + \varepsilon_{i2} - \varepsilon_{i1}$, then $\tilde{\varepsilon}_i \overset{iid}{\sim} N(0, \sigma_2^2 + 2\sigma_\varepsilon^2)$ and therefore the above model becomes the change score model. Thus, when there is no missing data, estimating β_3 using the mixed model is the same as estimating β_1 in the change score model (2.1). However, unlike the change score model, the mixed model can include data on subjects with just one of the two outcome values. Therefore, hypotheses test in the change score analysis can be generalized to testing

$$\begin{aligned} H_0 &: \quad \beta_3 = 0, \\ H_a &: \quad \beta_3 \neq 0 \end{aligned}$$

in the mixed model (2.7).

Multiple Imputation (MI)

The general idea of multiple imputation is to fill in the missing data several times and obtain several "complete" data sets, then conduct the statistical inferences based on those completed data sets. Compared to single imputation, multiple imputation accounts for variability in the estimate for the missing value. Therefore, the estimated variance of the parameters from MI is larger than single imputation, but should be closer to the true variance. Rigorous MI assumes: (1) an ignorable missing data mechanism, (2) proper imputations, and (3) congeniality [6, 17, 18].

The parameter vector ϕ may be separated into two parts $\phi = (\theta', \psi')'$ such that θ is the vector of parameters for the data generating process and ψ is the vector of parameters for the missing data mechanism. A missing data mechanism is said to be ignorable for Bayesian based inference if data are MAR and the prior distribution of θ and ψ has the form $p(\theta, \psi) = p(\theta)p(\psi)$ [6]. Parameters θ and ψ with the above form are said to be *a priori* independent [6]. When the missing data mechanism is ignorable,

$$\begin{aligned} f(\theta, \psi | Y_{\text{obs}}, M) &\propto f(\theta, \psi)\mathcal{L}(\theta, \psi | Y_{\text{obs}}, M) \\ &\propto f(\theta | Y_{\text{obs}}) f(\psi | Y_{\text{obs}}, M). \end{aligned}$$

Therefore, inference about θ can be made from $f(\theta|Y_{\text{obs}})$, ignoring the missing data mechanism.

A proper imputation is defined as drawing the missing data from the joint posterior distribution $f(\theta, Y_{\text{mis}}|Y_{\text{obs}})$ [17]. This means that the variabilities of both θ and Y_{mis} should be considered during the imputation process. To obtain valid inferences from proper imputations, a "correct" full model is required. When the full model does

not describe the data well, some short cuts of the proper imputation, which usually approximate $f(\theta, Y_{\text{mis}}|Y_{\text{obs}})$, may even provide more effective analyses than the proper imputation [6]. However, these approximate methods may neglect the variability of θ or Y_{mis}, or set up the joint distribution only based on a subset of fully observed values [6].

An imputation is said to be congenial if the model used to analyze the data can be derived from the imputation model. That is, imputation model has to include all variables in the analysis model. Therefore, to get a congenial imputation, the analysis model must be either equal to or nested within the imputation model [6, 17]. Parameter estimates may be biased when the imputation is not congenial [17].

For MI, consider

$$f(\theta|Y_{\text{obs}}) = \int f(\theta, Y_{\text{mis}}|Y_{\text{obs}})\,\mathrm{d}Y_{\text{mis}} = \int f(\theta|Y_{\text{mis}}, Y_{\text{obs}}) f(Y_{\text{mis}}|Y_{\text{obs}})\,\mathrm{d}Y_{\text{mis}}.$$

Treat $f(\theta|Y_{\text{mis}}, Y_{\text{obs}})$ as a function of Y_{mis} and let $f(Y_{\text{mis}}|Y_{\text{obs}})$ be the conditional p.d.f. of Y_{mis}. By Monte Carlo integration,

$$\begin{aligned} f(\theta|Y_{\text{obs}}) &= \int f(\theta|Y_{\text{mis}}, Y_{\text{obs}}) f(Y_{\text{mis}}|Y_{\text{obs}})\,\mathrm{d}Y_{\text{mis}} \\ &= E_{Y_{\text{mis}}|Y_{\text{obs}}}[f(\theta|Y_{\text{mis}}, Y_{\text{obs}})] \\ &\approx \frac{1}{N}\sum_{i=1}^{N} f(\theta|Y_{\text{mis}}^{(i)}, Y_{\text{obs}}), \end{aligned}$$

where, $Y_{\text{mis}}^{(i)}$'s are sampled from $(Y_{\text{mis}}|Y_{\text{obs}})$. The above formula implies that if we sample $Y_{\text{mis}}^{(i)}$'s several times and obtain several completed data sets without missing values, then the distribution of θ can be estimated from the average of conditional distributions of θ given those completed data sets; this is the idea behind MI. The general procedure for MI can be summarized as three steps:

Step 1: Fill in the missing data N times and get N complete data sets

Step 2: Analyze each of the N data sets separately using regular analysis methods

Step 3: Combine the analysis results for the N data sets to obtain the result for the original data set which has missing values.

In Step 1, since sampling $Y_{\text{mis}}^{(i)}$ directly from $(Y_{\text{mis}}|Y_{\text{obs}})$ is usually difficult, instead, we may sample $Y_{\text{mis}}^{(i)}$ from joint posterior distribution of $(\theta, Y_{\text{mis}}|Y_{\text{obs}})$. To do so, general procedures for Bayesian data augmentation, such as the Gibbs' sampler, can be used. The general steps for the Gibbs' sampler are as follows:

1. Sample initial values $\theta^{(0)}$ from an approximated posterior distribution $f(\theta|Y_{\text{obs}})$;
2. For $i \geq 1$,
   (a) sample $Y_{\text{mis}}^{(i)}$ from $(Y_{\text{mis}}|\theta^{(i-1)}, Y_{\text{obs}})$;
   (b) sample $\theta^{(i)}$ from $(\theta|Y_{\text{mis}}^{(i)}, Y_{\text{obs}})$.

It can be shown that as the iteration number t approaches infinity, $(\theta^{(t)}, Y_{\text{mis}}^{(t)})$ is approximately a draw from the joint posterior distribution of $(\theta, Y_{\text{mis}}|Y_{\text{obs}})$ [6,19,20].

In Step 2, regular analysis methods for complete data sets are applied to each filled-in data set. In Step 3, analysis results from the N filled-in data sets are combined. To estimate $E(\theta|Y_{\text{obs}})$ and $Var(\theta|Y_{\text{obs}})$, first recall that

$$\begin{aligned} E(X) &= E_Y[E(X|Y)], \\ Var(X) &= E[Var(X|Y)] + Var[E(X|Y)], \end{aligned}$$

where E_Y is the expectation operator for variable Y. In the above two equations, set $X = \theta$ and $Y = Y_{\text{mis}}$. Then,

$$\begin{aligned} E(\theta|Y_{\text{obs}}) &= E_{Y_{\text{mis}}}[E(\theta|Y_{\text{mis}}, Y_{\text{obs}})|Y_{\text{obs}}], \\ Var(\theta|Y_{\text{obs}}) &= E[Var(\theta|Y_{\text{mis}}, Y_{\text{obs}})|Y_{\text{obs}}] + Var[E(\theta|Y_{\text{mis}}, Y_{\text{obs}})|Y_{\text{obs}}]. \end{aligned}$$

Therefore,

$$\begin{aligned} E(\theta|Y_{\text{obs}}) &= \int E(\theta|Y_{\text{mis}}, Y_{\text{obs}}) \times f(Y_{\text{mis}}|Y_{\text{obs}})\, \mathrm{d}Y_{\text{mis}} \\ &\approx \frac{1}{N}\sum_{i=1}^{N} E(\theta|Y_{\text{mis}}^{(i)}, Y_{\text{obs}}) \\ &= \frac{1}{N}\sum_{i=1}^{N} \hat{\theta}^{(i)}, \end{aligned} \tag{2.8}$$

where $Y_{\text{mis}}^{(i)}$'s are sampled from $f(Y_{\text{mis}}|Y_{\text{obs}})$ and $\hat{\theta}^{(i)} = E(\theta|Y_{\text{mis}}^{(i)}, Y_{\text{obs}})$. The approximation in the second step is obtained from Monte Carlo integration. Similarly, one can get

$$\begin{aligned} Var(\theta|Y_{\text{obs}}) &\approx \frac{1}{N}\sum_{i=1}^{N} V_i + \frac{1}{N-1}\sum_{i=1}^{N}(\hat{\theta}_i - \overline{\theta})^2 \\ &= \overline{V} + B, \end{aligned} \tag{2.9}$$

where $V_i = Var(\theta|Y_{\text{mis}}^{(i)}, Y_{\text{obs}})$, $\overline{V} = \frac{1}{N}\sum_{i=1}^{N} V_i$ is the within-imputation variance, and $B = \frac{1}{N-1}\sum_{i=1}^{N}(\hat{\theta}_i - \overline{\theta})^2$ is the between-imputation variance. When N is small, an adjustment can be made in (2.9) to improve the approximation [6]:

$$Var(\theta|Y_{\text{obs}}) \approx \overline{V} + (1 + \frac{1}{N})B. \tag{2.10}$$

In SAS, Step 1 is performed using PROC MI, Step 2 is performed using a standard analysis procedure appropriate for the problem (e.g., PROC REG), and Step 3 is performed using PROC MIANALYZE.

If the variables $Y_1, \ldots, Y_k$ in the data set can be reordered such that y_{ij} is not missing when $y_{i,j+1}$ is not missing for all $j = 1, \ldots, k-1$, then the data set is said to have a monotone missing pattern. For data with a monotone missing pattern, PROC MI in SAS provides a MONOTONE statement with several choices to impute

the missing data: regression method, predicted mean method, or propensity score method for continuous variables; logistic regression method for binary or ordinal variables; discriminant function method for binary or nominal variables [21].

For the regression method in PROC MI, missing values are imputed sequentially, from the variable with the fewest to the most missing values, within each imputed data set [21]. To be clearer, let $Y_{i(\text{obs})}$ be variable Y_i with all observed values, $\boldsymbol{\theta}_j$ be all parameters from jth data generating process, $\boldsymbol{Y}_j^{(*)}$ be the vectors of simulated values for missing data in Y_j, and $\boldsymbol{\theta}_j^{(*)}$ be a draw of $\boldsymbol{\theta}_j$ based on its joint distribution with $\boldsymbol{Y}_{\text{obs}}$ $(j = 2, \ldots, k)$. If variables Y_1, $Y_2, \ldots, Y_k$ are reordered to a monotonic missing pattern (y_{ij} not missing if $y_{i,j+1}$ not missing for all $j = 1, \ldots, k-1$), then each complete data set in Step 1 is generated as follows:

For $j = 2, \ldots, k$, if Y_j has missing values, then

1. Sample $\boldsymbol{\theta}_j^{(*)}$ from $(\boldsymbol{\theta}_j | Y_{1(\text{obs})}, \ldots, Y_{j(\text{obs})})$;

2. Sample $\boldsymbol{Y}_j^{(*)}$ from $(Y_j | \boldsymbol{\theta}_j^{(*)}, Y_{1(\text{obs})}, \ldots, Y_{j(\text{obs})})$;

3. Include values of $\boldsymbol{Y}_j^{(*)}$ in $Y_{j(\text{obs})}$ and use updated $Y_{j(\text{obs})}$ in following steps.

For data with a non-monotone missing pattern, SAS 9.3 provides an MCMC statement in the PROC MI procedure to impute missing data using an MCMC method which assumes a multivariate normal distribution for the data [21]. The MCMC method can be used individually to impute the entire set of missing data; it can also be used jointly with a method for monotone missing pattern by only imputing enough data to obtain a monotone missing pattern [21]. For this study, we only focus on the regression method for monotone missing pattern, since we assume our data have a monotone missing pattern (only missing outcomes at post-test) and the

regression method for monotone missing pattern provides more flexibility than the MCMC method in assigning the imputation model [21].

Chapter 3: Analysis of the BePHIT Data

3.1 Study Description

The BePHIT (Behavior and Exercise for Physical Health Intervention) study was a randomized controlled study of a 12-week walking intervention conducted on post-menopausal women between January 2008 and March 2009 [22]. The primary outcome was the change in time for women to finish a one-mile walk.

After passing the selection criteria, 71 participants were stratified by BMI and randomized into either a coach group or a no-coach group [22]. For women in the coach group, a trained coach was assigned [22]. The role of the coach was to explain the intervention, provide the first week's steps goal, train subjects to use the pedometer and the IVR system, and offer help during the intervention [22]. Women in the no-coach group received similar instructions, training, and help, except that they were not informed that they had access to a coach [22].

Among the 71 randomized participants, 35 were assigned to the coach group and 36 to the no-coach group [22]. For the no-coach group, baseline walking time was only available for 35 patients. In total, 12 (17%) patients dropped out before the post walking test, 4 of whom were in the coach group and 8 in the no-coach group.

Table 3.1: Summary of Loss-to-Follow-up

	Total	Coach	No coach
Started	70	35	35
Dropped Out	12	4	8
Finished	58	31	27

The drop out rate was the same for the two groups ($p = 0.20$). Table 3.1 summarizes the change in the number of participants through out the study.

3.2 Analyses

Logistic regression was used to determine whether or not dropping out was related to pre-test values of any baseline measures. The estimates of the univariate logistic models and their significance test results are listed in Table 3.2. Among all the factors, only waist-hip ratio had a significant effect on missing post-test values ($p = 0.04$). People with larger waist-hip ratio tended to be more likely to drop out the study. Since waist-hip ratio was related to the dropout rate, the missing data mechanism was not MCAR.

David et al. used linear mixed models to analyze the pre-post changes [22]. Here we analyzed the data using a mixed model, ANCOVA after multiple imputation (MI), and the corresponding CC (complete-case) analyses (the mixed model with CC and ANCOVA with CC). Results from each method were compared. The ANCOVA model is valid here, since the primary outcome, time for one-mile walk, can be assumed to be measured without error. For both the mixed model and ANCOVA after MI, all 70 participants with baseline walking test were included. For the two CC analyses,

Table 3.2: Logistic Regression Estimates for Drop Out

Variable	Estimate	SE	Odds Ratio	p-value
Design and Outcome				
Treatment	-0.83	0.67	0.44	0.21
Pre-test	0.11	0.16	1.12	0.47
Baseline Anthropometrics				
Pulse rate	0.04	0.03	1.05	0.12
Waist/hip	12.29	5.93	>999.99	0.04*
BMI	0.11	0.08	1.12	0.16
Baseline Psychometrics				
PRETHOUG	0.32	0.52	1.38	0.54
PREEXSTA	0.21	0.42	1.23	0.62
Social support from family	-0.16	0.48	0.86	0.75
Social support from friends	-0.94	0.71	0.39	0.19
PRESE	-0.17	0.17	0.84	0.32
PREQ16	0.19	0.21	1.21	0.38
Exercise goals	0.12	0.36	1.12	0.74
Exercise planning	-0.36	0.54	0.70	0.50

* Significant at 0.05 significance level.

only participants with both baseline and post-intervention walking time were used ($n = 58$).

An unstructured (UN) variance-covariance matrix for the pre- and post-test values was used for the mixed models. In MI, post-test values were imputed by the 12 baseline measures listed in Table 3.2 including treatment. Missing values were imputed 20 times. Imputed data were analyzed using the ANCOVA model (2.2) and results were combined.

The imputation model was fit using the completely observed cases to check whether the variables were predictive of post-test. The estimates from the imputation model

Table 3.3: Regression Estimates for Imputation Model

Variable	Estimate	SE	p-value
Intercept	-3.07	4.23	0.4712
Design and Outcome			
Treatment	0.01	0.40	0.9827
Pre-test	0.78	0.11	<.0001*
Baseline Anthropometrics			
Pulse rate	-0.01	0.02	0.6068
Waist/hip	2.56	4.01	0.5271
BMI	0.16	0.05	0.0040*
Baseline Psychometrics			
PRETHOUG	-0.28	0.39	0.4826
PREEXSTA	0.15	0.28	0.6011
Social support from family	0.29	0.30	0.3467
Social support from friends	-0.26	0.36	0.4750
PRESE	-0.13	0.14	0.3744
PREQ16	0.14	0.14	0.3076
Exercise goals	0.29	0.34	0.4100
Exercise planning	-0.54	0.55	0.3288

* Significant at 0.05 significance level.

are listed in Table 3.3. The r-squared of the imputation model was 0.6975. However, only two variables, pre-test walk time and BMI, were significant at the 0.05 level.

Estimates and tests of a treatment effect for the four methods are summarized in Table 3.4. All four methods showed no coach effect on the change in one-mile walk time. This result is consistent with those reported by David et al. [22].

As we can see, the results from the two mixed models were very similar. However, there were obvious differences in both the estimates and hypothesis tests between the two ANCOVA analyses. When MI was used, the estimated effect for treatment changed from -0.08 to -0.24, and its corresponding standard error changed from 0.38 to 0.63.

Table 3.4: Estimates Comparison from 4 Analysis Methods

Method	Variable	Estimate	SE	t	p-value
Mixed Model, CC	Trt. × Post	-0.05	0.38	-0.13	0.8995
Mixed Model, AA	Trt. × Post	-0.05	0.38	-0.13	0.8959
ANCOVA, CC	Treatment	-0.08	0.38	-0.22	0.8251
ANCOVA, MI	Treatment	-0.24	0.63	-0.39	0.7005

Abbreviations: CC, complete-case analysis; AA, all-available-case analysis.

ANCOVA after MI gave the most significant test of treatment among all four methods. The treatment effect tests from the two ANCOVA methods were more significant than from the two mixed models. Within each model, test of treatment effect using all available cases for the analysis was more significant than using complete cases only.

