A Comparison for Longitudinal Data Missing Due to Truncation

Rong Liu
Virginia Commonwealth University

Follow this and additional works at: https://scholarscompass.vcu.edu/etd
Part of the Biostatistics Commons

© The Author

Downloaded from
https://scholarscompass.vcu.edu/etd/1096
A COMPARISON FOR LONGITUDINAL DATA MISSING DUE TO TRUNCATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

Rong Liu
M.D. Shandong Medical University, P.R. China, 1997
M.P.H. Beijing Microbiology and Epidemiology Institute, P.R. China, 2000
M.S. University of Virginia, Virginia, 2004

Co-Directors:

Dr. Viswanathan Ramakrishnan
Associate Professor
Department of Biostatistics

Dr. Kellie J. Archer
Assistant Professor
Department of Biostatistics

Virginia Commonwealth University
Richmond, Virginia
February, 2006
Acknowledgement

First, I would like to begin by acknowledging and thanking the members of my dissertation committee: my advisor, Dr. Viswanathan Ramakrishnan, for his invaluable guidance and support throughout this research, my knowledge of biostatistics has grown greatly because of him, Dr. Kellie Archer, for her statistical insight on bootstrap sampling, Dr. Charlie Kish, for his constant encouragement, Dr. R.K. Elswick, for keeping me on the track to graduate, Dr Saba Masho, for being flexible about the meeting schedule and thoughtful insight on this topic.

Secondly, I would like to thank the numerous faculties, staff and students in the department of biostatistics at VCU. I have enjoyed working with this group of fine people. I would also thank the department and Wyeth Consumer Healthcare, in particular, Charlie Kish and Carol Summitt, for providing me with the one-year internship opportunity. It has been a most valuable and positive learning experience for me.

Finally, I would like to thank my family for helping to make me the person I am today, especially my father who taught me to love learning, my mother who convinced me to reach for the stars and my brother, who has been so good to take care of our parents while I am on the other side of the ocean and the only thing I can offer is talking. The last ‘thank-you’ goes to my best friend and dearest husband Jackie, thank you for making me your wife and thank you for being always present, always supportive.
Table of Contents

List of Tables ....................................................................................................................... viii

List of Figures ....................................................................................................................... xi

Chapter

Chapter 1 Introduction and Prospectus ............................................................................... 1

1.1 Introduction ..................................................................................................................... 1

1.2 Types of Missing Data .................................................................................................. 3

1.3 Ignorable and Nonignorable Missing Data ................................................................. 3

1.4 Other Types of Missing Data ....................................................................................... 5

1.5 Analyzing Missing Data .............................................................................................. 6

1.6 Examples of Nonignorable Missing Data ..................................................................... 7

1.6.1 A panel study on methadone treatment practices ................................................. 7

1.6.2 An anti-psychotic drug study .................................................................................. 8

1.7 Prospectus ..................................................................................................................... 10

Chapter 2 Missing Due to Truncation Method .................................................................... 12

2.1 Introduction .................................................................................................................. 12

2.2 The MDT Data Structure ............................................................................................ 13

2.3 Likelihood and Estimation .......................................................................................... 14

2.4 Missing Data at Several Time Points .......................................................................... 23
2.5 Data Missing Due to Truncation at both Tails of the Distribution ........23
2.6 MDT with MAR and MCAR .................................................................24
2.7 Other Methods Useful for Longitudinal Missing Data ..................25
   2.7.1 Last observation carried forward..............................................25
   2.7.2 Individual regression prediction method.................................26
   2.7.3 Repeated measures mixed model .............................................26
   2.7.4 Multiple imputation.................................................................27
   2.7.5 Treating missing as censored ..................................................28
Chapter 3 Analysis of IMPS Data by MDT Method ................................30
   3.1 Introduction ................................................................................30
   3.2 Application of the MDT Method ..................................................33
   3.3 Analysis Results by MDT Method ...............................................40
   3.4 Discussion ..................................................................................44
Chapter 4 Simulation Study for MDT Method .......................................46
   4.1 Introduction ................................................................................46
   4.2 The MDT Method .......................................................................47
   4.3 Likelihood and Estimation ...........................................................48
   4.4 Other Methods Useful for Longitudinal Missing Data .................53
      4.4.1 Last observation carried forward method (LOCF)...............54
      4.4.2 Individual regression prediction method (PRED) ...............54
4.4.3 Repeated measures mixed model method (MIXED) ..................55
4.5 Design of the Simulation Study .............................................55
4.5.1 Factors in the simulation study ..........................................56
4.5.2 Generation of multivariate normal random sample with missing 58
4.5.3 Comparison of measures ................................................59
4.6 Simulation Results .............................................................61
4.7 Discussion ........................................................................63