Chapter 4: Simulation Study

4.1 Data Generation

The simulation studies were based on the BePHIT study discussed in Chapter 3. The coefficient values used in the data generation process were, for the most part, estimated from the BePHIT data. The data sets for the simulation studies were generated from the following model:

$$Y_{ij} = 17.946 - 0.811 Post_{ij} - 2WH_i - Trt_i \times Post_{ij} + b_{i0} + \varepsilon_{ij}, \qquad (4.1)$$

$i = 1, \ldots, n$; $j = 1, 2$; $b_{i0} \overset{iid}{\sim} N(0, 3.069)$; $\varepsilon_i \overset{iid}{\sim} N(0, 1.005)$. Here Y_{ij} is the outcome of subject i at time j and $Post_j$ and Trt_i are indicator variables of time point and treatment group assignment, respectively. The variable WH_i is the waist-hip ratio of subject i, which was generated from $N(0.938, 0.121)$.

In contrast to the mixed model fit to the BePHIT data, the treatment effect was removed in (4.1) so that the average outcome value of the two groups was same at baseline. Also, a main effect of waist-hip ratio was added because waist-hip ratio was significantly related to missing follow-ups in the BePHIT study. Its coefficient was adjusted to -2 so that, for all the analysis models, waist-hip ratio was, on average, significantly associated with the outcome. In the original analyses of the BePHIT

study, the interaction effect of treatment and time was not significant. However, in order to conduct power comparisons in the simulation studies, the coefficient for this interaction effect was adjusted to -1 so that, on average, it was significant for all the analysis models.

Two different group sizes were considered: $n = 35$ (i.e., the BePHIT sample size) and $n = 100$ per group. For each group size, 500 data sets were simulated. Post data were then set to be missing at 20%, 30%, 40%, and 50%. At each percentage of missingness, missing data were generated as missing completely at random (MCAR) and missing at random (MAR). For MAR, missingness was generated under three different conditions: dependent on waist-hip ratio, dependent on both waist-hip ratio and pre-test level, and dependent on waist-hip ratio, pre-test level, and treatment group assignment. To generate the missing data, a Bernoulli indicator was drawn for each subject with the probability defined by the following logistic regression model:

$$\text{log-odds of missing follow-up} = \gamma_0 + \gamma_1 WH_i + \gamma_2 Y_{i1} + \gamma_3 Trt_i. \tag{4.2}$$

The γ's were first estimated from the BePHIT data. However, since the analyses of drop out rate in the original study showed very different effects of the three terms (Table 3.2), the coefficients of WH_i and Trt_i were adjusted so that their effects would be similar to the pre-test level's. The coefficient of WH_i was adjusted by forcing the odds ratio of drop out to be two for a one standard deviation change in waist-hip ratio; the coefficient of Trt_i was adjusted by forcing the odds ratio of drop out for the treatment group versus the control group to be four. The γ's were set to 0 when the missingness did not depend on their corresponding variables. After each coefficient was determined, the same value was carried on to the more complex MAR mechanisms. Finally, γ_0 was adjusted for each percentage of missingness. Table 4.1

Table 4.1: Coefficients for Log-odds of Missing Follow-up

% Missingness	Missing Mechanism	γ_0	γ_1	γ_2	γ_3
20	MCAR	-1.386	0	0	0
	MAR (WH)	-3.252	1.989	0	0
	MAR (WH + Pre)	-4.754	1.989	0.094	0
	MAR (WH + Pre + Trt)	-5.448	1.989	0.094	1.386
30	MCAR	-0.847	0	0	0
	MAR (WH)	-2.713	1.989	0	0
	MAR (WH + Pre)	-4.215	1.989	0.094	0
	MAR (WH + Pre + Trt)	-4.909	1.989	0.094	1.386
40	MCAR	-0.405	0	0	0
	MAR (WH)	-2.271	1.989	0	0
	MAR (WH + Pre)	-3.774	1.989	0.094	0
	MAR (WH + Pre + Trt)	-4.467	1.989	0.094	1.386
50	MCAR	0	0	0	0
	MAR (WH)	-1.866	1.989	0	0
	MAR (WH + Pre)	-3.368	1.989	0.094	0
	MAR (WH + Pre + Trt)	-4.061	1.989	0.094	1.386

Abbreviations: MAR (WH), missingness dependent on waist-hip ratio; MAR (WH + Pre), missingness dependent on waist-hip ratio and pre-test level; MAR (WH + Pre + Trt), missingness dependent on waist-hip ratio, pre-test level, and treatment assignment.

summarizes the values of the coefficients for log-odds function (4.2) for each scenario used in the simulations.

4.2 Analyses

For each scenario mentioned above, simulated data sets were analyzed by both mixed models and ANCOVA after multiple imputation (MI). Since outcome values were generated from a model containing waist-hip ratio, in addition to the mixed model and ANCOVA model mentioned in Chapter 2, each method was also conducted with waist-hip ratio in the analysis model. For comparison purposes, we also

performed complete-case (CC) analyses using the same mixed models and ANCOVA models.

The mixed model used for analysis was:

$$Y_{ij} = \beta_0 + \beta_1 Trt_i + \beta_2 Post_{ij} + \beta_3 Trt_i \times Post_{ij} + \beta_4 WH_i + b_{i1} + \varepsilon_{ij}, \tag{4.3}$$

where $b_{i1} \overset{iid}{\sim} N(0, \sigma_1^2)$ and $\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$. The Kenward-Roger method was used for computing denominator degrees of freedom [23–25]. The variance-covariance structure was assumed to be compound symmetry (CS).

The ANCOVA model used for analysis was:

$$Y_{i2} = \beta_0 + \beta_1 Y_{i1} + \beta_2 Trt_i + \beta_3 WH_i + \varepsilon_i, \tag{4.4}$$

where $\varepsilon_i \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$. The parameter β_3 was set to 0 when waist-hip ratio was not included in the analysis model.

For MI, missing follow-ups were imputed 20 times using the following imputation model:

$$Y_{i2} = \beta_0 + \beta_1 Y_{i1} + \beta_2 Trt_i + \beta_3 WH_i + \beta_4 Y_{i1} \times Trt_i + \beta_5 Y_{i1} \times WH_i + \beta_6 Trt_i \times WH_i.$$

Thus, the predictors in the imputation model were all the main effects that were used in generating the data and all their two-way interactions.

The statistical package used for both data generation and analyses was SAS 9.3 (SAS Inc., Cary, NC).

4.3 Results

As mentioned in Chapter 2, the hypothesis test of main interest was

$$\begin{aligned} H_0 &: \quad \beta_3 = 0, \\ H_a &: \quad \beta_3 \neq 0 \end{aligned}$$

for mixed model (4.3) and

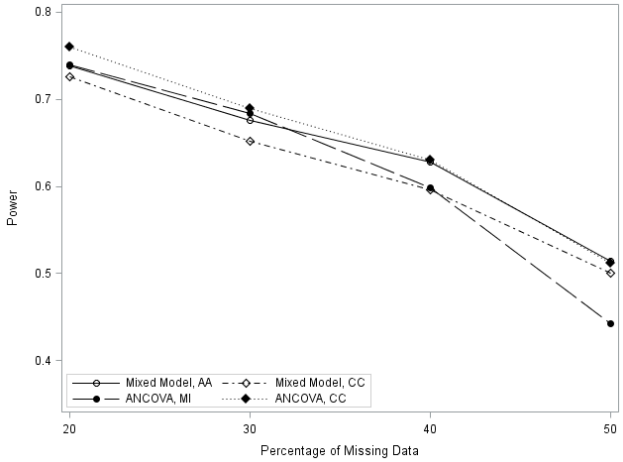
$$H_0 : \quad \beta_2 = 0,$$
$$H_a : \quad \beta_2 \neq 0.$$

for ANCOVA model (4.4). Loosely speaking, the two tests were both testing the treatment effect on the outcome levels; however, there are some subtle differences between them: the test for the mixed model was testing the treatment effect on the differences between pre- and post-test outcomes, while the test for the ANCOVA model was testing the treatment effect on the post-test response controlling for the pre-test level.

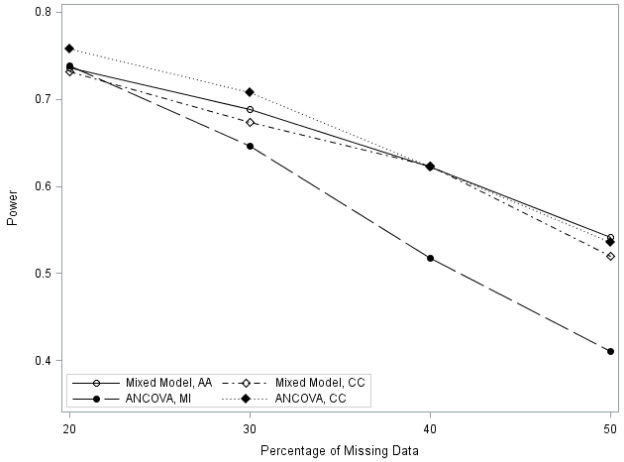
Figures 4.1 and 4.2 compare the power of the above tests for each method when waist-hip ratio was included in the analysis models. Here α was set to 0.05. As expected, the power of each test was larger when the sample size was larger. Also, it is clear that the power decreased when the percentage of missingness increased. Under MCAR, ANCOVA after multiple imputation (MI) was more powerful than the mixed model using all available data when percentage of missingness was small (20% and 30%). Under MAR, ANCOVA after MI and the mixed model using all available data were still comparable when post data were missing at 20%. However, the power of ANCOVA after MI decreased dramatically when the percentage of missingness increased. Therefore, ANCOVA after MI became the least powerful method when 50% of the post data were missing.

When $n = 35$, the complete-case (CC) ANCOVA was apparently the best in terms of the power when percentage of missingness was small (20% and 30%). When missing proportion became larger (40% and 50%), the all-available-case mixed model yielded similar results to CC ANCOVA.

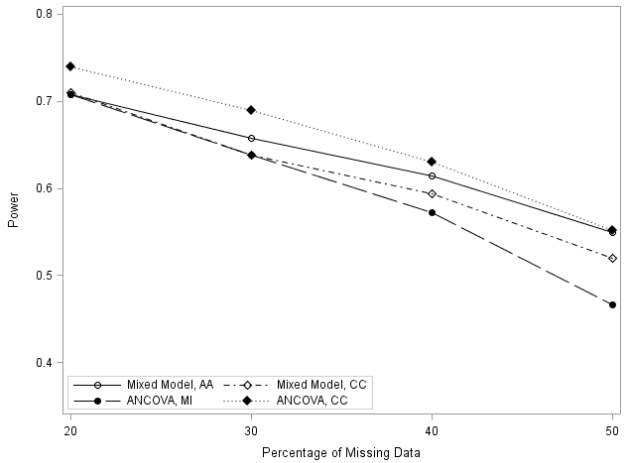
Figure 4.1: Power Comparisons for Analysis Models including WH and $n = 35$



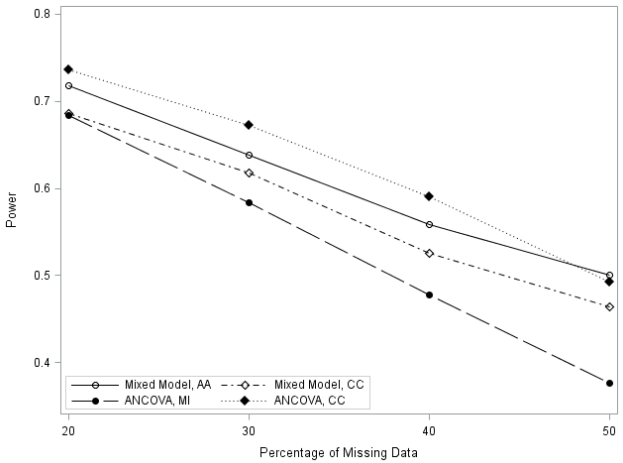
(a) Power Comparison under MCAR



(b) Power Comparison under MAR (WH)

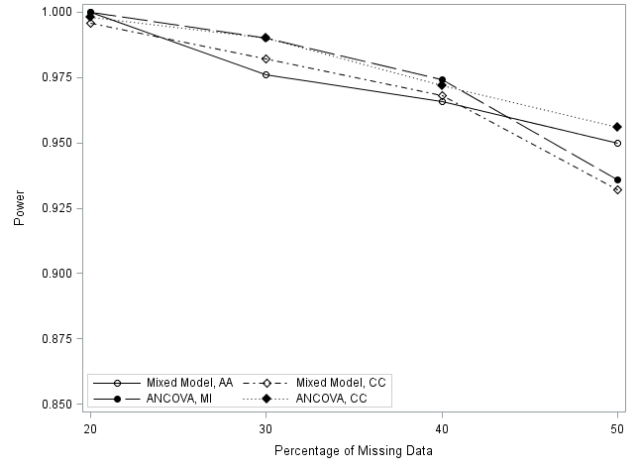


(c) Power Comparison under MAR (WH + Pre)

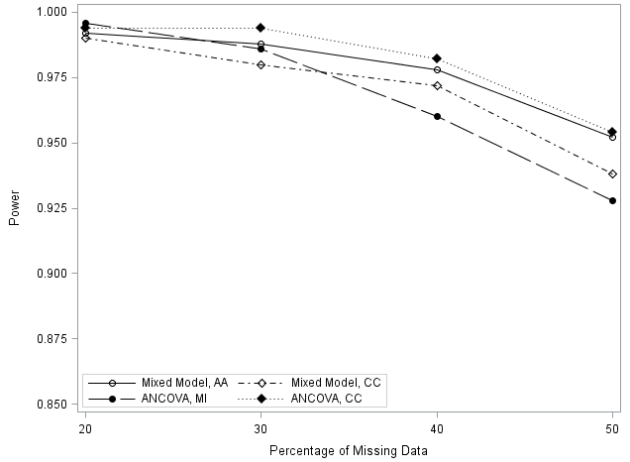


(d) Power Comparison under MAR (WH + Pre + Trt)

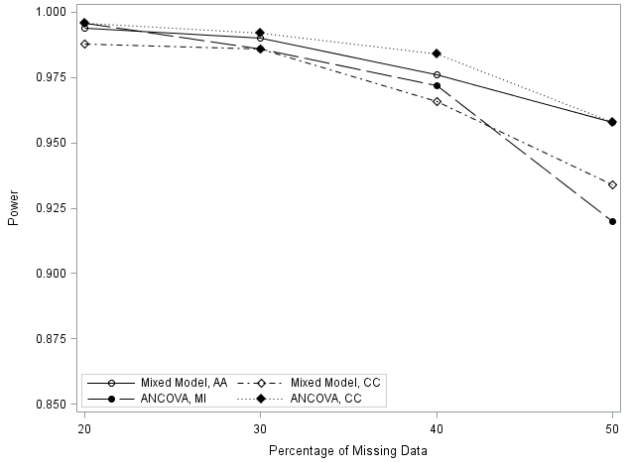
Figure 4.2: Power Comparisons for Analysis Models including WH and $n = 100$



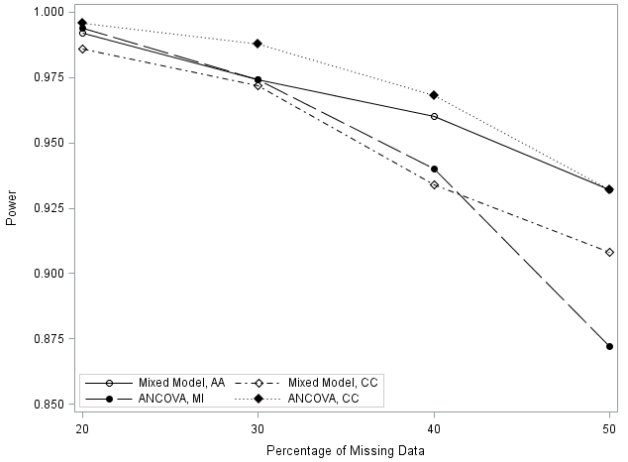
(a) Power Comparison under MCAR



(b) Power Comparison under MAR (WH)



(c) Power Comparison under MAR (WH + Pre)



(d) Power Comparison under MAR (WH + Pre + Trt)

When $n = 100$, the power of each test was very high (mostly > 0.9). CC ANCOVA and the all-available-case mixed model had very similar power under each scenario. The powers for ANCOVA after MI and the CC mixed model dropped faster when the percentage of missingness was large (40% and 50%). Under MAR, ANCOVA after MI became the worst in power when 50% of the post data were missing.

Tables 4.2 and 4.3 summarize the results of the simulation studies when 20% of post data were set to missing. Overall speaking, all methods produced unbiased estimates when waist-hip ratio was included in the analysis model. Also, the estimates were more accurate at the larger sample size ($n = 100$ per group). For the mixed models, the estimated interaction between treatment and time point was a little biased for MCAR data with 35 subjects per group (mean -0.97 vs. the truth -1). Similar results were also observed for the estimated treatment effect in the ANCOVA models under the same design setup (means -0.97, -0.98 vs. the truth -1). The estimated effect of pre-test level was a little biased (mean 0.78 vs. the truth 0.75) when waist-hip ratio was removed from the ANCOVA model. Also, for the ANCOVA models, the estimated effect of waist-hip ratio was biased for data with MCAR or MAR dependent on waist-hip ratio when the sample size was 35 (-0.45 for MCAR, -0.42 for MAR dependent on waist-hip ratio vs. the truth -0.49).

The all-available-case mixed model consistently had smaller standard errors of the estimates than the CC mixed model, while the reverse was true for the ANCOVA analyses.

Similar results with less power and larger standard errors were observed when the percentage of missing follow-ups was set to 30%, 40%, and 50%. The results of these

simulation studies along with the figures of power comparisons when waist-hip ratio was not included in the analysis model can be found in Appendix A.