Chapter 5 MDT Method in conjunction with Multiple Imputation ..........65
5.1 Introduction .......................................................................65
5.2 Review of Methods for Continuous Repeated Measures with
    Nonignorable Dropout .........................................................66
      5.2.1 Systematic Difference from Ignorable Imputation ..........67
      5.2.2 Semiparametric Nonresponse Model ............................68
      5.2.3 Likelihood-Based Methods .........................................68
5.3 MDT Method in conjunction with Multiple Imputation (MI)..........73
5.4 Analysis of IMPS Data using MDT Method in conjunction with MI,79
5.5 Comparison with MI Procedure treating all Missing as MAR ..........86

Chapter 6 Summary and Extension .............................................88
6.1 Dissertation Summary .......................................................88
6.2 Extension and Suggestions for Future Research .......................90
6.2.1 Extend MDT method to Multivariate Outcomes

6.2.2 Allow for the Threshold to be Random

References

Appendices

Appendix A: IMPS Data Listing and Histogram Plots: Chapter 3

Appendix B: Simulation Results: Chapter 4

Appendix C: Programs for Application of MDT method to IMPS data

C.1 Program for iterating between E-step and M-step

C.2 Program for E-Step of EM algorithm to estimate the MDT

C.2 Program for M-Step of EM algorithm

Appendix D: Programs for Simulation Study of MDT Method and Other Methods

D.1 Program for missing due to truncation method

D.2 Program for individual regression prediction method

D.3 Program for repeated measure mixed model method

D.4 Program for last observation carried forward

Appendix E: Programs for MDT Method in conjunction with Multiple Imputation including Bootstrap Algorithm

E.1 Program for bootstrap sampling of IMPS data and application of MDT to bootstrap samples
E.2 Program for combining the MDT method with multiple imputation for IMPS data set.
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>IMPS data in an antipsychotic drug study: sample summary</td>
<td>9</td>
</tr>
<tr>
<td>1.2</td>
<td>IMPS data in an antipsychotic drug study for ages 6-21 group: sample summary</td>
<td>10</td>
</tr>
<tr>
<td>3.1</td>
<td>Subjects with missing data in placebo group</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Subjects with missing data in treatment group</td>
<td>35</td>
</tr>
<tr>
<td>3.3</td>
<td>Initial values with AR(1) covariance structure</td>
<td>37</td>
</tr>
<tr>
<td>3.4</td>
<td>Covariance structure selection with time as categorical variable</td>
<td>41</td>
</tr>
<tr>
<td>3.5</td>
<td>Covariance structure selection with SQRT(Time) as covariate</td>
<td>42</td>
</tr>
<tr>
<td>5.1</td>
<td>Variance information using MDT in conjunction with MI</td>
<td>85</td>
</tr>
<tr>
<td>5.2</td>
<td>Parameter estimates using MDT in conjunction with MI</td>
<td>86</td>
</tr>
<tr>
<td>5.3</td>
<td>Multiple imputation variance Information using MI assuming MAR</td>
<td>87</td>
</tr>
<tr>
<td>5.4</td>
<td>Multiple imputation parameter Estimates using MI assuming MAR</td>
<td>87</td>
</tr>
<tr>
<td>A.1</td>
<td>IMPS data listing</td>
<td>102</td>
</tr>
<tr>
<td>B.1</td>
<td>Mean ($\mu_4 = 2.6$) estimates (s.d) from different methods for linear response and MDT at last time point</td>
<td>108</td>
</tr>
<tr>
<td>B.2</td>
<td>Variance ($\sigma^2 = 2$) estimates (s.d) from different methods for linear response and MDT at last time point</td>
<td>109</td>
</tr>
</tbody>
</table>
Table B.3: Correlation estimates (s.d) from different methods for linear response and