Table 4.2: Mixed Model – 20% Missingness

Scenario	Method			Results														
				Intercept			Treatment			Post			Trt × Post			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth[a]				17.95			-0.81			0			-1			-2		
MCAR	$n = 35$	w/ WH	CC	17.93	0.79	23.09	-0.82	0.27	-3.13	-0.00	0.54	-0.01	-0.97	0.38	-2.60	-1.98	0.74	-2.72
			AA	17.95	0.72	25.44	-0.83	0.26	-3.16	0.00	0.49	0.00	-0.97	0.37	-2.63	-2.00	0.67	-3.02
		w/o WH	CC	16.08	0.40	40.34	-0.82	0.27	-3.13	-0.01	0.57	-0.02	-0.97	0.38	-2.60	.	.	.
			AA	16.08	0.36	44.84	-0.83	0.26	-3.16	0.00	0.51	-0.00	-0.97	0.37	-2.63	.	.	.
	$n = 100$	w/ WH	CC	17.96	0.46	38.90	-0.82	0.16	-5.18	-0.01	0.32	-0.03	-0.99	0.22	-4.42	-2.00	0.43	-4.67
			AA	17.96	0.42	43.09	-0.82	0.16	-5.25	-0.01	0.29	-0.04	-0.99	0.22	-4.47	-2.01	0.39	-5.15
		w/o WH	CC	16.08	0.24	67.39	-0.82	0.16	-5.18	-0.01	0.34	-0.02	-0.99	0.22	-4.42	.	.	.
			AA	16.08	0.21	75.48	-0.82	0.16	-5.24	-0.01	0.30	-0.03	-0.99	0.22	-4.46	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	17.93	0.79	22.99	-0.81	0.27	-3.02	-0.01	0.55	-0.02	-0.99	0.38	-2.64	-1.98	0.78	-2.57
			AA	17.94	0.72	25.30	-0.81	0.27	-3.06	0.00	0.49	0.00	-1.00	0.38	-2.67	-1.99	0.68	-2.99
		w/o WH	CC	16.18	0.41	40.04	-0.81	0.27	-3.02	-0.01	0.58	-0.02	-0.99	0.38	-2.64	.	.	.
			AA	16.08	0.36	44.85	-0.79	0.27	-2.97	0.00	0.51	-0.00	-1.00	0.38	-2.67	.	.	.
	$n = 100$	w/ WH	CC	17.95	0.46	39.03	-0.81	0.16	-5.05	-0.00	0.32	-0.01	-1.00	0.23	-4.44	-2.00	0.45	-4.45
			AA	17.96	0.42	42.97	-0.81	0.16	-5.13	-0.01	0.29	-0.04	-1.00	0.22	-4.49	-2.00	0.39	-5.12
		w/o WH	CC	16.18	0.24	67.31	-0.81	0.16	-5.05	-0.01	0.34	-0.02	-1.00	0.23	-4.44	.	.	.
			AA	16.08	0.21	75.58	-0.79	0.16	-4.98	-0.01	0.30	-0.03	-1.00	0.22	-4.48	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	17.94	0.79	23.08	-0.81	0.27	-3.03	0.02	0.55	0.03	-0.98	0.38	-2.61	-2.07	0.78	-2.74
			AA	17.96	0.72	25.36	-0.82	0.27	-3.13	0.00	0.49	0.00	-0.98	0.38	-2.63	-2.01	0.68	-3.02
		w/o WH	CC	16.09	0.41	39.80	-0.81	0.27	-3.03	0.02	0.58	0.03	-0.98	0.38	-2.61	.	.	.
			AA	16.08	0.36	44.82	-0.80	0.27	-3.05	0.00	0.51	-0.00	-0.98	0.38	-2.63	.	.	.
	$n = 100$	w/ WH	CC	17.95	0.46	39.13	-0.79	0.16	-5.00	-0.01	0.32	-0.02	-0.99	0.23	-4.42	-2.09	0.45	-4.67
			AA	17.96	0.42	42.98	-0.81	0.16	-5.18	-0.01	0.29	-0.04	-0.99	0.22	-4.47	-2.01	0.39	-5.14
		w/o WH	CC	16.09	0.24	66.93	-0.79	0.16	-5.00	-0.01	0.34	-0.02	-0.99	0.23	-4.42	.	.	.
			AA	16.08	0.21	75.52	-0.79	0.16	-5.04	-0.01	0.30	-0.03	-0.99	0.22	-4.47	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	17.99	0.80	22.85	-0.81	0.26	-3.20	-0.07	0.57	-0.13	-0.98	0.39	-2.52	-2.09	0.79	-2.71
			AA	17.96	0.72	25.35	-0.82	0.25	-3.27	0.00	0.49	0.00	-0.99	0.39	-2.60	-2.01	0.68	-3.02
		w/o WH	CC	16.09	0.39	41.92	-0.81	0.26	-3.20	0.01	0.59	0.02	-0.98	0.39	-2.52	.	.	.
			AA	16.08	0.36	44.84	-0.81	0.25	-3.22	0.00	0.51	-0.00	-0.97	0.39	-2.54	.	.	.
	$n = 100$	w/ WH	CC	17.99	0.47	38.81	-0.80	0.15	-5.33	-0.09	0.33	-0.26	-0.99	0.23	-4.28	-2.09	0.45	-4.64
			AA	17.96	0.42	42.98	-0.82	0.15	-5.44	-0.01	0.29	-0.04	-1.00	0.23	-4.42	-2.01	0.39	-5.14
		w/o WH	CC	16.08	0.23	70.61	-0.80	0.15	-5.33	0.01	0.35	0.01	-0.99	0.23	-4.28	.	.	.
			AA	16.08	0.21	75.53	-0.80	0.15	-5.36	-0.01	0.30	-0.03	-0.98	0.23	-4.33	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; AA, all-available-case analysis.
[a] True values used when generating the data.

Table 4.3: ANCOVA – 20% Missingness

Scenario	Method			Results											
				Intercept			Pre			Treatment			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth[a]				3.62			0.75			-1			-0.49		
MCAR	$n = 35$	w/ WH	CC	3.56	1.73	2.09	0.75	0.09	8.42	-0.98	0.36	-2.74	-0.45	0.56	-0.82
			MI	3.57	1.80	2.03	0.75	0.10	8.10	-0.97	0.37	-2.64	-0.45	0.58	-0.80
		w/o WH	CC	2.75	1.41	1.98	0.78	0.09	9.23	-0.98	0.36	-2.75	.	.	.
			MI	2.76	1.47	1.91	0.78	0.09	8.84	-0.97	0.37	-2.64	.	.	.
	$n = 100$	w/ WH	CC	3.58	1.00	3.59	0.75	0.05	14.39	-0.99	0.21	-4.70	-0.49	0.32	-1.51
			MI	3.58	1.02	3.55	0.75	0.05	14.20	-0.99	0.21	-4.66	-0.49	0.33	-1.50
		w/o WH	CC	2.70	0.82	3.32	0.78	0.05	15.70	-0.99	0.21	-4.67	.	.	.
			MI	2.71	0.83	3.27	0.78	0.05	15.46	-0.99	0.21	-4.63	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	3.49	1.74	2.04	0.76	0.09	8.39	-1.00	0.37	-2.77	-0.42	0.58	-0.73
			MI	3.48	1.84	1.94	0.76	0.10	7.95	-1.00	0.39	-2.64	-0.42	0.62	-0.68
		w/o WH	CC	2.77	1.43	1.98	0.78	0.09	9.13	-1.00	0.36	-2.78	.	.	.
			MI	2.74	1.55	1.83	0.78	0.09	8.57	-1.00	0.39	-2.63	.	.	.
	$n = 100$	w/ WH	CC	3.61	1.01	3.59	0.75	0.05	14.21	-1.00	0.21	-4.72	-0.48	0.34	-1.44
			MI	3.61	1.03	3.53	0.75	0.05	13.89	-1.00	0.22	-4.58	-0.48	0.34	-1.40
		w/o WH	CC	2.79	0.83	3.36	0.78	0.05	15.41	-1.00	0.21	-4.70	.	.	.
			MI	2.75	0.86	3.22	0.78	0.05	15.00	-1.00	0.22	-4.56	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	3.53	1.74	2.08	0.76	0.09	8.41	-0.98	0.36	-2.74	-0.47	0.58	-0.83
			MI	3.53	1.82	1.99	0.76	0.10	8.01	-0.99	0.38	-2.64	-0.46	0.61	-0.78
		w/o WH	CC	2.72	1.41	1.97	0.78	0.09	9.24	-0.99	0.36	-2.75	.	.	.
			MI	2.72	1.50	1.86	0.78	0.09	8.73	-0.99	0.38	-2.63	.	.	.
	$n = 100$	w/ WH	CC	3.58	1.01	3.58	0.76	0.05	14.31	-1.00	0.21	-4.70	-0.49	0.34	-1.47
			MI	3.57	1.03	3.50	0.76	0.05	13.97	-0.99	0.22	-4.59	-0.49	0.34	-1.44
		w/o WH	CC	2.73	0.82	3.33	0.78	0.05	15.63	-0.99	0.21	-4.68	.	.	.
			MI	2.69	0.85	3.20	0.78	0.05	15.21	-0.99	0.22	-4.57	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	3.59	1.77	2.07	0.75	0.09	8.26	-0.99	0.37	-2.69	-0.48	0.59	-0.82
			MI	3.58	1.91	1.94	0.75	0.10	7.72	-0.99	0.40	-2.54	-0.47	0.64	-0.75
		w/o WH	CC	2.76	1.43	1.97	0.78	0.09	9.09	-0.98	0.37	-2.65	.	.	.
			MI	2.74	1.57	1.81	0.78	0.10	8.40	-0.99	0.40	-2.54	.	.	.
	$n = 100$	w/ WH	CC	3.62	1.03	3.55	0.75	0.05	14.06	-1.01	0.22	-4.63	-0.50	0.34	-1.47
			MI	3.60	1.07	3.40	0.75	0.06	13.53	-1.01	0.23	-4.51	-0.49	0.35	-1.39
		w/o WH	CC	2.74	0.83	3.30	0.78	0.05	15.37	-0.99	0.22	-4.52	.	.	.
			MI	2.72	0.87	3.15	0.78	0.05	14.77	-1.01	0.23	-4.48	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; MI, multiple imputation.

[a] True values used when generating the data.

Chapter 5: Discussion

In the BePHIT data analyses, p-values were smaller for ANCOVA than for mixed models when testing the treatment effects. Also, for both ANCOVA and mixed models, p-values were smaller for all-available-data analyses than CC analyses. These results were expected since they were consistent with the theory presented in Chapter 2. However, the treatment effects estimated by ANCOVA after MI and CC ANCOVA were remarkably different (-0.08 for CC vs. -0.24 for MI), as were the standard errors (0.38 for CC vs. 0.63 for MI). It is also interesting that the p-value from CC ANCOVA was smaller than the p-value from the all-available-data mixed model. Thus, for the BePHIT data, the benefit of switching the CC mixed model to the CC ANCOVA model outweighed the benefit of switching the CC mixed model to the all-available-data mixed model.

The simulation studies showed that all four methods, ANCOVA after multiple imputation (MI), ANCOVA for complete cases (CC), the all-available-case mixed model, and the CC mixed model, were generally unbiased when waist-hip ratio was included in the analysis model. The power of ANCOVA after MI dropped the fastest as the proportion of missing data increased. In almost all simulated scenarios, ANCOVA after MI was the least powerful method when 50% of the data were missing.

As expected, the simulation studies showed that ANCOVA was more powerful than the mixed model for CC analyses, and the mixed model became more powerful when all available data were used. However, it was not expected that CC ANCOVA had the largest power under almost all scenarios. Similar estimated results were also observed, though not included in the appendix, when the imputation model was set to the same as the analysis model under the scenario of MAR dependent on waist-hip ratio when 20% of the data were missing.

The fact that ANCOVA after MI was consistently less powerful than CC ANCOVA suggested that MI was not gaining any extra information on ANCOVA. When data were missing at 20%, the averaged r-squared of the imputation model was consistently greater than 60% under each scenario, which indicates that the imputation model did a moderately good job in explaining the post-test values. Under each scenario for 20% missingness, the average fraction missing information (FMI) for MI was less than 10%.

The results from the simulation studies were not consistent with the analyses of the BePHIT data. The simulation studies suggested that ANCOVA after MI was almost always the least powerful method for testing the treatment effect, while this approach resulted in the smallest p-value in the BePHIT analysis.

This discrepancy may be the result of using different imputation models between the simulation studies and the BePHIT analyses. The imputation model in the simulations included only the variables used in generating the data and their two-way interactions, while the imputation model for the BePHIT studies also included some baseline anthropometric and psychometric measures. This resulted in the difference in the r-squared of the imputation model for the two studies. The r-squared was 0.6975

for the BePHIT study, while the averaged r-squared was smaller for each simulated scenario when 20% of the data were missing (between 0.61 and 0.65).

5.1 Limitations

The simulations were based on the BePHIT data. Therefore, the results from the simulation studies may not be generalized to other pre-post studies. When analyzing both the BePHIT data and the simulated data, the imputation models for MI and the variance-covariance structure in the mixed models were arbitrarily assumed. For the BePHIT data, sensitivity analysis for the MAR assumption, which is assumed by both MI and the mixed model, was not conducted, either.

Only four missing data mechanisms were considered in the simulations. These scenarios were chosen based on the analyses of the BePHIT study. However, many other missing mechanisms are possible in biomedical studies. For example, if the subject dropped out the study because she knew that her post-test one-mile walk would be very bad, the data would be not missing at random (NMAR). This is a very reasonable way for subjects to drop out. However, this scenario, along with many other scenarios of missing data were not considered.

Since we did not generate any baseline measures (other than the outcome and waist-hip ratio) for the simulation data, the imputation model was not the same for the BePHIT analyses and the simulation studies. Also, the variance-covariance structure for the mixed models were different. These differences could explain the differences in results.

5.2 Future Work

Further investigations are needed to explain the result that ANCOVA after MI had the least power among all four methods used in the simulation studies. For the next step of this study, the imputation models used for the BePHIT study and the simulation studies need to be evaluated further. In the simulation studies, more baseline measures may need to be generated to use as predictors in imputation models and analysis models.

Appendix A: Additional Simulation Results

This section consists of two parts. The first part contains the figures of power comparisons when waist-hip ratio was not included in the analysis models. The second part contains the simulation results when data were missing at 30%, 40%, and 50% under each scenario.