**MDT at last time point** .................................................................110

Table B.4: Mean (μ₃) estimates (s.d) from different methods for linear response and MDT

at last two time points. .................................................................................................111

Table B.5: Mean (μ₄=2.6) estimates (s.d) from different methods for linear response and

MDT at last two time points. .................................................................................................112

Table B.6: Variance (σ²=2) estimates (s.d) from different methods for linear response

and MDT at last two time points.......................................................................................113

Table B.7: Correlation estimates (s.d) from different methods for linear response and

MDT at last two time points. .................................................................................................114

Table B.8: Mean (μ₄=2.6) estimates (s.d) from different forms of response function and

20% MDT at last time point.................................................................115

Table B.9: Variance (σ²=2) estimates (s.d) from different forms of response function and

20% MDT at last time point.......................................................................................116

Table B.10: Correlation estimates (s.d) from different forms of response function and

20% MDT at last time point.......................................................................................117

Table B.11: Mean (μ₄=2.6) estimates (s.d) from linear response function and MAR at last
time point. .................................................................................................118
Table B.12: Variance ($\sigma^2=2$) estimates (s.d) from linear response function and MAR at last time point ................................................................. 119

Table B.13: Correlation estimates (s.d) from linear response function and MAR at last time point ........................................................................................................ 120
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>MDT data at time $T$</td>
<td>21</td>
</tr>
<tr>
<td>3.1</td>
<td>Placebo group for ages 16-21</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>Treatment group for ages 16-21</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>Flowchart of MDT method</td>
<td>36</td>
</tr>
<tr>
<td>3.4</td>
<td>LSmeans comparisons from repeated measures mixed model</td>
<td>43</td>
</tr>
<tr>
<td>4.1</td>
<td>The Three response functions simulated</td>
<td>57</td>
</tr>
<tr>
<td>5.1</td>
<td>Flowchart of MDT method in conjunction with MI</td>
<td>84</td>
</tr>
<tr>
<td>A.1</td>
<td>Histogram of IMPS scores for placebo group at 3 weeks and 6 weeks</td>
<td>106</td>
</tr>
<tr>
<td>A.2</td>
<td>Histogram of IMPS Scores for treatment group at 3 weeks and 6 weeks</td>
<td>107</td>
</tr>
<tr>
<td>B.1</td>
<td>Mean ($\mu_4$) estimates from different methods for linear response, MDT at last time point and AR(1)=0.2, 0.4 and 0.8</td>
<td>121</td>
</tr>
<tr>
<td>B.2</td>
<td>Variance estimates from different methods for linear response, MDT at last time point and AR(1)=0.2, 0.4 and 0.8</td>
<td>122</td>
</tr>
<tr>
<td>B.3</td>
<td>Correlation estimates from different methods for linear response, MDT at last time point and AR(1)=0.2, 0.4 and 0.8</td>
<td>123</td>
</tr>
<tr>
<td>B.4</td>
<td>Mean ($\mu_1$) estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8</td>
<td>124</td>
</tr>
</tbody>
</table>
Figure B.5: Mean ($\mu_4$) estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8.................................125

Figure B.6: Variance estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8.................................126

Figure B.7: Correlation estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8.................................127

Figure B.8: Variance estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8. Enlarge the left corners of plots in Figure B.6.................................................................128

Figure B.9: Correlation estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8. Enlarge the left corners of plots in Figure B.7.................................................................129

Figure B.10: Mean ($\mu_4$) estimates from different forms of response function, 20% MDT at last time point and AR(1)=0.2, 0.4 and 0.8.................................130

Figure B.11: Variance estimates from different forms of response function, 20% MDT at last time point and AR(1)=0.2, 0.4 and 0.8.................................131

Figure B.12: Correlation estimates from different forms of response function, 20% MDT at last time point and AR(1)=0.2, 0.4 and 0.8.................................132

Figure B.13: Mean ($\mu_4$) estimates from different methods, MAR at last time point and AR(1)=0.2, 0.4 and 0.8 .................................................................133
Figure B.14: Variance estimates from different methods, MAR at last time point and

AR(1)=0.2, 0.4 and 0.8 .................................................................134

Figure B.15: Correlation estimates from different methods, MAR at last time point and

AR(1)=0.2, 0.4 and 0.8. .................................................................135
Abstract

A COMPARISON FOR LONGITUDINAL DATA MISSING DUE TO TRUNCATION

By Rong Liu, Ph.D.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2006

Co-Directors:

Dr. Viswanathan Ramakrishnan
Associate Professor
Department of Biostatistics

Dr. Kellie J. Archer
Assistant Professor
Department of Biostatistics

Many longitudinal clinical studies suffer from patient dropout. Often the dropout is nonignorable and the missing mechanism needs to be incorporated in the analysis. The methods handling missing data make various assumptions about the missing mechanism, and their utility in practice depends on whether these assumptions apply in a specific application. Ramakrishnan and Wang (2005) proposed a method (MDT) to handle
nonignorable missing data, where missing is due to the observations exceeding an
unobserved threshold. Assuming that the observations arise from a truncated normal
distribution, they suggested an EM algorithm to simplify the estimation.

In this dissertation the EM algorithm is implemented for the MDT method when
data may include missing at random (MAR) cases. A data set, where the missing data
occur due to clinical deterioration and/or improvement is considered for illustration. The
missing data are observed at both ends of the truncated normal distribution. A simulation
study is conducted to compare the performance of other relevant methods. The factors
chosen for the simulation study included, the missing data mechanisms, the forms of
response functions, missing at one or two time points, dropout rates, sample sizes and
different correlations with AR(1) structure. It was found that the choice of the method for
dealing with the missing data is important, especially when a large proportion is missing.
The MDT method seems to perform the best when there is reason to believe that the
assumption of truncated normal distribution is appropriate.

A multiple imputation (MI) procedure under the MDT method to accommodate the
uncertainty introduced by imputation is also proposed. The proposed method combines the
MDT method with Rubin's (1987) MI method. A procedure to implement the MI method
is described.
Chapter 1
Introduction and Prospectus

1.1 Introduction

Missing data is a common problem in longitudinal clinical trials. It is caused by patients who do not complete the study schedule and drop out from the study for known or unknown reasons. Possible reasons for patients dropping out of the study include death, adverse reactions, unpleasant study procedures, lack of improvement, early recovery, and other factors related or unrelated to trial procedure or treatments. The exact reasons may or may not be available at the time of data collection. In either case, the loss of information from missing data could introduce bias or reduce power for detecting treatment effect.

There are numerous approaches in the literature which are useful in handling missing data. These statistical approaches depend on certain assumptions regarding the mechanism by which the missing data arise. Thus the primary step necessary to appropriately handle missing data is to clearly characterize the missing data mechanism.

Consider an $N \times p$ data matrix $Y$ with $N$ subjects observed on $p$ variables from a $p$ dimensional multivariate probability distribution $p(Y | \theta)$, where $\theta$ may be a scalar or
vector-valued parameter. Suppose there is an $N \times p$ matrix $I$, where the $(i,j)^{th}$ element, $I_{i,j}$ is an indicator of whether or not an observation is missing. That is,

$$I_{i,j} = \begin{cases} 1 & \text{if } y_{i,j} \text{ is missing} \\ 0 & \text{if } y_{i,j} \text{ is observed} \end{cases}$$