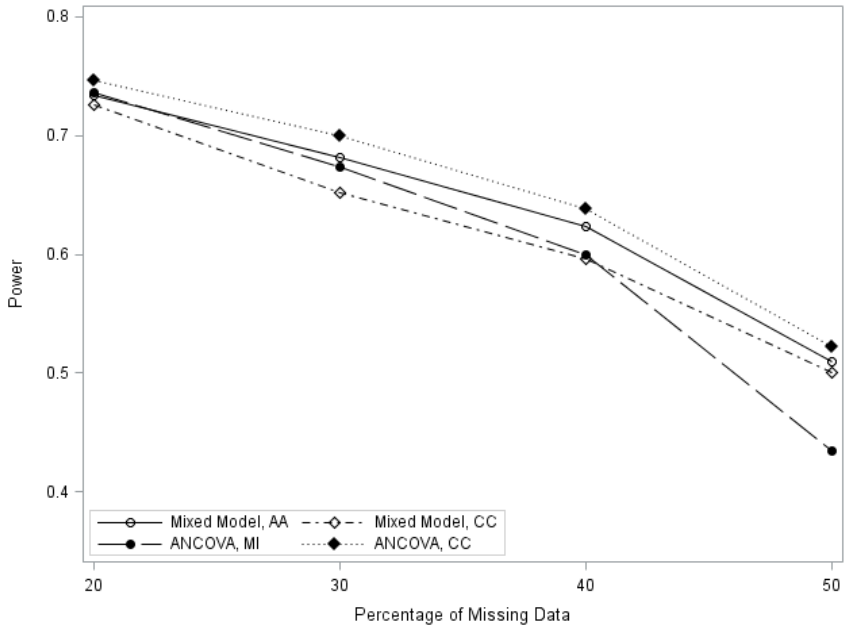


Figure A.1: Power Comparison under MCAR, $n = 35$, WH not in Analysis Models

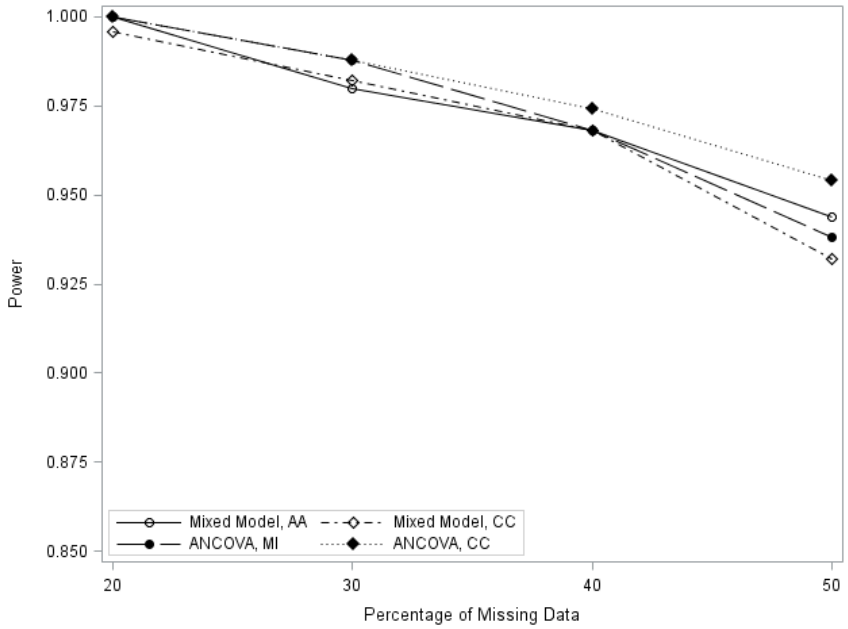


Figure A.2: Power Comparison under MCAR, $n = 100$, WH not in Analysis Models

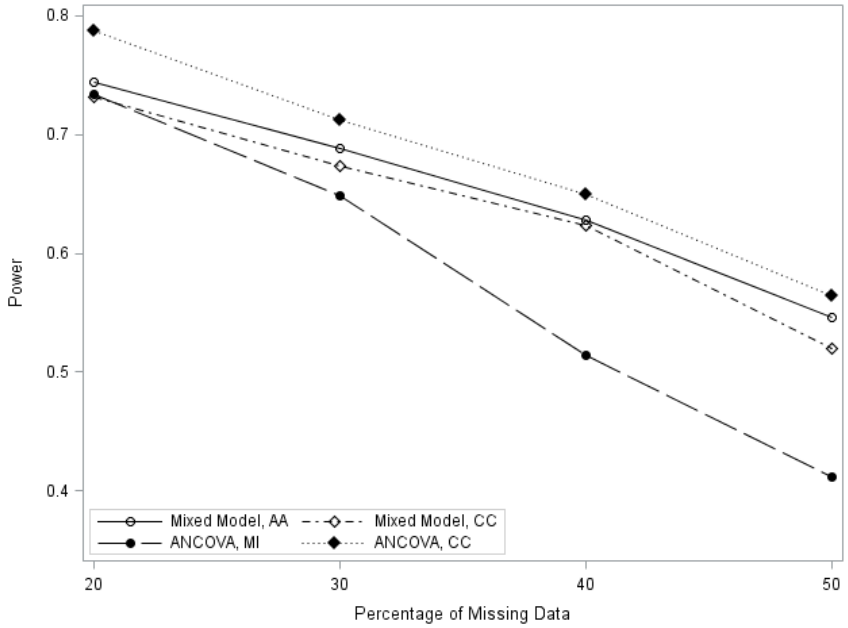


Figure A.3: Power Comparison under MAR (WH), $n = 35$, WH not in Analysis Models

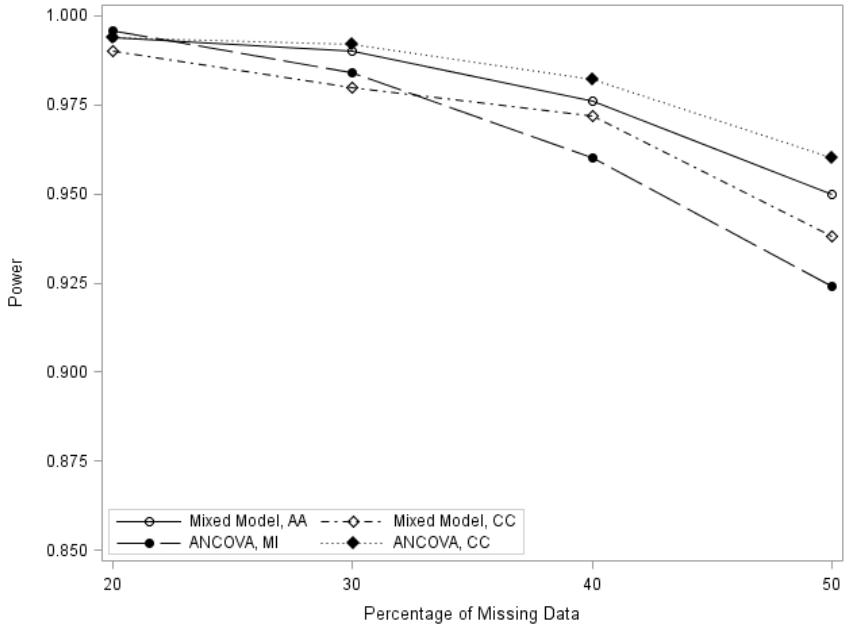


Figure A.4: Power Comparison under MAR (WH), $n = 100$, WH not in Analysis Models

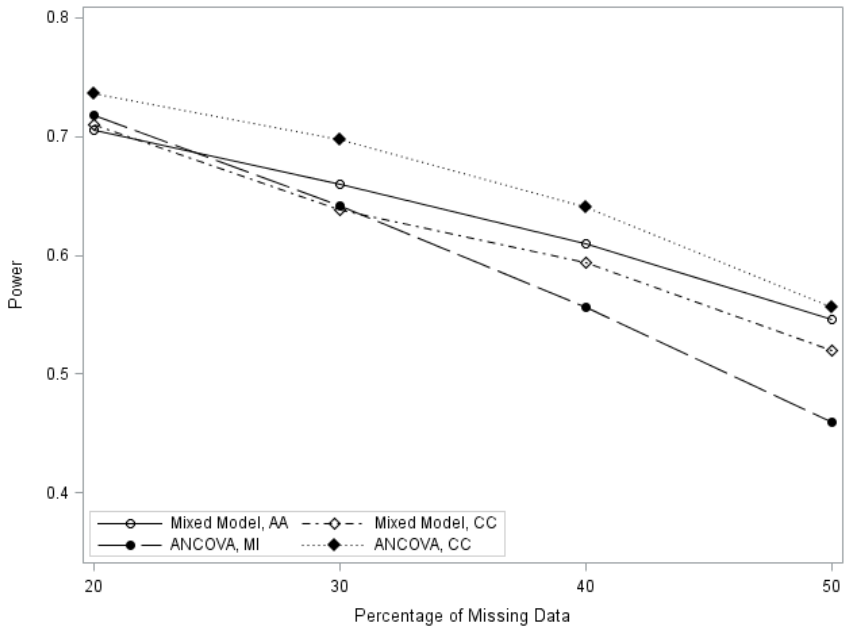


Figure A.5: Power Comparison under MAR (WH + Pre), $n = 35$, WH not in Analysis Models

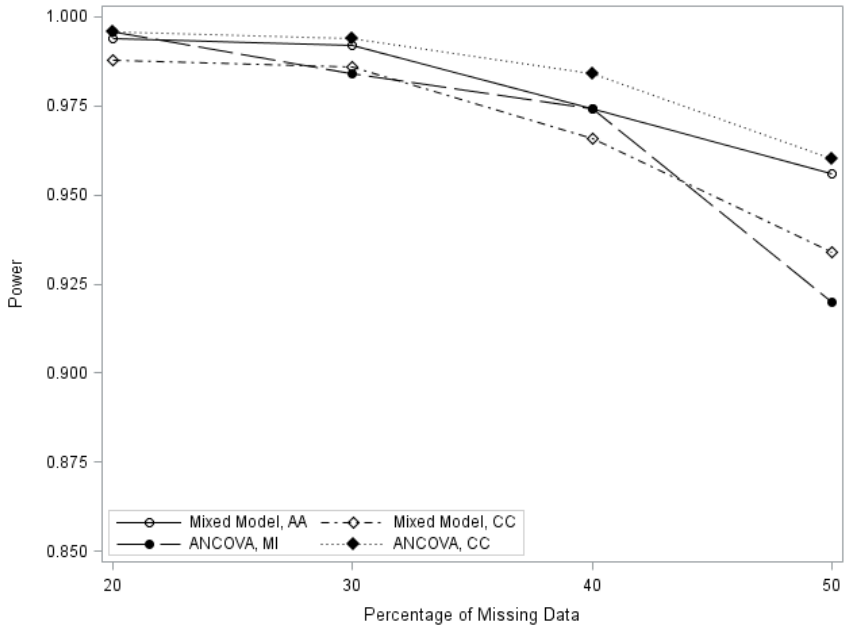


Figure A.6: Power Comparison under MAR (WH + Pre), $n = 100$, WH not in Analysis Models

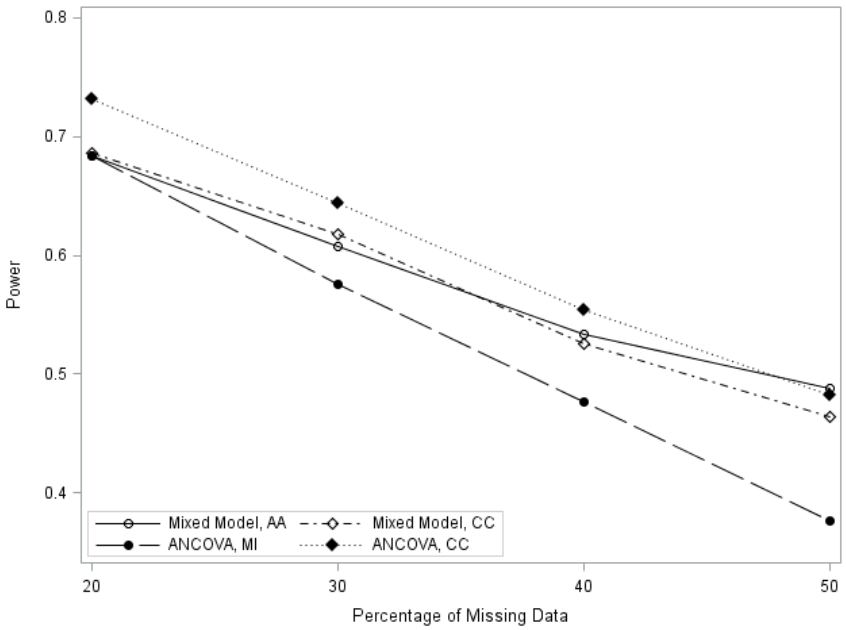


Figure A.7: Power Comparison under MAR (WH + Pre + Trt), $n = 35$, WH not in Analysis Models

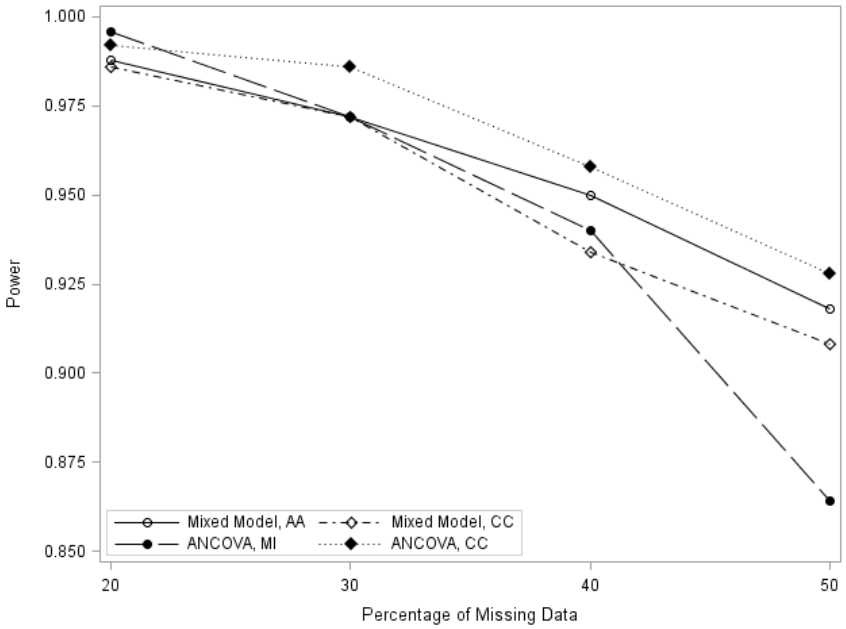


Figure A.8: Power Comparison under MAR (WH + Pre + Trt), $n = 100$, WH not in Analysis Models

Table A.1: Mixed Model – 30% Missingness

Scenario	Method			Results															
				Intercept			Treatment			Post			Trt × Post			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth				17.95	0.78	23.12	-0.81	0.27	-3.01	0			-1			-2		
MCAR	$n = 35$	w/ WH	CC	17.91	0.85	21.57	-0.82	0.28	-2.94	-0.01	0.58	-0.01	-0.98	0.40	-2.46	-1.96	0.80	-2.51
			AA	17.95	0.72	25.28	-0.83	0.28	-2.98	0.00	0.49	-0.00	-0.98	0.40	-2.51	-2.00	0.68	-2.99
		w/o WH	CC	16.08	0.43	37.89	-0.82	0.28	-2.94	-0.01	0.61	-0.02	-0.98	0.40	-2.46	.	.	.
			AA	16.08	0.36	44.84	-0.82	0.28	-2.97	0.00	0.51	-0.00	-0.98	0.40	-2.51	.	.	.
	$n = 100$	w/ WH	CC	17.96	0.50	36.41	-0.82	0.17	-4.85	-0.01	0.34	-0.02	-0.99	0.24	-4.16	-2.00	0.46	-4.37
			AA	17.96	0.42	42.89	-0.82	0.17	-4.94	-0.01	0.29	-0.04	-0.99	0.24	-4.23	-2.00	0.39	-5.11
		w/o WH	CC	16.08	0.26	63.01	-0.82	0.17	-4.85	-0.01	0.36	-0.02	-0.99	0.24	-4.16	.	.	.
			AA	16.08	0.21	75.52	-0.82	0.17	-4.93	-0.01	0.30	-0.03	-0.99	0.24	-4.22	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	17.95	0.84	21.74	-0.81	0.29	-2.85	-0.01	0.59	-0.03	-0.99	0.41	-2.46	-2.00	0.85	-2.41
			AA	17.95	0.72	25.17	-0.81	0.28	-2.90	0.00	0.49	0.00	-1.00	0.40	-2.51	-2.00	0.68	-2.98
		w/o WH	CC	16.23	0.44	37.61	-0.81	0.29	-2.85	-0.02	0.62	-0.04	-0.99	0.41	-2.46	.	.	.
			AA	16.08	0.36	44.82	-0.78	0.28	-2.78	0.00	0.51	-0.00	-1.00	0.40	-2.51	.	.	.
	$n = 100$	w/ WH	CC	17.95	0.49	36.88	-0.81	0.17	-4.76	-0.00	0.35	-0.00	-1.00	0.24	-4.16	-2.00	0.49	-4.12
			AA	17.96	0.42	42.76	-0.81	0.17	-4.85	-0.01	0.29	-0.04	-1.00	0.24	-4.23	-2.01	0.40	-5.10
		w/o WH	CC	16.22	0.26	63.21	-0.81	0.17	-4.76	-0.00	0.36	-0.01	-1.00	0.24	-4.16	.	.	.
			AA	16.08	0.21	75.60	-0.78	0.17	-4.66	-0.01	0.30	-0.03	-1.00	0.24	-4.22	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	17.93	0.84	21.92	-0.80	0.29	-2.82	0.01	0.59	0.01	-0.98	0.41	-2.44	-2.09	0.83	-2.58
			AA	17.96	0.72	25.26	-0.82	0.28	-2.95	0.00	0.49	0.00	-0.98	0.40	-2.48	-2.01	0.68	-3.00
		w/o WH	CC	16.10	0.44	37.41	-0.80	0.29	-2.82	0.01	0.62	0.01	-0.98	0.41	-2.44	.	.	.
			AA	16.08	0.36	44.85	-0.79	0.28	-2.85	0.00	0.51	-0.00	-0.98	0.40	-2.48	.	.	.
	$n = 100$	w/ WH	CC	17.95	0.49	37.07	-0.78	0.17	-4.61	-0.01	0.34	-0.01	-1.00	0.24	-4.17	-2.12	0.48	-4.41
			AA	17.96	0.42	42.77	-0.81	0.17	-4.85	-0.01	0.29	-0.04	-1.00	0.24	-4.24	-2.01	0.39	-5.10
		w/o WH	CC	16.09	0.26	62.68	-0.78	0.17	-4.61	-0.01	0.36	-0.01	-1.00	0.24	-4.17	.	.	.
			AA	16.08	0.21	75.52	-0.78	0.17	-4.67	-0.01	0.30	-0.03	-1.00	0.24	-4.23	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	17.99	0.85	21.64	-0.81	0.27	-3.08	-0.09	0.61	-0.14	-0.98	0.42	-2.35	-2.11	0.85	-2.56
			AA	17.97	0.72	25.22	-0.83	0.26	-3.17	0.00	0.49	-0.00	-1.00	0.42	-2.43	-2.02	0.68	-3.01
		w/o WH	CC	16.10	0.40	40.36	-0.81	0.27	-3.08	0.03	0.64	0.04	-0.98	0.42	-2.35	.	.	.
			AA	16.08	0.36	44.81	-0.81	0.26	-3.09	0.00	0.51	-0.00	-0.97	0.42	-2.37	.	.	.
	$n = 100$	w/ WH	CC	17.99	0.49	36.86	-0.80	0.16	-5.09	-0.10	0.36	-0.29	-0.99	0.25	-3.97	-2.11	0.49	-4.37
			AA	17.96	0.42	42.79	-0.81	0.16	-5.26	-0.01	0.29	-0.04	-1.01	0.24	-4.14	-2.01	0.40	-5.11
		w/o WH	CC	16.09	0.24	67.92	-0.80	0.16	-5.09	0.01	0.38	0.03	-0.99	0.25	-3.97	.	.	.
			AA	16.08	0.21	75.56	-0.80	0.16	-5.13	-0.01	0.30	-0.03	-0.98	0.24	-4.03	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; AA, all-available-case analysis.
[a] True values used when generating the data.