The matrix $I$ has a probability distribution $p(I | \xi, Y)$, conditional on the response $Y$, where $\xi$ is an unknown scalar or vector-valued parameter of the missing data mechanism. When $Y$ is not fully observed, denote the observed part of $Y$ by $Y_{\text{obs}}$ and the missing part by $Y_{\text{mis}}$. The joint probability distribution of the response variables and the missing indicator can be expressed as the product of the marginal distribution of the response variable and the conditional distribution of missing indicator given the response variables. That is,

$$p(Y, I | \theta, \xi) = p(Y | \theta) p(I | \xi, Y).$$

(1.1)

(The notation $p(Y | \theta)$ is used in place of the conventional notation $f(Y; \theta)$ in order to include the Bayesian approach, if necessary.) There are two sets of parameters, the parameter of interest $\theta$ and the nuisance parameter $\xi$. In general inferences on $\theta$ should be based on the joint probability of $Y$ and $I$ as in (1.1). That is, the inference not only should depend on the distribution of $Y$ but also should depend on how the probability model for missing data is defined.

A distinction is made between three types of missing data through the conditional distribution $p(I | \xi, Y)$ (Rubin, 1976; Little & Rubin, 2002).
1.2 Types of missing data

a) Missing at random

When the missing data mechanism depends on the observed data but not on the unobserved data, that is $p(I|Y, \xi) = p(I|Y_{\text{obs}}, \xi)$, the data are missing at random (MAR). The MAR therefore can be predicted from just the available responses.

b) Missing completely at random

When the missing data mechanism is independent of both the unobserved and the observed data the data are missing completely at random (MCAR). In this case $p(I|Y, \xi) = \xi$. Missing values for a variable under MCAR can therefore be predicted by a random sample of the observed data for that variable. Notice that the MCAR is a special case of MAR.

c) Missing not at random

When the missing data mechanism depends on both the observed and the missing responses the data are missing not at random (MNAR). In this case $p(I|Y, \xi) \neq p(I|Y_{\text{obs}}, \xi)$. Therefore missing data cannot be imputed with the observed data alone. Further knowledge of the missing data mechanism or assumptions regarding the missing mechanism is required for imputing the missing data.

1.3 Ignorable and nonignorable missing data

Suppose the model parameter $\theta$ and the missing data parameter $\xi$ are from the parameter space $\Theta_1$ and $\Theta_2$, respectively. The model parameter $\theta$ and the missing data parameter $\xi$ are said to be distinct if from a frequentist perspective, the joint parameter
space of \((\theta, \xi)\), say \(\Theta\), is the product of \(\Theta_1\) and \(\Theta_2\), and from a Bayesian perspective, the joint prior distribution of parameters \(\theta\) and \(\xi\) is the product of priors of \(\theta\) and \(\xi\). If both MAR and distinctness hold, the missing data mechanism is defined to be ignorable (Rubin, 1987, pp. 51; Little and Rubin, 2002, pp. 119). If a missing data mechanism does not satisfy the ignorability definition, the missing data mechanism is nonignorable.

Since \(Y_{mis}\) is unknown, the full likelihood function of this distribution can not be evaluated. Therefore the inference is based on the observed data likelihood function. By definition, the observed data likelihood function is proportional to the joint distribution in (1.1) integrated over \(Y_{mis}\). That is,

\[
L(\theta, \xi | Y, I) \propto p(Y_{obs}, I | \theta, \xi),
\]

where

\[
p(Y_{obs}, I | \theta, \xi) = \int p(Y, I | \theta, \xi) dY_{mis}. \tag{1.3}
\]

In the case of ignorable missing, under the MAR assumption, (1.3) yields

\[
p(Y_{obs}, I | \theta, \xi) = \int p((Y_{obs}, Y_{mis}), I | \theta, \xi) dY_{mis}
= \int p(I | Y_{obs}, \xi) p((Y_{obs}, Y_{mis}) | \theta) dY_{mis}
= p(I | Y_{obs}, \xi) \int p((Y_{obs}, Y_{mis}) | \theta) dY_{mis}
= p(I | Y_{obs}, \xi) p(Y_{obs} | \theta).
\]