Table A.2: ANCOVA – 30% Missingness

Scenario	Method			Results											
				Intercept			Pre			Treatment			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth				3.62			0.75			-1			-0.49		
MCAR	$n = 35$	w/ WH	CC	3.51	1.86	1.93	0.76	0.10	7.88	-0.99	0.39	-2.59	-0.43	0.60	-0.73
			MI	3.52	1.98	1.82	0.76	0.10	7.44	-0.99	0.41	-2.47	-0.43	0.64	-0.70
		w/o WH	CC	2.74	1.51	1.85	0.78	0.09	8.64	-0.99	0.39	-2.61	.	.	.
			MI	2.74	1.63	1.72	0.78	0.10	8.08	-0.99	0.41	-2.47	.	.	.
	$n = 100$	w/ WH	CC	3.57	1.07	3.36	0.75	0.06	13.47	-1.00	0.23	-4.42	-0.48	0.35	-1.39
			MI	3.58	1.10	3.30	0.75	0.06	13.17	-1.00	0.23	-4.34	-0.48	0.35	-1.38
		w/o WH	CC	2.71	0.88	3.12	0.78	0.05	14.69	-0.99	0.23	-4.40	.	.	.
			MI	2.73	0.90	3.05	0.78	0.05	14.33	-0.99	0.23	-4.32	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	3.46	1.88	1.90	0.76	0.10	7.84	-1.00	0.39	-2.59	-0.44	0.64	-0.70
			MI	3.46	2.05	1.76	0.76	0.11	7.21	-1.00	0.43	-2.39	-0.43	0.69	-0.63
		w/o WH	CC	2.74	1.54	1.83	0.78	0.09	8.55	-1.00	0.39	-2.60	.	.	.
			MI	2.67	1.75	1.62	0.78	0.11	7.74	-1.00	0.43	-2.39	.	.	.
	$n = 100$	w/ WH	CC	3.61	1.08	3.37	0.75	0.06	13.33	-1.00	0.23	-4.41	-0.49	0.36	-1.35
			MI	3.61	1.12	3.25	0.75	0.06	12.76	-1.00	0.24	-4.23	-0.49	0.38	-1.32
		w/o WH	CC	2.80	0.89	3.14	0.78	0.05	14.44	-1.00	0.23	-4.41	.	.	.
			MI	2.73	0.95	2.90	0.78	0.06	13.71	-1.00	0.24	-4.21	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	3.57	1.87	1.96	0.75	0.10	7.84	-0.98	0.39	-2.56	-0.47	0.63	-0.78
			MI	3.57	2.02	1.82	0.75	0.11	7.24	-0.99	0.42	-2.39	-0.47	0.67	-0.72
		w/o WH	CC	2.76	1.51	1.86	0.78	0.09	8.64	-0.99	0.39	-2.57	.	.	.
			MI	2.74	1.68	1.69	0.78	0.10	7.88	-0.99	0.42	-2.38	.	.	.
	$n = 100$	w/ WH	CC	3.58	1.08	3.35	0.76	0.06	13.37	-1.00	0.23	-4.43	-0.49	0.36	-1.36
			MI	3.58	1.12	3.22	0.76	0.06	12.83	-1.00	0.24	-4.24	-0.49	0.37	-1.32
		w/o WH	CC	2.74	0.88	3.13	0.78	0.05	14.64	-1.00	0.23	-4.41	.	.	.
			MI	2.71	0.92	2.97	0.78	0.06	14.02	-0.99	0.24	-4.21	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	3.59	1.90	1.94	0.75	0.10	7.71	-1.01	0.41	-2.51	-0.49	0.64	-0.78
			MI	3.59	2.16	1.74	0.75	0.12	6.88	-1.00	0.45	-2.27	-0.48	0.72	-0.69
		w/o WH	CC	2.73	1.54	1.82	0.78	0.09	8.49	-0.98	0.40	-2.45	.	.	.
			MI	2.73	1.79	1.59	0.78	0.11	7.46	-1.00	0.46	-2.25	.	.	.
	$n = 100$	w/ WH	CC	3.64	1.09	3.34	0.75	0.06	13.16	-1.01	0.24	-4.31	-0.50	0.37	-1.39
			MI	3.63	1.17	3.14	0.75	0.06	12.38	-1.01	0.25	-4.09	-0.49	0.39	-1.29
		w/o WH	CC	2.75	0.89	3.10	0.78	0.05	14.41	-0.98	0.24	-4.19	.	.	.
			MI	2.74	0.96	2.89	0.78	0.06	13.47	-1.01	0.25	-4.07	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; AA, all-available-case analysis.

[a] True values used when generating the data.

Table A.3: Mixed Model – 40% Missingness

Scenario	Method			Results															
				Intercept			Treatment			Post			Trt × Post			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth				17.95	0.78	23.12	-0.81	0.27	-3.01	0			-1			-2		
MCAR	$n = 35$	w/ WH	CC	17.90	0.93	19.85	-0.83	0.31	-2.73	0.01	0.63	0.02	-0.98	0.44	-2.29	-1.96	0.87	-2.32
			AA	17.95	0.72	25.12	-0.83	0.30	-2.79	0.00	0.49	-0.00	-0.98	0.43	-2.33	-1.99	0.68	-2.96
		w/o WH	CC	16.07	0.47	35.02	-0.83	0.31	-2.73	0.01	0.66	0.01	-0.98	0.44	-2.29	.	.	.
			AA	16.08	0.36	44.81	-0.83	0.30	-2.78	0.00	0.51	-0.00	-0.98	0.43	-2.33	.	.	.
	$n = 100$	w/ WH	CC	17.96	0.54	33.73	-0.82	0.18	-4.49	-0.00	0.37	-0.01	-1.00	0.26	-3.86	-2.01	0.50	-4.05
			AA	17.96	0.42	42.66	-0.82	0.18	-4.60	-0.01	0.29	-0.04	-0.99	0.25	-3.95	-2.00	0.40	-5.07
		w/o WH	CC	16.08	0.28	58.39	-0.82	0.18	-4.49	-0.00	0.39	-0.01	-1.00	0.26	-3.86	.	.	.
			AA	16.08	0.21	75.50	-0.82	0.18	-4.59	-0.01	0.30	-0.03	-0.99	0.25	-3.94	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	17.97	0.90	20.49	-0.81	0.31	-2.67	-0.03	0.64	-0.05	-1.00	0.44	-2.30	-2.01	0.93	-2.24
			AA	17.96	0.73	25.07	-0.81	0.30	-2.72	0.00	0.49	0.00	-1.00	0.43	-2.36	-2.01	0.69	-2.98
		w/o WH	CC	16.27	0.47	35.22	-0.81	0.31	-2.67	-0.04	0.67	-0.06	-1.00	0.44	-2.30	.	.	.
			AA	16.08	0.36	44.82	-0.77	0.30	-2.58	0.00	0.51	-0.00	-1.00	0.43	-2.36	.	.	.
	$n = 100$	w/ WH	CC	17.95	0.52	34.69	-0.81	0.18	-4.42	-0.01	0.37	-0.01	-1.00	0.26	-3.86	-1.99	0.53	-3.79
			AA	17.96	0.42	42.56	-0.81	0.18	-4.53	-0.01	0.29	-0.04	-1.00	0.25	-3.95	-2.01	0.40	-5.07
		w/o WH	CC	16.26	0.28	58.88	-0.81	0.18	-4.42	-0.01	0.39	-0.02	-1.00	0.26	-3.86	.	.	.
			AA	16.08	0.21	75.59	-0.77	0.18	-4.30	-0.01	0.30	-0.03	-1.00	0.26	-3.94	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	17.90	0.89	20.64	-0.79	0.31	-2.60	0.01	0.63	0.01	-0.99	0.44	-2.29	-2.10	0.90	-2.40
			AA	17.96	0.72	25.16	-0.82	0.30	-2.76	0.00	0.49	0.00	-0.98	0.43	-2.34	-2.01	0.68	-2.99
		w/o WH	CC	16.11	0.47	34.87	-0.79	0.31	-2.60	0.00	0.66	-0.00	-0.99	0.44	-2.29	.	.	.
			AA	16.08	0.36	44.89	-0.78	0.30	-2.63	0.00	0.51	-0.00	-0.99	0.43	-2.34	.	.	.
	$n = 100$	w/ WH	CC	17.92	0.52	34.91	-0.77	0.18	-4.24	-0.00	0.37	-0.00	-1.00	0.26	-3.87	-2.13	0.52	-4.11
			AA	17.96	0.42	42.59	-0.81	0.18	-4.53	-0.01	0.29	-0.04	-1.00	0.25	-3.95	-2.01	0.40	-5.08
		w/o WH	CC	16.10	0.28	58.35	-0.77	0.18	-4.24	-0.00	0.39	-0.01	-1.00	0.26	-3.87	.	.	.
			AA	16.08	0.21	75.53	-0.77	0.18	-4.31	-0.01	0.30	-0.03	-1.00	0.25	-3.95	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	17.95	0.90	20.46	-0.81	0.28	-2.94	-0.12	0.67	-0.17	-0.96	0.46	-2.11	-2.10	0.92	-2.38
			AA	17.97	0.73	25.13	-0.83	0.28	-3.06	0.00	0.49	0.00	-0.98	0.45	-2.21	-2.02	0.68	-2.99
		w/o WH	CC	16.10	0.42	38.57	-0.81	0.28	-2.94	0.01	0.70	0.02	-0.96	0.46	-2.11	.	.	.
			AA	16.08	0.36	44.85	-0.80	0.28	-2.96	0.00	0.51	-0.00	-0.95	0.45	-2.15	.	.	.
	$n = 100$	w/ WH	CC	17.96	0.52	34.83	-0.79	0.17	-4.81	-0.10	0.39	-0.26	-0.99	0.27	-3.68	-2.11	0.52	-4.07
			AA	17.96	0.42	42.61	-0.81	0.16	-5.03	-0.01	0.29	-0.04	-1.01	0.26	-3.85	-2.01	0.40	-5.08
		w/o WH	CC	16.09	0.25	64.76	-0.79	0.17	-4.81	0.03	0.41	0.07	-0.99	0.27	-3.68	.	.	.
			AA	16.08	0.21	75.57	-0.79	0.16	-4.87	-0.01	0.30	-0.03	-0.98	0.26	-3.73	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; AA, all-available-case analysis.
[a] True values used when generating the data.

Table A.4: ANCOVA – 40% Missingness

Scenario	Method			Results											
				Intercept			Pre			Treatment			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth				3.62			0.75			-1			-0.49		
MCAR	$n = 35$	w/ WH	CC	3.44	2.03	1.74	0.76	0.11	7.27	-0.99	0.42	-2.39	-0.41	0.66	-0.63
			MI	3.47	2.23	1.61	0.76	0.12	6.68	-0.98	0.46	-2.22	-0.40	0.72	-0.58
		w/o WH	CC	2.72	1.64	1.70	0.78	0.10	7.98	-0.99	0.42	-2.41	.	.	.
			MI	2.75	1.84	1.55	0.78	0.11	7.23	-0.98	0.46	-2.21	.	.	.
	$n = 100$	w/ WH	CC	3.56	1.17	3.07	0.76	0.06	12.42	-1.00	0.25	-4.09	-0.47	0.37	-1.26
			MI	3.57	1.21	2.99	0.75	0.06	12.08	-1.00	0.25	-3.99	-0.47	0.39	-1.23
		w/o WH	CC	2.71	0.95	2.87	0.78	0.06	13.57	-1.00	0.25	-4.07	.	.	.
			MI	2.72	0.98	2.81	0.78	0.06	13.18	-1.00	0.25	-3.97	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	3.50	2.04	1.79	0.76	0.11	7.26	-1.01	0.42	-2.42	-0.47	0.69	-0.69
			MI	3.47	2.31	1.59	0.76	0.12	6.41	-1.00	0.48	-2.15	-0.46	0.78	-0.61
		w/o WH	CC	2.75	1.67	1.70	0.78	0.10	7.93	-1.01	0.42	-2.43	.	.	.
			MI	2.65	2.01	1.42	0.79	0.12	6.84	-1.00	0.49	-2.14	.	.	.
	$n = 100$	w/ WH	CC	3.61	1.16	3.13	0.75	0.06	12.36	-1.00	0.25	-4.09	-0.48	0.39	-1.24
			MI	3.60	1.23	2.99	0.75	0.07	11.65	-1.00	0.26	-3.83	-0.48	0.41	-1.19
		w/o WH	CC	2.81	0.97	2.93	0.78	0.06	13.39	-1.00	0.25	-4.09	.	.	.
			MI	2.73	1.05	2.65	0.78	0.06	12.46	-1.00	0.27	-3.82	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	3.60	2.01	1.84	0.75	0.11	7.28	-0.99	0.42	-2.40	-0.48	0.68	-0.72
			MI	3.57	2.28	1.64	0.76	0.12	6.49	-1.00	0.47	-2.19	-0.48	0.75	-0.67
		w/o WH	CC	2.80	1.63	1.76	0.78	0.10	8.04	-0.99	0.42	-2.42	.	.	.
			MI	2.73	1.91	1.49	0.78	0.12	7.05	-1.00	0.47	-2.18	.	.	.
	$n = 100$	w/ WH	CC	3.61	1.16	3.13	0.75	0.06	12.39	-1.00	0.25	-4.11	-0.50	0.39	-1.28
			MI	3.60	1.22	2.99	0.75	0.07	11.77	-1.00	0.26	-3.86	-0.50	0.40	-1.25
		w/o WH	CC	2.76	0.95	2.93	0.78	0.06	13.58	-1.00	0.25	-4.09	.	.	.
			MI	2.71	1.02	2.71	0.78	0.06	12.82	-1.00	0.26	-3.84	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	3.61	2.05	1.81	0.75	0.11	7.16	-0.99	0.45	-2.26	-0.49	0.69	-0.73
			MI	3.57	2.50	1.53	0.75	0.13	6.04	-0.98	0.52	-1.97	-0.47	0.83	-0.60
		w/o WH	CC	2.74	1.66	1.71	0.78	0.10	7.92	-0.96	0.44	-2.21	.	.	.
			MI	2.73	2.07	1.42	0.78	0.13	6.56	-0.98	0.52	-1.96	.	.	.
	$n = 100$	w/ WH	CC	3.64	1.17	3.14	0.75	0.06	12.29	-1.01	0.26	-3.98	-0.51	0.39	-1.30
			MI	3.64	1.28	2.89	0.75	0.07	11.27	-1.01	0.27	-3.72	-0.49	0.42	-1.18
		w/o WH	CC	2.76	0.95	2.92	0.78	0.06	13.46	-0.98	0.26	-3.86	.	.	.
			MI	2.75	1.06	2.65	0.78	0.07	12.25	-1.01	0.28	-3.69	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; AA, all-available-case analysis.

[a] True values used when generating the data.

Table A.5: Mixed Model – 50% Missingness

Scenario	Method			Results															
				Intercept			Treatment			Post			Trt × Post			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth				17.95	0.78	23.12	-0.81	0.27	-3.01	0			-1			-2		
MCAR	$n = 35$	w/ WH	CC	17.88	1.02	18.16	-0.83	0.34	-2.52	0.01	0.70	0.01	-0.98	0.48	-2.08	-1.94	0.96	-2.10
			AA	17.95	0.73	24.99	-0.84	0.33	-2.59	0.00	0.49	0.00	-0.97	0.47	-2.13	-1.99	0.69	-2.94
		w/o WH	CC	16.08	0.51	32.07	-0.83	0.34	-2.52	0.00	0.72	0.00	-0.98	0.48	-2.08	.	.	.
			AA	16.08	0.36	44.81	-0.83	0.33	-2.58	0.00	0.51	-0.00	-0.97	0.47	-2.12	.	.	.
	$n = 100$	w/ WH	CC	17.95	0.59	30.79	-0.82	0.20	-4.11	0.01	0.41	0.01	-0.99	0.28	-3.53	-2.00	0.55	-3.68
			AA	17.96	0.43	42.45	-0.82	0.19	-4.24	-0.01	0.29	-0.04	-0.99	0.27	-3.63	-2.00	0.40	-5.04
		w/o WH	CC	16.07	0.30	53.37	-0.82	0.20	-4.11	0.01	0.43	0.02	-0.99	0.28	-3.53	.	.	.
			AA	16.08	0.21	75.50	-0.82	0.20	-4.23	-0.01	0.30	-0.03	-0.99	0.28	-3.61	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	18.00	0.97	19.22	-0.83	0.34	-2.52	-0.05	0.70	-0.07	-0.98	0.48	-2.08	-2.04	1.02	-2.08
			AA	17.96	0.73	24.96	-0.83	0.33	-2.57	0.00	0.49	0.00	-0.99	0.47	-2.16	-2.01	0.69	-2.97
		w/o WH	CC	16.32	0.51	32.58	-0.83	0.34	-2.52	-0.06	0.73	-0.08	-0.98	0.48	-2.08	.	.	.
			AA	16.08	0.36	44.79	-0.78	0.33	-2.42	0.00	0.51	-0.00	-0.99	0.47	-2.16	.	.	.
	$n = 100$	w/ WH	CC	17.94	0.56	32.45	-0.81	0.20	-4.05	0.01	0.41	0.02	-1.01	0.28	-3.58	-2.00	0.58	-3.49
			AA	17.96	0.43	42.40	-0.81	0.20	-4.19	-0.01	0.29	-0.04	-1.00	0.28	-3.67	-2.01	0.40	-5.04
		w/o WH	CC	16.29	0.30	54.29	-0.81	0.20	-4.05	0.00	0.43	0.00	-1.01	0.28	-3.58	.	.	.
			AA	16.08	0.21	75.60	-0.76	0.20	-3.92	-0.01	0.30	-0.03	-1.00	0.28	-3.66	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	17.87	0.97	19.07	-0.78	0.34	-2.35	0.02	0.70	0.02	-0.99	0.48	-2.12	-2.11	1.00	-2.19
			AA	17.97	0.73	25.03	-0.82	0.33	-2.53	0.00	0.49	0.00	-0.99	0.46	-2.17	-2.02	0.69	-2.98
		w/o WH	CC	16.11	0.51	31.98	-0.78	0.34	-2.35	0.01	0.73	0.01	-0.99	0.48	-2.12	.	.	.
			AA	16.08	0.36	44.88	-0.77	0.33	-2.39	0.00	0.51	-0.00	-0.99	0.47	-2.16	.	.	.
	$n = 100$	w/ WH	CC	17.90	0.56	32.43	-0.76	0.20	-3.85	-0.00	0.40	-0.00	-1.00	0.28	-3.56	-2.15	0.57	-3.80
			AA	17.96	0.43	42.39	-0.81	0.19	-4.18	-0.01	0.29	-0.04	-1.00	0.27	-3.66	-2.01	0.40	-5.05
		w/o WH	CC	16.10	0.30	53.62	-0.76	0.20	-3.85	-0.00	0.43	-0.00	-1.00	0.28	-3.56	.	.	.
			AA	16.08	0.21	75.51	-0.76	0.19	-3.93	-0.01	0.30	-0.03	-1.00	0.27	-3.65	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	17.93	0.97	19.13	-0.80	0.29	-2.79	-0.11	0.74	-0.14	-0.96	0.51	-1.93	-2.10	1.00	-2.18
			AA	17.97	0.73	25.02	-0.83	0.29	-2.93	0.00	0.49	0.00	-0.98	0.49	-2.03	-2.02	0.69	-2.98
		w/o WH	CC	16.10	0.45	36.51	-0.80	0.29	-2.79	0.03	0.77	0.05	-0.96	0.51	-1.93	.	.	.
			AA	16.08	0.36	44.84	-0.80	0.29	-2.82	0.00	0.51	-0.00	-0.95	0.49	-1.97	.	.	.
	$n = 100$	w/ WH	CC	17.95	0.56	32.68	-0.79	0.17	-4.55	-0.11	0.43	-0.25	-0.98	0.30	-3.33	-2.12	0.57	-3.77
			AA	17.97	0.43	42.44	-0.82	0.17	-4.81	-0.01	0.29	-0.04	-1.01	0.29	-3.51	-2.01	0.40	-5.06
		w/o WH	CC	16.10	0.26	61.24	-0.79	0.17	-4.55	0.03	0.45	0.07	-0.98	0.30	-3.33	.	.	.
			AA	16.08	0.21	75.55	-0.79	0.17	-4.61	-0.01	0.30	-0.03	-0.98	0.29	-3.40	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; AA, all-available-case analysis.
[a] True values used when generating the data.

Table A.6: ANCOVA – 50% Missingness

Scenario	Method			Results											
				Intercept			Pre			Treatment			Waist-hip Ratio		
				Est.	SE	t	Est.	SE	t	Est.	SE	t	Est.	SE	t
Truth				3.62			0.75			-1			-0.49		
MCAR	$n = 35$	w/ WH	CC	3.40	2.24	1.57	0.76	0.12	6.63	-0.98	0.47	-2.16	-0.39	0.72	-0.57
			MI	3.44	2.57	1.41	0.76	0.14	5.87	-0.97	0.52	-1.94	-0.40	0.84	-0.51
		w/o WH	CC	2.70	1.81	1.53	0.78	0.11	7.30	-0.98	0.46	-2.18	.	.	.
			MI	2.73	2.13	1.35	0.78	0.13	6.34	-0.97	0.52	-1.94	.	.	.
	$n = 100$	w/ WH	CC	3.53	1.28	2.79	0.76	0.07	11.37	-1.00	0.27	-3.73	-0.47	0.41	-1.15
			MI	3.52	1.34	2.69	0.76	0.07	10.96	-0.99	0.28	-3.62	-0.46	0.43	-1.11
		w/o WH	CC	2.69	1.04	2.61	0.78	0.06	12.43	-0.99	0.27	-3.72	.	.	.
			MI	2.69	1.09	2.51	0.78	0.07	11.93	-0.99	0.28	-3.60	.	.	.
MAR (WH)	$n = 35$	w/ WH	CC	3.46	2.24	1.62	0.76	0.12	6.65	-0.99	0.47	-2.19	-0.47	0.76	-0.64
			MI	3.47	2.70	1.38	0.76	0.15	5.58	-0.99	0.56	-1.86	-0.46	0.89	-0.54
		w/o WH	CC	2.70	1.84	1.53	0.78	0.11	7.29	-0.99	0.46	-2.20	.	.	.
			MI	2.64	2.37	1.23	0.78	0.14	5.96	-0.99	0.56	-1.84	.	.	.
	$n = 100$	w/ WH	CC	3.63	1.26	2.90	0.75	0.07	11.37	-1.01	0.27	-3.78	-0.49	0.43	-1.14
			MI	3.65	1.36	2.74	0.75	0.07	10.46	-1.00	0.29	-3.47	-0.49	0.45	-1.10
		w/o WH	CC	2.83	1.06	2.71	0.78	0.06	12.33	-1.01	0.27	-3.78	.	.	.
			MI	2.76	1.19	2.39	0.78	0.07	11.15	-1.00	0.29	-3.46	.	.	.
MAR (WH + Pre)	$n = 35$	w/ WH	CC	3.61	2.20	1.69	0.75	0.12	6.67	-0.98	0.46	-2.19	-0.49	0.75	-0.67
			MI	3.61	2.62	1.47	0.75	0.14	5.71	-0.99	0.54	-1.91	-0.50	0.87	-0.61
		w/o WH	CC	2.79	1.79	1.61	0.78	0.11	7.38	-0.99	0.46	-2.22	.	.	.
			MI	2.74	2.23	1.31	0.78	0.14	6.16	-0.99	0.55	-1.90	.	.	.
	$n = 100$	w/ WH	CC	3.59	1.27	2.87	0.76	0.07	11.38	-1.00	0.27	-3.77	-0.50	0.43	-1.17
			MI	3.55	1.36	2.67	0.76	0.07	10.59	-1.00	0.29	-3.47	-0.49	0.45	-1.11
		w/o WH	CC	2.74	1.04	2.68	0.78	0.06	12.49	-1.00	0.27	-3.76	.	.	.
			MI	2.67	1.14	2.40	0.78	0.07	11.49	-1.00	0.29	-3.45	.	.	.
MAR (WH + Pre + Trt)	$n = 35$	w/ WH	CC	3.56	2.22	1.66	0.76	0.12	6.65	-0.98	0.49	-2.03	-0.50	0.75	-0.68
			MI	3.52	2.93	1.32	0.76	0.16	5.28	-0.98	0.61	-1.71	-0.49	0.97	-0.54
		w/o WH	CC	2.73	1.79	1.56	0.78	0.11	7.35	-0.95	0.48	-2.00	.	.	.
			MI	2.65	2.45	1.19	0.78	0.15	5.72	-0.98	0.62	-1.70	.	.	.
	$n = 100$	w/ WH	CC	3.66	1.26	2.91	0.75	0.07	11.37	-1.01	0.28	-3.60	-0.52	0.43	-1.22
			MI	3.64	1.44	2.57	0.75	0.08	10.03	-1.00	0.31	-3.27	-0.50	0.47	-1.08
		w/o WH	CC	2.76	1.03	2.69	0.78	0.06	12.47	-0.98	0.28	-3.49	.	.	.
			MI	2.73	1.20	2.34	0.78	0.07	10.90	-1.00	0.32	-3.25	.	.	.

Abbreviations: Est., averaged estimated coefficients; SE, averaged standard errors; t, averaged t-values; MCAR, missing completely at random; MAR (WH), missing at random dependent on waist-hip ratio; MAR (WH + Pre), missing at random dependent on waist-hip ratio and pre-test value; MAR (WH + Pre + Trt), missing at random dependent on waist-hip ratio, pre-test value, and treatment group; w/(o) WH, with(out) waist-hip ratio in the analysis model. CC, complete-case analysis; MI, multiple imputation.

[a] True values used when generating the data.

Appendix B: SAS Code

SAS Code for the BePHIT Analyses

```
libname data "u:\master thesis\data";
*data constructed in the file thesis_1018;

*datasets construction;
data test;
    if not eof then do;
    set data.walktest (firstobs = 2 rename = (date = post_date t = post_t)) end =
        eof;
    end;
    set data.walktest (rename = (date = pre_date t = pre_t));
    by subid pre_date pre_t;
    if last.subid then do;
    post_date = .;
    post_t = .;
    end;
    if not first.subid then delete;
    label pre_date = 'Pre Date' pre_t = 'Pre t' post_date = 'Post Date' post_t = '
        Post t';
run;

data data.test;
    retain subid pre_date pre_t post_date post_t;
    set test;
    if post_date ne . then days = post_date - pre_date;
run;

data data.walktest_post;
    set data.walktest;
    by subid;
    if not first.subid then post = 1;
    else post = 0;
run;

proc sort data = data.walktest_post out = walktest_post;
    by subid;
proc sort data = data.randomization out = randomization;
```

```
    by subid;
proc sort data = data.baseline_vitals out = baseline_vitals;
    by subid;
proc sort data = data.prepsych_wenna out = prepsych_wenna;
    by subid;
proc sort data = data.test out = test;
    by subid;
run;

data data.ancova;
    merge test(in = a) prepsych_wenna baseline_vitals randomization;
    by subid;
    if a;
    if intervention = 'A' then treatment = 1;
    else treatment = 0;
    if post_t = . then dropout = 1;
    else dropout = 0;
run;

data data.mixed;
    merge walktest_post(in = a) prepsych_wenna baseline_vitals randomization;
    by subid;
    if a;
    if intervention = 'A' then treatment = 1;
    else treatment = 0;
run;

*****cc & or data;
data data.ancova_cc;
    set data.ancova;
    where post_t ne .;
run;

data subid_cc;
    set data.ancova_cc;
    keep subid dropout;
run;
proc sort data = subid_cc;
    by subid;
proc sort data = data.mixed out = mixed_sorted;
    by subid;
run;

data data.mixed_cc;
    merge subid_cc (in = a) mixed_sorted;
    by subid;
    if a;
run;

*check;
proc sort data = data.mixed_cc out = mixed_cc_check nodupkey;
```

```
    by subid;
run;
data subid_mixed_cc;
    set mixed_cc_check;
    keep subid;
run;
proc compare data = subid_cc compare = subid_mixed_cc;
run;
*****end of cc & or data;


*data analyzed in the file bephitresults_0512;
*****All Available cases;
*mixed model;
proc mixed data = data.mixed;
    class subid treatment post;
    model t = treatment post treatment * post / s ddfm = kenwardroger;
    repeated post / type = un subject = subid;
    lsmeans treatment * post;
    estimate 'Delta' treatment * post 1 -1 -1 1;
run;

*ANCOVA;
proc mi data = data.ancova nimpute = 20 out = data.ancova_mi2;
    class treatment;
    monotone reg (post_t / details);
    var pre_t treatment prethoug preexsta presfam presfrnd prese preq16 pregls
        preplan pulse_rate waist_hip bmi post_t;
run;

proc reg data = data.ancova_mi2 outest = data.ancova_mi_reg2 covout noprint;
    model post_t = pre_t treatment;
    by _imputation_;
run;
quit;

proc mianalyze data = data.ancova_mi_reg2;
    modeleffects intercept pre_t treatment;
run;
*****end of AA;


*****CC;
*MIXED model;
proc mixed data = data.mixed_cc;
    class subid treatment post;
    model t = treatment post treatment * post / s ddfm = kenwardroger;
    repeated post / type = un subject = subid;
    lsmeans treatment * post;
    estimate 'Delta' treatment * post 1 -1 -1 1;
run;
```

```
*ANCOVA;
proc reg data = data.ancova_cc;
    model post_t = pre_t treatment;
run;
quit;
*****end of CC;

*****OR;
%macro or(var);
proc logistic data = data.ancova;
    class dropout (ref = "0");
    model dropout = &var.;
    ods select parameterestimates;
run;
quit;
%mend;

%or(pre_t);
%or(treatment);
%or(prethoug);
%or(preexsta);
%or(presfam);
%or(presfrnd);
%or(prese);
%or(preq16);
%or(pregls);
%or(preplan);
%or(pulse_rate);
%or(waist_hip);
%or(bmi);
*****end of OR;


*****Checking imputation model for BePHIT;
*imputation model in BePHIT;
proc reg data = data.ancova;
    model post_t = pre_t treatment prethoug preexsta presfam presfrnd prese preq16
         pregls preplan pulse_rate waist_hip bmi;
run;
quit;
*****end of checking imputation model for BePHIT;

proc means data = data.ancova;
   var post_t pre_t;
run;
```