Under the assumption that parameter \(\theta\) and parameter \(\xi\) are distinct, likelihood based inferences about \(\theta\) will be unaffected by \(\xi\) or \(p(I | Y_{obs}, \xi)\). That is, the joint observed data distribution \(p(Y_{obs}, I | \theta, \xi)\) can be replaced by the marginal observed data distribution \(p(Y_{obs} | \theta)\) for the purposes of inferences on \(\theta\).
For nonignorable data, inferences on \( \theta \) can not be based on the marginal observed data distribution \( p(Y_{\text{obs}} \mid \theta) \) alone as in the ignorable missing case. If a specific model for the missing data mechanism is known, the full likelihood \( L(\theta, \xi \mid Y_{\text{obs}}, I) \) needs to be defined and inferences can be based on this.

### 1.4 Other types of missing data

The missing pattern can also be categorized into monotone missing and non-monotone missing. This distinction is useful in longitudinal data. The missing pattern is said to be monotone if, whenever an element \( y_{ij} \) is missing, \( y_{ik} \) is also missing for all points of time \( k > j \). Otherwise the missing pattern is called non-monotone (Little & Rubin, 2002, pp6). Monotone missing often arise in clinical trials with repeated measures. For example, a subject may drop out of the trial prior to the end of the trial and does not return, so that all the measures at the subsequent points of time are also missing. Let \( n_j \) denote the number of observed values at time \( j \), then if the missing data follow a monotone pattern then the condition \( n_1 \geq n_2 \geq \ldots \geq n_p \) must be true. The joint observed-data likelihood for \( \theta \) in this case can be factored into the independent observed data likelihood for \( \theta_1, \theta_2, \ldots, \theta_p \) as follows.

\[
L(\theta \mid Y) = \prod_{i=1}^{n} p(y_{i1}, \ldots, y_{ip} \mid \theta)
\]

\[
= \prod_{i=1}^{n} \prod_{j=1}^{p} p(y_{ij} \mid y_{i1}, \ldots, y_{i,j-1}, \theta_j)
\]

\[(1.4)\]

\[
= \prod_{i=1}^{n} \prod_{j=1}^{p} p(y_{ij} \mid y_{i1}, \ldots, y_{i,j-1}, \theta_j).
\]
It reduces the problem of inference about $\theta$ to a sequence of independent univariate distributions given the previous observed data (Schafer, 1997, pp219).

In practice, the pattern of missing data is rarely monotone but it is often close to monotone. For non-monotone missing, when the missing proportion is not large, the observations that violate monotone pattern may be discarded to create monotone pattern. Or we can impute enough missing values and create a monotone pattern so that the methods for monotone missing data can be applied.

1.5 Analyzing missing data

Most of the missing data methods impute the missing data and then the analysis is performed using complete data inferential methods. Some methods are based on a single imputation and other methods are based on multiple imputations. In general, the basic idea of imputation is to fill in the missing data by using values based on a certain model along with assumptions on $Y$ and $I$ in (1.1). The advantage of imputation methods is that once the missing data are filled-in (imputed) all the statistical tools available for complete data could be easily applied. As mentioned before, appropriate imputation methods depend on the missing data mechanism as well as the missing pattern. Most available imputation methods deal with monotone missing pattern (Hao & Krisnamoorthy 2001, Wu & Perlman, 2000, Molenberghs & Michiels, 1998). When nonmonotone missing occurs some programs (such as SAS MI) use simulation methods such as the Markov Chain Monte Carlo (MCMC) to impute either all the missing values or just enough missing values to make the imputed data sets have only monotone missing so that other more flexible imputation methods can be applied.
Most methods available in commercial computer programs are applicable only when the dropouts could be treated as missing at random. Some widely used methods include complete-case analysis, last observation carried forward, and regression prediction and so on. These methods will be further described in chapter 2.

Statistical packages (such as SPSS, SAS) that can be used for longitudinal data with missing data are now widely available. These analyses tools are valuable in that they incorporate all the available information in the data. It can reduce or even eliminate the bias resulting from an analysis confined to the complete cases. However, as mentioned the estimates from these models assume that the missing data are MAR, which sometimes may not be plausible. Some examples where this may be the case, are given here.