Simulation Macros – File "simulation_macro_0125.sas"

```
/*Macros in this file*/
/*%library -> Call the libraries*/
/*%data20 -> Generate data set with 20% missingness*/
/*%data30 -> Generate data set with 30% missingness*/
/*%data40 -> Generate data set with 40% missingness*/
/*%data50 -> Generate data set with 50% missingness*/
/*%mixed -> Mixed model for fully observed data, complete cases, and available
    cases*/
/*%ancova -> ANCOVA for fully observed data, complete cases, and available cases (
    including MI)*/
/*%result -> Analyzing the data using either mixed model or ANCOVA*/

*library;
%macro library(mm = , dd = );
options DLCREATEDIR;
libname data "u:\master thesis\data"; /*library for experiment data*/
libname sim "u:\master thesis\simulation\sim&mm.&dd."; /*library for simulation
    data*/
%mend;

*Data Generation;
/*libname = library name*/
/*dsname = name for created data set*/
/*seed = starting seed value*/
/*nreps = number of replicates*/
/*nsub = number of subjects in each group*/
/*b0 = estimated coefficient for intercept*/
/*b1 = estimated coefficient for time*/
/*b2 = estimated coefficient for waist_hip*/
/*b3 = estimated coefficient for treatment*/
/*b4 = estimated coefficient for treatment * time*/
/*b5 = estimated coefficient for waist_hip * time*/
/*var_b = estimated variance for b_i*/
/*var_e = estimated variance for e_ij*/

*Generate data set with 20% missingness;
%macro data20(libname = , dsname = , seed = , nreps = , nsub = , b0 = , b1 = , b2
    = , b3 = , b4 = , b5 = , var_b = , var_e = );
data &dsname._&nsub.;
    retain nreps subid y post waist_hip treatment;
    call streaminit(&seed);
    do nreps = 1 to &nreps;
    do subid = 1 to 2 * &nsub;
    if subid <= &nsub then treatment = 0;
    else treatment = 1;
    waist_hip = rand('normal', 0.9381308532, 0.3485664653);
    b_i = rand('normal', 0, sqrt(&var_b));
    do post = 0 to 1;
    e_ij = rand('normal', 0, sqrt(&var_e));
    y = b_i + &b0 + &b1 * post + &b2 * waist_hip + &b3 * treatment + &b4 * (
        treatment * post) + &b5 * (waist_hip * post) + e_ij;
```

```
    output;
    end;
    end;
    end;
run;

/*Set 20% post scores to missing*/
data &libname..&dsname._&nsub.;
    set &dsname._&nsub.;
    call streaminit(32574435);
    y_mcar = y;
    y_mar_wh = y;
    y_mar_wh_pre = y;
    y_mar_wh_pre_tx = y;
    p_wh = exp(-3.251829 + 1.988565 * waist_hip) / (exp(-3.251829 + 1.988565 *
        waist_hip) + 1);
    p_wh_pre = exp(-4.754387 + 1.988565 * waist_hip + 0.0935 * y) / (exp(-4.754387
         + 1.988565 * waist_hip + 0.0935 * y) + 1);
    p_wh_pre_tx = exp(-5.447534 + 1.988565 * waist_hip + 0.0935 * y + 1.386294 *
        treatment) / (exp(-5.447534 + 1.988565 * waist_hip + 0.0935 * y + 1.386294
        * treatment) + 1);
    lag_p_wh_pre = lag(p_wh_pre);
    lag_p_wh_pre_tx = lag(p_wh_pre_tx);
    if post = 1 then do;
    mcar = rand('bernoulli', .2);
    if mcar = 1 then y_mcar = .;
    mar_wh = rand('bernoulli', p_wh);
    if mar_wh = 1 then y_mar_wh = .;
    mar_wh_pre = rand('bernoulli', lag_p_wh_pre);
    if mar_wh_pre = 1 then y_mar_wh_pre = .;
    mar_wh_pre_tx = rand('bernoulli', lag_p_wh_pre_tx);
    if mar_wh_pre_tx = 1 then y_mar_wh_pre_tx = .;
    end;
    rename waist_hip = wh;
run;

/*Construct dataset for MI*/
data test;
    if not eof then do;
    set &libname..&dsname._&nsub. (firstobs = 2 rename = (y = y_post_c y_mcar = y_
        post_mcar y_mar_wh = y_post_mar_wh y_mar_wh_pre = y_post_mar_wh_pre y_mar_
        wh_pre_tx = y_post_mar_wh_pre_tx e_ij = e_ij_post)) end = eof;
    end;
    set &libname..&dsname._&nsub. (rename = (y_mcar = y_pre e_ij = e_ij_pre));
    by nreps subid;
    if last.subid then y_post = . ;
    if not first.subid then delete;
run;
data &libname..&dsname._&nsub._ancova;
    retain nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post
        _mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
```

```
    set test;
    keep nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post_
        mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
run;
%mend;


*Generate data set with 30% missingness;
%macro data30(libname = , dsname = , seed = , nreps = , nsub = , b0 = , b1 = , b2
    = , b3 = , b4 = , b5 = , var_b = , var_e = );
data &dsname._&nsub.;
    retain nreps subid y post waist_hip treatment;
    call streaminit(&seed);
    do nreps = 1 to &nreps;
    do subid = 1 to 2 * &nsub;
    if subid <= &nsub then treatment = 0;
    else treatment = 1;
    waist_hip = rand('normal', 0.9381308532, 0.3485664653);
    b_i = rand('normal', 0, sqrt(&var_b));
    do post = 0 to 1;
    e_ij = rand('normal', 0, sqrt(&var_e));
    y = b_i + &b0 + &b1 * post + &b2 * waist_hip + &b3 * treatment + &b4 * (
        treatment * post) + &b5 * (waist_hip * post) + e_ij;
    output;
    end;
    end;
    end;
run;

/*Set 30% post scores to missing*/
data &libname..&dsname._&nsub.;
    set &dsname._&nsub.;
    call streaminit(32574435);
    y_mcar = y;
    y_mar_wh = y;
    y_mar_wh_pre = y;
    y_mar_wh_pre_tx = y;
    p_wh = exp(-2.712832 + 1.988565 * waist_hip) / (exp(-2.712832 + 1.988565 *
        waist_hip) + 1);
    p_wh_pre = exp(-4.21539 + 1.988565 * waist_hip + 0.0935 * y) / (exp(-4.21539 +
         1.988565 * waist_hip + 0.0935 * y) + 1);
    p_wh_pre_tx = exp(-4.908537 + 1.988565 * waist_hip + 0.0935 * y + 1.386294 *
        treatment) / (exp(-4.908537 + 1.988565 * waist_hip + 0.0935 * y + 1.386294
        * treatment) + 1);
    lag_p_wh_pre = lag(p_wh_pre);
    lag_p_wh_pre_tx = lag(p_wh_pre_tx);
    if post = 1 then do;
    mcar = rand('bernoulli', .3);
    if mcar = 1 then y_mcar = .;
    mar_wh = rand('bernoulli', p_wh);
    if mar_wh = 1 then y_mar_wh = .;
```

```
    mar_wh_pre = rand('bernoulli', lag_p_wh_pre);
    if mar_wh_pre = 1 then y_mar_wh_pre = .;
    mar_wh_pre_tx = rand('bernoulli', lag_p_wh_pre_tx);
    if mar_wh_pre_tx = 1 then y_mar_wh_pre_tx = .;
    end;
    rename waist_hip = wh;
run;

/*Construct dataset for MI*/
data test;
    if not eof then do;
    set &libname..&dsname._&nsub.
    (firstobs = 2 rename = (y = y_post_c y_mcar = y_post_mcar y_mar_wh = y_post_
        mar_wh y_mar_wh_pre = y_post_mar_wh_pre y_mar_wh_pre_tx = y_post_mar_wh_pre
        _tx e_ij = e_ij_post)) end = eof;
    end;
    set &libname..&dsname._&nsub.
    (rename = (y_mcar = y_pre e_ij = e_ij_pre));
    by nreps subid;
    if last.subid then y_post = . ;
    if not first.subid then delete;
run;
data &libname..&dsname._&nsub._ancova;
    retain nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post
        _mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
    set test;
    keep nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post_
        mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
run;
%mend;


*Generate data set with 40% missingness;
%macro data40(libname = , dsname = , seed = , nreps = , nsub = , b0 = , b1 = , b2
    = , b3 = , b4 = , b5 = , var_b = , var_e = );
data &dsname._&nsub.;
    retain nreps subid y post waist_hip treatment;
    call streaminit(&seed);
    do nreps = 1 to &nreps;
    do subid = 1 to 2 * &nsub;
    if subid <= &nsub then treatment = 0;
    else treatment = 1;
    waist_hip = rand('normal', 0.9381308532, 0.3485664653);
    b_i = rand('normal', 0, sqrt(&var_b));
    do post = 0 to 1;
    e_ij = rand('normal', 0, sqrt(&var_e));
    y = b_i + &b0 + &b1 * post + &b2 * waist_hip + &b3 * treatment + &b4 * (
        treatment * post) + &b5 * (waist_hip * post) + e_ij;
    output;
    end;
    end;
```

```
    end;
run;

/*Set 40% post scores to missing*/
data &libname..&dsname._&nsub.;
    set &dsname._&nsub.;
    call streaminit(32574435);
    y_mcar = y;
    y_mar_wh = y;
    y_mar_wh_pre = y;
    y_mar_wh_pre_tx = y;
    p_wh = exp(-2.271 + 1.988565 * waist_hip) / (exp(-2.271 + 1.988565 * waist_hip
        ) + 1);
    p_wh_pre = exp(-3.773558 + 1.988565 * waist_hip + 0.0935 * y) / (exp(-3.773558
         + 1.988565 * waist_hip + 0.0935 * y) + 1);
    p_wh_pre_tx = exp(-4.466705 + 1.988565 * waist_hip + 0.0935 * y + 1.386294 *
        treatment) / (exp(-4.466705 + 1.988565 * waist_hip + 0.0935 * y + 1.386294
        * treatment) + 1);
    lag_p_wh_pre = lag(p_wh_pre);
    lag_p_wh_pre_tx = lag(p_wh_pre_tx);
    if post = 1 then do;
    mcar = rand('bernoulli', .4);
    if mcar = 1 then y_mcar = .;
    mar_wh = rand('bernoulli', p_wh);
    if mar_wh = 1 then y_mar_wh = .;
    mar_wh_pre = rand('bernoulli', lag_p_wh_pre);
    if mar_wh_pre = 1 then y_mar_wh_pre = .;
    mar_wh_pre_tx = rand('bernoulli', lag_p_wh_pre_tx);
    if mar_wh_pre_tx = 1 then y_mar_wh_pre_tx = .;
    end;
    rename waist_hip = wh;
run;

/*Construct dataset for MI*/
data test;
    if not eof then do;
    set &libname..&dsname._&nsub. (firstobs = 2 rename = (y = y_post_c y_mcar = y_
        post_mcar y_mar_wh = y_post_mar_wh y_mar_wh_pre = y_post_mar_wh_pre y_mar_
        wh_pre_tx = y_post_mar_wh_pre_tx e_ij = e_ij_post)) end = eof;
    end;
    set &libname..&dsname._&nsub. (rename = (y_mcar = y_pre e_ij = e_ij_pre));
    by nreps subid;
    if last.subid then y_post = . ;
    if not first.subid then delete;
run;
data &libname..&dsname._&nsub._ancova;
    retain nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post
        _mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
    set test;
    keep nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post_
        mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
```

```
run;
%mend;


*Generate data set with 50% missingness;
%macro data50(libname = , dsname = , seed = , nreps = , nsub = , b0 = , b1 = , b2
    = , b3 = , b4 = , b5 = , var_b = , var_e = );
data &dsname._&nsub.;
    retain nreps subid y post waist_hip treatment;
    call streaminit(&seed);
    do nreps = 1 to &nreps;
    do subid = 1 to 2 * &nsub;
    if subid <= &nsub then treatment = 0;
    else treatment = 1;
    waist_hip = rand('normal', 0.9381308532, 0.3485664653);
    b_i = rand('normal', 0, sqrt(&var_b));
    do post = 0 to 1;
    e_ij = rand('normal', 0, sqrt(&var_e));
    y = b_i + &b0 + &b1 * post + &b2 * waist_hip + &b3 * treatment + &b4 * (
        treatment * post) + &b5 * (waist_hip * post) + e_ij;
    output;
    end;
    end;
    end;
run;

/*Set 50% post scores to missing*/
data &libname..&dsname._&nsub.;
    set &dsname._&nsub.;
    call streaminit(32574435);
    y_mcar = y;
    y_mar_wh = y;
    y_mar_wh_pre = y;
    y_mar_wh_pre_tx = y;
    p_wh = exp(-1.865534 + 1.988565 * waist_hip) / (exp(-1.865534 + 1.988565 *
        waist_hip) + 1);
    p_wh_pre = exp(-3.368092 + 1.988565 * waist_hip + 0.0935 * y) / (exp(-3.368092
         + 1.988565 * waist_hip + 0.0935 * y) + 1);
    p_wh_pre_tx = exp(-4.06124 + 1.988565 * waist_hip + 0.0935 * y + 1.386294 *
        treatment) / (exp(-4.06124 + 1.988565 * waist_hip + 0.0935 * y + 1.386294 *
         treatment) + 1);
    lag_p_wh_pre = lag(p_wh_pre);
    lag_p_wh_pre_tx = lag(p_wh_pre_tx);
    if post = 1 then do;
    mcar = rand('bernoulli', .5);
    if mcar = 1 then y_mcar = .;
    mar_wh = rand('bernoulli', p_wh);
    if mar_wh = 1 then y_mar_wh = .;
    mar_wh_pre = rand('bernoulli', lag_p_wh_pre);
    if mar_wh_pre = 1 then y_mar_wh_pre = .;
    mar_wh_pre_tx = rand('bernoulli', lag_p_wh_pre_tx);
```

```
    if mar_wh_pre_tx = 1 then y_mar_wh_pre_tx = .;
    end;
    rename waist_hip = wh;
run;

/*Construct dataset for MI*/
data test;
    if not eof then do;
    set &libname..&dsname._&nsub. (firstobs = 2 rename = (y = y_post_c y_mcar = y_
        post_mcar y_mar_wh = y_post_mar_wh y_mar_wh_pre = y_post_mar_wh_pre y_mar_
        wh_pre_tx = y_post_mar_wh_pre_tx e_ij = e_ij_post)) end = eof;
    end;
    set &libname..&dsname._&nsub.
    (rename = (y_mcar = y_pre e_ij = e_ij_pre));
    by nreps subid;
    if last.subid then y_post = . ;
    if not first.subid then delete;
run;
data &libname..&dsname._&nsub._ancova;
    retain nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post
        _mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
    set test;
    keep nreps subid y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_post_
        mar_wh_pre_tx y_pre treatment wh b_i e_ij_post e_ij_pre;
run;
%mend;


*Mixed model;
/*libname = library name*/
/*dsname = data set name*/
    /*dsname = sim20, sim30, sim40, or sim50*/
/*nsub = number of subject per group*/
    /*nsub = 35 or 100*/
/*mech = missing data mechanism*/
    /*mech = , mcar, mar_wh, mar_wh_pre, or mar_wh_pre_tx*/
/*waist_hip = whether or not include waist hip in the analysis model*/
    /*waist_hip = wh or */
/*cc = whether or not it is for complete case analysis*/
    /*cc = cc or */

%macro mixed(libname = , dsname = , nsub = , mech = , waist_hip = , cc = );
proc sort data = &libname..&dsname._&nsub. out = sim_&nsub._mixed;
    by nreps subid post;
run;
%if &mech. = %then %do; /*Fully observed data*/
proc mixed data = sim_&nsub._mixed noclprint noitprint;
    class subid;
    model y = post &waist_hip. treatment treatment * post / s ddfm = kenwardroger;
    repeated / type = cs subject = subid;
    by nreps;
```

```
    ods output SolutionF = &libname..&dsname._&nsub._&waist_hip.;
run;
%end;
%else %if &cc. ne %then %do; /*Complete Cases*/
data &mech._&cc._&nsub._mixed;
    set sim_&nsub._mixed;
    if y_&mech. = . then getout = subid;
run;
data &mech._&cc._&nsub._mixed2;
    set &mech._&cc._&nsub._mixed;
    set &mech._&cc._&nsub._mixed (firstobs = 2 keep = getout rename = (getout =
        next_getout)) &mech._&cc._&nsub._mixed (obs = 1 drop = _all_);
    next_getout = ifn(last.subid, (.), next_getout );
run;
data &mech._&cc._&nsub._mixed;
    set &mech._&cc._&nsub._mixed2;
    if getout ne . or next_getout ne . then delete;
    drop getout next_getout;
run;
proc mixed data = &mech._&cc._&nsub._mixed noclprint noitprint;
    class subid;
    model y_&mech. = post &waist_hip. treatment treatment * post / s ddfm =
        kenwardroger;
    repeated / type = cs subject = subid;
    by nreps;
    ods output SolutionF = &libname..&dsname._&nsub._&waist_hip.;
run;
%end;
%else %do; /*Available Cases*/
proc mixed data = sim_&nsub._mixed noclprint noitprint;
    class subid;
    model y_&mech. = post &waist_hip. treatment treatment * post / s ddfm =
        kenwardroger;
    repeated / type = cs subject = subid;
    by nreps;
    ods output SolutionF = &libname..&dsname._&nsub._&waist_hip.;
run;
%end;
proc means data = &libname..&dsname._&nsub._&waist_hip. mean;
    var estimate stderr tvalue;
    class effect;
    output out=sumstat;
run;
data &mech._&cc._&nsub._&waist_hip.;
    set sumstat;
    where _stat_ = "MEAN" & effect ne " ";
    drop _type_ _freq_ _stat_;
    condition = "&mech._&cc._&nsub._&waist_hip.";
run;
data &mech._&cc._&nsub._&waist_hip.;
```

```
    retain condition int_est int_se int_t post_est post_se post_t tx_est tx_se tx_
        t ptx_est ptx_se ptx_t wh_est wh_se wh_t;
    set &mech._&cc._&nsub._&waist_hip.;
    by condition;
    if effect = "Intercept" then do;
    int_est = estimate;
    int_se = stderr;
    int_t = tvalue;
    end;
    if effect = "post" then do;
    post_est = estimate;
    post_se = stderr;
    post_t = tvalue;
    end;
    if effect = "post*treatment" then do;
    ptx_est = estimate;
    ptx_se = stderr;
    ptx_t = tvalue;
    end;
    if effect = "treatment" then do;
    tx_est = estimate;
    tx_se = stderr;
    tx_t = tvalue;
    end;
    if effect = "wh" then do;
    wh_est = estimate;
    wh_se = stderr;
    wh_t = tvalue;
    end;
    if last.condition then output;
    keep condition int_est int_se int_t post_est post_se post_t tx_est tx_se tx_t
        ptx_est ptx_se ptx_t wh_est wh_se wh_t;
run;
%mend;


*ANCOVA;
/*libname = library name*/
/*dsname = data set name*/
    /*dsname = sim20, sim30, sim40, or sim50*/
/*nsub = number of subject per group*/
    /*nsub = 35 or 100*/
/*mech = missing data mechanism*/
    /*mech = , mcar, mar_wh, mar_wh_pre, or mar_wh_pre_tx*/
/*waist_hip = whether or not include waist hip in the analysis model*/
    /*waist_hip = wh or */
/*cc = whether or not it is for complete case analysis*/
    /*cc = cc or */

%macro ancova(libname = , dsname = , nsub = , mech = , waist_hip = , cc = );
%if &mech. = %then %do; /*Fully observed data*/
```

```
proc reg data = &libname..&dsname._&nsub._ancova outest = &dsname._&nsub._ancova_&
    waist_hip. tableout noprint;
    model y_post_c = y_pre &waist_hip. treatment;
    by nreps;
run;
quit;
data &libname..&dsname._&nsub._ancova_&waist_hip.;
    set &dsname._&nsub._ancova_&waist_hip.;
    where _type_ in ('PARMS' 'STDERR' 'T');
    keep nreps _type_ intercept y_pre &waist_hip. treatment;
run;
proc means data = &libname..&dsname._&nsub._ancova_&waist_hip. mean;
    var intercept y_pre &waist_hip. treatment;
    class _type_;
    output out=sumstat;
run;
data sumstat;
    retain _type_ intercept treatment wh y_pre;
    set sumstat;
    where _stat_ = "MEAN" & _type_ ne " ";
    drop _freq_ _stat_;
run;
proc transpose data = sumstat out = &mech._&nsub._&waist_hip.;
run;
data &mech._&cc._&nsub._&waist_hip.;
    set &mech._&nsub._&waist_hip.;
    condition = "&mech._&cc._&nsub._&waist_hip.";
    rename col1 = estimate col2 = stderr col3 = tvalue;
    drop _label_;
run;
%end;
%else %if &cc. ne %then %do; /*Complete Cases*/
data &mech._&cc._&nsub._ancova;
    set &libname..&dsname._&nsub._ancova;
    if y_post_&mech. = . then delete;
run;
proc reg data = &mech._&cc._&nsub._ancova outest = &mech._&cc._&nsub._ancova_&
    waist_hip. tableout noprint;
    model y_post_&mech. = y_pre &waist_hip. treatment;
    by nreps;
run;
quit;
data &libname..ancova_&nsub._&mech._&cc._&waist_hip.;
    set &mech._&cc._&nsub._ancova_&waist_hip.;
    where _type_ in ('PARMS' 'STDERR' 'T');
    keep nreps _type_ intercept y_pre &waist_hip. treatment;
run;
proc means data = &libname..ancova_&nsub._&mech._&cc._&waist_hip. mean;
    var intercept y_pre &waist_hip. treatment;
    class _type_;
    output out=sumstat;
```

```
run;
data sumstat;
    retain _type_ intercept treatment wh y_pre;
    set sumstat;
    where _stat_ = "MEAN" & _type_ ne " ";
    drop _freq_ _stat_;
run;
proc transpose data = sumstat out = &mech._&cc._&nsub._&waist_hip.;
run;
data &mech._&cc._&nsub._&waist_hip.;
    set &mech._&cc._&nsub._&waist_hip.;
    condition = "&mech._&cc._&nsub._&waist_hip.";
    rename col1 = estimate col2 = stderr col3 = tvalue;
    drop _label_;
run;
%end;
%else %do; /*Available Cases*/
proc mi data = &libname..&dsname._&nsub._ancova nimpute = 20 seed = 32435345 out =
     &libname..&mech._&nsub._mi;
    class treatment;
    monotone reg (y_post_&mech. = y_pre treatment wh y_pre * treatment y_pre * wh
        treatment * wh / details);
    var y_pre treatment wh y_post_&mech.;
    by nreps;
    ods output VarianceInfo = fmi;
run;
proc reg data = &libname..&mech._&nsub._mi outest = &libname..&mech._&nsub._&waist
    _hip._reg covout noprint;
    model y_post_&mech. = y_pre treatment &waist_hip.;
    by nreps _imputation_;
run;
quit;
proc mianalyze data = &libname..&mech._&nsub._&waist_hip._reg;
    modeleffects intercept y_pre treatment &waist_hip.;
    ods output ParameterEstimates = ancova_&nsub._&mech._&waist_hip.;
    by nreps;
run;
data &libname..ancova_&nsub._&mech._&waist_hip.;
    set ancova_&nsub._&mech._&waist_hip.;
    keep nreps parm estimate stderr tvalue;
run;
proc means data = &libname..ancova_&nsub._&mech._&waist_hip. mean;
    var estimate stderr tvalue;
    class parm;
    output out=sumstat;
run;
data &mech._&cc._&nsub._&waist_hip.;
    set sumstat;
    where _stat_ = "MEAN" & parm ne " ";
    drop _type_ _freq_ _stat_;
    condition = "&mech._&cc._&nsub._&waist_hip.";
```

```
run;
proc means data = fmi;
    var fracmiss;
run;
%end;
data &mech._&cc._&nsub._&waist_hip.;
    retain condition int_est int_se int_t pre_est pre_se pre_t tx_est tx_se tx_t
        wh_est wh_se wh_t;
    set &mech._&cc._&nsub._&waist_hip.;
    by condition;
    if _name_ = "intercept" then do;
    int_est = estimate;
    int_se = stderr;
    int_t = tvalue;
    end;
    if _name_ = "y_pre" then do;
    pre_est = estimate;
    pre_se = stderr;
    pre_t = tvalue;
    end;
    if _name_ = "treatment" then do;
    tx_est = estimate;
    tx_se = stderr;
    tx_t = tvalue;
    end;
    if _name_ = "wh" then do;
    wh_est = estimate;
    wh_se = stderr;
    wh_t = tvalue;
    end;
    if parm = "intercept" then do;
    int_est = estimate;
    int_se = stderr;
    int_t = tvalue;
    end;
    if parm = "y_pre" then do;
    pre_est = estimate;
    pre_se = stderr;
    pre_t = tvalue;
    end;
    if parm = "treatment" then do;
    tx_est = estimate;
    tx_se = stderr;
    tx_t = tvalue;
    end;
    if parm = "wh" then do;
    wh_est = estimate;
    wh_se = stderr;
    wh_t = tvalue;
    end;
    if last.condition then output;
```

```
    keep condition int_est int_se int_t pre_est pre_se pre_t tx_est tx_se tx_t wh_
        est wh_se wh_t;
run;
%mend;


*Analysis Result;
/*method = analysis method*/
    /*method = mixed or ancova*/
/*mis = % of missingness*/
    /*mis = 20, 30, 40, or 50*/

%macro result(method = , mis = );
/*%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = , waist_hip = wh,
    cc = );*/
/*%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = , waist_hip = , cc
     = );*/
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mcar, waist_hip = wh
    , cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mcar, waist_hip = wh
    , cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mcar, waist_hip = ,
    cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mcar, waist_hip = ,
    cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh, waist_hip =
    wh, cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh, waist_hip =
    wh, cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh, waist_hip =
    , cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh, waist_hip =
    , cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre, waist_
    hip = wh, cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre, waist_
    hip = wh, cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre, waist_
    hip = , cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre, waist_
    hip = , cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre_tx, waist
    _hip = wh, cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre_tx, waist
    _hip = wh, cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre_tx, waist
    _hip = , cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 35, mech = mar_wh_pre_tx, waist
    _hip = , cc = );
```

```
/*%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = , waist_hip = wh,
     cc = );*/
/*%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = , waist_hip = ,
    cc = );*/
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mcar, waist_hip =
    wh, cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mcar, waist_hip =
    wh, cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mcar, waist_hip = ,
     cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mcar, waist_hip = ,
     cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh, waist_hip =
     wh, cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh, waist_hip =
     wh, cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh, waist_hip =
     , cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh, waist_hip =
     , cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre, waist_
    hip = wh, cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre, waist_
    hip = wh, cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre, waist_
    hip = , cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre, waist_
    hip = , cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre_tx,
    waist_hip = wh, cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre_tx,
    waist_hip = wh, cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre_tx,
    waist_hip = , cc = cc);
%&method.(libname = sim, dsname = sim&mis., nsub = 100, mech = mar_wh_pre_tx,
    waist_hip = , cc = );

data sim.&method.&mis.;
    set mcar_cc_35_wh mcar__35_wh
        mcar_cc_35_ mcar__35_
        mcar_cc_100_wh mcar__100_wh
        mcar_cc_100_ mcar__100_
        mar_wh_cc_35_wh mar_wh__35_wh
        mar_wh_cc_35_ mar_wh__35_
        mar_wh_cc_100_wh mar_wh__100_wh
        mar_wh_cc_100_ mar_wh__100_
        mar_wh_pre_cc_35_wh mar_wh_pre__35_wh
        mar_wh_pre_cc_35_ mar_wh_pre__35_
        mar_wh_pre_cc_100_wh mar_wh_pre__100_wh
        mar_wh_pre_cc_100_ mar_wh_pre__100_
        mar_wh_pre_tx_cc_35_wh mar_wh_pre_tx__35_wh
```

```
        mar_wh_pre_tx_cc_35_ mar_wh_pre_tx__35_
        mar_wh_pre_tx_cc_100_wh mar_wh_pre_tx__100_wh
        mar_wh_pre_tx_cc_100_ mar_wh_pre_tx__100_;
run;

%if &method. = mixed %then %do;
ods rtf file = "U:\&method.&mis..rtf" bodytitle;
proc print data = sim.&method.&mis.;
    format int_est 7.2 int_se 7.2 int_t 7.2 post_est 7.2 post_se 7.2 post_t 7.2 tx
        _est 7.2 tx_se 7.2 tx_t 7.2 ptx_est 7.2 ptx_se 7.2 ptx_t 7.2 wh_est 7.2 wh_
        se 7.2 wh_t 7.2;
run;
ods rtf close;
%end;

%else %if &method. = ancova %then %do;
ods rtf file = "U:\&method.&mis..rtf" bodytitle;
proc print data = sim.&method.&mis.;
    format int_est 7.2 int_se 7.2 int_t 7.2 pre_est 7.2 pre_se 7.2 pre_t 7.2 tx_
        est 7.2 tx_se 7.2 tx_t 7.2 wh_est 7.2 wh_se 7.2 wh_t 7.2;
run;
ods rtf close;
%end;
%mend;
```

Simulation Results – File "simulation_0125.sas"

```
%include "u:\master thesis\simulation\simulation_macro_0125.sas";
%library(mm = 01, dd = 25);

*b0 - b5, var_b, and var_e come from the following mixed model;
proc mixed data = data.mixed noclprint noitprint;
    class subid;
    model t = post waist_hip treatment treatment * post / s ddfm = kenwardroger;
    repeated / type = cs subject = subid;
    ods output SolutionF = est_para_original CovParms = est_cov_original;
run;

*Coefficient for p_wh;
proc logistic data = data.ancova;
    class dropout (ref = "0");
    model dropout = waist_hip;
    ods select parameterestimates;
run;
quit;

*Coefficient for p_wh_pre;
proc logistic data = data.ancova;
    class dropout (ref = "0");
```

```
   model dropout = waist_hip pre_t;
   ods select parameterestimates;
run;
quit;


*Data generation;
%data20(libname = sim, dsname = sim20, seed = 513471, nreps = 500, nsub = 35,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);
%data20(libname = sim, dsname = sim20, seed = 513471, nreps = 500, nsub = 100,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);

%data30(libname = sim, dsname = sim30, seed = 513471, nreps = 500, nsub = 35,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);
%data30(libname = sim, dsname = sim30, seed = 513471, nreps = 500, nsub = 100,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);

%data40(libname = sim, dsname = sim40, seed = 513471, nreps = 500, nsub = 35,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);
%data40(libname = sim, dsname = sim40, seed = 513471, nreps = 500, nsub = 100,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);

%data50(libname = sim, dsname = sim50, seed = 513471, nreps = 500, nsub = 35,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);
%data50(libname = sim, dsname = sim50, seed = 513471, nreps = 500, nsub = 100,
     b0 = 17.9464, b1 = -0.8111, b2 = -2, b3 = 0, b4 = -1, b5 = 0, var_b = 3.0689,
          var_e = 1.0047);


*Check % missingness;
proc means data = sim.sim20_35;
   var mcar mar_wh mar_wh_pre mar_wh_pre_tx;
run;

proc means data = sim.sim20_100;
   var mcar mar_wh mar_wh_pre mar_wh_pre_tx;
run;

*Check correlation between pre and post scores;
proc corr data = sim.sim20_35_ancova outp = correlation;
   var y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre y_pre;
   by nreps;
run;
data correlation;
```

```
    set correlation;
    where _name_ = "y_pre";
    keep nreps y_post_c y_post_mcar y_post_mar_wh y_post_mar_wh_pre;
run;
proc means data = correlation;
run;


*Analyses;
%result(method = mixed, mis = 20);
%result(method = ancova, mis = 20);

%result(method = mixed, mis = 30);
%result(method = ancova, mis = 30);

%result(method = mixed, mis = 40);
%result(method = ancova, mis = 40);

%result(method = mixed, mis = 50);
%result(method = ancova, mis = 50);
```

Simulation Figures – File "simulation_plots_0322.sas"

```
%include "u:\master thesis\simulation\simulation_macro_0125.sas";
%library(mm = 01, dd = 25);

%macro plotsdata(method = , mis = , nsub = , mech = , waist_hip = , cc = );
%&method.(libname = sim, dsname = sim&mis., nsub = &nsub., mech = &mech., waist_
    hip = &waist_hip., cc = &cc.);
%if &method. = mixed %then %do;
data sim&mis._&nsub._&waist_hip._&cc.;
    set sim.sim&mis._&nsub._&waist_hip.;
    where effect = "post*treatment";
    if probt < 0.05 then &method._&cc. = 1;
    else &method._&cc. = 0;
run;
%end;
%else %if &method. = ancova %then %do;
%if &cc. ne %then %do;
data sim&mis._&nsub._&waist_hip._&cc.;
    set sim.ancova_&nsub._&mech._&cc._&waist_hip.;
    where _type_ = "PVALUE";
    if treatment < 0.05 then &method._&cc. = 1;
    else &method._&cc. = 0;
run;
%end;
%else %do;
data sim&mis._&nsub._&waist_hip._&cc.;
    set sim.ancova_&nsub._&mech._&waist_hip.;
```

```
    where parm = "treatment";
    if probt < 0.05 then &method._&cc. = 1;
    else &method._&cc. = 0;
run;
%end;
%end;
proc means data = sim&mis._&nsub._&waist_hip._&cc.;
    var &method._&cc.;
    output out = sumstat;
run;
data &method.&mis._&nsub._&mech._&waist_hip._&cc.;
    set sumstat;
    mis = &mis.;
    where _stat_ = "MEAN";
    drop _type_;
run;
%mend;


%macro data_method(method = , nsub = , mech = , waist_hip = , cc = );
%plotsdata(method = &method., mis = 20, nsub = &nsub., mech = &mech., waist_hip =
    &waist_hip., cc = &cc.);
%plotsdata(method = &method., mis = 30, nsub = &nsub., mech = &mech., waist_hip =
    &waist_hip., cc = &cc.);
%plotsdata(method = &method., mis = 40, nsub = &nsub., mech = &mech., waist_hip =
    &waist_hip., cc = &cc.);
%plotsdata(method = &method., mis = 50, nsub = &nsub., mech = &mech., waist_hip =
    &waist_hip., cc = &cc.);
data &method._&nsub._&mech._&waist_hip._&cc.;
    set &method.20_&nsub._&mech._&waist_hip._&cc. &method.30_&nsub._&mech._&waist_
        hip._&cc. &method.40_&nsub._&mech._&waist_hip._&cc. &method.50_&nsub._&mech
        ._&waist_hip._&cc.;
run;
%mend;


%macro data_plot1(nsub = , mech = , waist_hip = );
%data_method(method = mixed, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = );
%data_method(method = mixed, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = cc);
%data_method(method = ancova, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = );
%data_method(method = ancova, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = cc);

data plot_&nsub._&mech._&waist_hip.;
    merge mixed_&nsub._&mech._&waist_hip._ mixed_&nsub._&mech._&waist_hip._cc
        ancova_&nsub._&mech._&waist_hip._ ancova_&nsub._&mech._&waist_hip._cc;
run;
```

```
title;

ods html file = "plots for thesis.html" gpath = "u:\";
ods graphics on / imagename = "plot_&nsub._&mech._&waist_hip." noborder;
proc sgplot data = plot_&nsub._&mech._&waist_hip.;
    xaxis label = "Percentage of Missing Data" type = discrete;
    yaxis label = "Power" min = .35 max = .8 minor;
    series x = mis y = mixed_ / markers lineattrs = (color = black pattern = 1)
        markerattrs = (color = black symbol = circle) legendlabel = "Mixed Model,
        AA" name = "Mixed Model, AA";
    series x = mis y = mixed_cc / markers lineattrs = (color = black pattern = 41)
         markerattrs = (color = black symbol = diamond) legendlabel = "Mixed Model,
         CC" name = "Mixed Model, CC";
    series x = mis y = ancova_ / markers lineattrs = (color = black pattern = 5)
        markerattrs = (color = black symbol = circlefilled)legendlabel = "ANCOVA,
        MI" name = "ANCOVA, MI";
    series x = mis y = ancova_cc / markers lineattrs = (color = black pattern =
        34) markerattrs = (color = black symbol = diamondfilled)legendlabel = "
        ANCOVA, CC" name = "ANCOVA, CC";
    keylegend "Mixed Model, AA" "Mixed Model, CC" "ANCOVA, MI" "ANCOVA, CC" /
        location = inside position = bottomleft across = 2;
    /*title "Power Comparison for &mech. (n = &nsub./Group &waist_hip.)";*/
run;
ods graphics off;
ods html close;
%mend;


%macro data_plot2(nsub = , mech = , waist_hip = );
%data_method(method = mixed, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = );
%data_method(method = mixed, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = cc);
%data_method(method = ancova, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = );
%data_method(method = ancova, nsub = &nsub., mech = &mech., waist_hip = &waist_hip
    ., cc = cc);

data plot_&nsub._&mech._&waist_hip.;
    merge mixed_&nsub._&mech._&waist_hip._
    mixed_&nsub._&mech._&waist_hip._cc
    ancova_&nsub._&mech._&waist_hip._
    ancova_&nsub._&mech._&waist_hip._cc;
run;

title;

ods html file = "plots for thesis.html" gpath = "u:\";
ods graphics on / imagename = "plot_&nsub._&mech._&waist_hip." noborder;
proc sgplot data = plot_&nsub._&mech._&waist_hip.;
    xaxis label = "Percentage of Missing Data" type = discrete;
```

```
    yaxis label = "Power" min = .85 max = 1 minor;
    series x = mis y = mixed_ / markers lineattrs = (color = black pattern = 1)
        markerattrs = (color = black symbol = circle) legendlabel = "Mixed Model,
        AA" name = "Mixed Model, AA";
    series x = mis y = mixed_cc / markers lineattrs = (color = black pattern = 41)
         markerattrs = (color = black symbol = diamond) legendlabel = "Mixed Model,
         CC" name = "Mixed Model, CC";
    series x = mis y = ancova_ / markers lineattrs = (color = black pattern = 5)
        markerattrs = (color = black symbol = circlefilled)legendlabel = "ANCOVA,
        MI" name = "ANCOVA, MI";
    series x = mis y = ancova_cc / markers lineattrs = (color = black pattern =
        34) markerattrs = (color = black symbol = diamondfilled)legendlabel = "
        ANCOVA, CC" name = "ANCOVA, CC";
    keylegend "Mixed Model, AA" "Mixed Model, CC" "ANCOVA, MI" "ANCOVA, CC" /
        location = inside position = bottomleft across = 2;
run;
ods graphics off;
ods html close;
%mend;

%data_plot1(nsub = 35, mech = mcar, waist_hip = );
%data_plot1(nsub = 35, mech = mcar, waist_hip = wh);
%data_plot2(nsub = 100, mech = mcar, waist_hip = );
%data_plot2(nsub = 100, mech = mcar, waist_hip = wh);

%data_plot1(nsub = 35, mech = mar_wh, waist_hip = );
%data_plot1(nsub = 35, mech = mar_wh, waist_hip = wh);
%data_plot2(nsub = 100, mech = mar_wh, waist_hip = );
%data_plot2(nsub = 100, mech = mar_wh, waist_hip = wh);

%data_plot1(nsub = 35, mech = mar_wh_pre, waist_hip = );
%data_plot1(nsub = 35, mech = mar_wh_pre, waist_hip = wh);
%data_plot2(nsub = 100, mech = mar_wh_pre, waist_hip = );
%data_plot2(nsub = 100, mech = mar_wh_pre, waist_hip = wh);

%data_plot1(nsub = 35, mech = mar_wh_pre_tx, waist_hip = );
%data_plot1(nsub = 35, mech = mar_wh_pre_tx, waist_hip = wh);
%data_plot2(nsub = 100, mech = mar_wh_pre_tx, waist_hip = );
%data_plot2(nsub = 100, mech = mar_wh_pre_tx, waist_hip = wh);
```

Bibliography

[1] King AC, Oman RF, Brassington GS, Bliwise DL, Haskell WL. Moderate-intensity exercise and self-rated quality of sleep in older adults: a randomized controlled trial. Jama. 1997;277(1):32–37.

[2] Allison PD. Change scores as dependent variables in regression analysis. Sociological methodology. 1990;20:93–114.

[3] Cronbach LJ, Furby L. How we should measure "change": Or should we? Psychological bulletin. 1970;74(1):68.

[4] Glidden DV, Shiboski SC, McCulloch CE. Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. Springer; 2011.

[5] Oakes JM, Feldman HA. Statistical Power for Nonequivalent Pretest-Posttest Designs The Impact of Change-Score versus ANCOVA Models. Evaluation Review. 2001;25(1):3–28.

[6] Little RJ, Rubin DB. Statistical analysis with missing data. Wiley; 2002.

[7] Lachin JM. Statistical considerations in the intent-to-treat principle. Controlled clinical trials. 2000;21(3):167–189.

[8] White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Statistics in medicine. 2010;29(28):2920–2931.

[9] Salim A, Mackinnon A, Christensen H, Griffiths K. Comparison of data analysis strategies for intent-to-treat analysis in pre-test–post-test designs with substantial dropout rates. Psychiatry research. 2008;160(3):335–345.

[10] Campbell DT, Stanley JC, Gage NL. Experimental and quasi-experimental designs for research. Houghton Mifflin Boston; 1963.

[11] Pape UJ, Millett C, Lee JT, Car J, Majeed A. Disentangling secular trends and policy impacts in health studies: use of interrupted time series analysis. Journal of the Royal Society of Medicine. 2013;106(4):124–129.

[12] Harris CW, et al. Problems in measuring change. University of Wisconsin Press Madison; 1963.

[13] Werts CE, Linn RL. A general linear model for studying growth. Psychological Bulletin. 1970;73(1):17–22.

[14] Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ: British Medical Journal. 2009;339(7713):157–160.

[15] Panel on Handling Missing Data in Clinical Trials; National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials. The National Academies Press; 2010. Available from: `http://www.nap.edu/openbook.php?record_id=12955`.

[16] Ibrahim JG, Chu H, Chen MH. Missing data in clinical studies: issues and methods. Journal of Clinical Oncology. 2012;30(26):3297–3303.

[17] Schafer JL. Analysis of incomplete multivariate data. CRC press; 1997.

[18] Allison PD. Multiple imputation for missing data: A cautionary tale. Philadelphia: University of Pennsylvania. 1999;.

[19] Casella G, George EI. Explaining the Gibbs sampler. The American Statistician. 1992;46(3):167–174.

[20] Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation. Journal of the American statistical Association. 1987;82(398):528–540.

[21] Yuan Y. Multiple imputation using SAS software. Journal of Statistical Software. 2011;45(6):1–25.

[22] David P, Buckworth J, Pennell ML, Katz ML, DeGraffinreid CR, Paskett ED. A walking intervention for postmenopausal women using mobile phones and interactive voice response. Journal of telemedicine and telecare. 2012;18(1):20–25.

[23] Gomez EV, Schaalje GB, Fellingham GW. Performance of the Kenward–Roger method when the covariance structure is selected using AIC and BIC. Communications in Statistics — Simulation and Computation®. 2005;34(2):377–392.

[24] Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics. 1997;53(3):983–997.

[25] Littell RC. SAS for mixed models. SAS institute; 2006.