1.6 Examples of nonignorable missing data

1.6.1 A panel study on methadone treatment practices

Methadone treatment is an important vehicle to reduce drug use and prevent human immunodeficiency virus (HIV) transmission. A panel study of a random sample 172 methadone treatment units nationwide was conducted to investigate how the methadone treatment has improved from 1988 to 1995. The 172 methadone units were phone surveyed about personnel, clients, and methadone treatment practices in 1988. An additional two waves of data collection took place in 1990 and 1995, with only 140 (81%) units responding in 1990 and 116 (67%) units responding in 1995. The effectiveness of the methadone treatment is measured by adequate dose level (typically 60 - 120 mg/day), sufficient treatment duration, and a small percentage of clients receiving progressive smaller doses. These three outcome variables are measured repeatedly over time.
Analyses of this dataset pose several challenges. First, the data consist of multivariate longitudinal outcomes whose joint effects capture the effectiveness of critical treatment practices. Second, a substantial percentage of units (33%) didn’t respond during the follow-up. These drop-out units tend to be the units with less effective treatment practices, thus making the dropout mechanism nonignorable. Third, several of the covariates vary over time. For the unit that dropped out from the study, these time-varying covariates were missing at the time of dropout. Therefore, analysis of these methadone data requires addressing the three issues simultaneously. Roy and Lin (2000 & 2002) developed a statistical model for multivariate longitudinal outcomes, while accommodating nonignorable dropouts and dropout-related missing time-varying covariates. Their work will be summarized in chapter 6.

1.6.2 An anti-psychotic drug study

A collaborative study conducted by the National Institute of Mental Health. Three anti-psychotic drugs, chlorpromazine, fluphenazine and thioridazine were compared to a placebo in a sample of inpatients between ages 16-45 on the overall severity of illness. Sample sizes for the drug groups were 110, 112, and 107, respectively. Here they will be combined into one treatment group since previous study showed that similar effects for these three anti-psychotic drugs are expected (Gibbons & Hedeker, 1988). The sample size for the placebo group is 107. Symptom severity was measured by the Inpatient Multidimensional Psychiatric Scale (IMPS) (Lorr & Klett, 1966). The patients were followed-up at four time points (baseline, 1 week, 3 weeks, and 6 weeks) and the IMPS were collected at each time point. During the course of the study, 33% (35/107) of the
subjects dropped out in the placebo group and 17% (56/329) of the subjects dropped out in
the treatment group, six of the 56 (11%) “treatment” patients dropped out due to
deterioration compared to 83% (29/35) in the placebo group. Twenty three percent (13/56)
of the patients in the treatment group dropped out because of improvement compared to
0% (0/35) in the placebo group (Table 1.1). The missing data in the “worsening” category
were recorded as “treatment failure” in the hospital records (they would have had high
IMPS measurements if they had stayed in the study). The data in the “improved” category
resulted from “hospital discharge” in the hospital records; they would have had low IMPS
measurements if observed. The rest of the missing data could be classified as missing at
random.

<table>
<thead>
<tr>
<th>Completion Status</th>
<th>Treatment (N=329)</th>
<th>Placebo (N=107)</th>
<th>Total (N=436)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Complete</td>
<td>273</td>
<td>83</td>
<td>72</td>
</tr>
<tr>
<td>Incomplete</td>
<td>56</td>
<td>17</td>
<td>35</td>
</tr>
<tr>
<td>Worsening</td>
<td>6</td>
<td>11</td>
<td>29</td>
</tr>
<tr>
<td>Improved</td>
<td>13</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>MAR</td>
<td>37</td>
<td>66</td>
<td>6</td>
</tr>
</tbody>
</table>

In the age range 16-21 (Table 1.2), number of subjects that completed the study
was 61 in the treatment group, and 18 in the placebo group. Eighty three percent (5/6) of
the patients dropped out due to “treatment failure” in the placebo group. It is much higher
than the dropout in the treatment group (20% (2/10)). Ten percent (1/10) of the subjects
dropped out because of major improvement in treatment group and no subject dropped out
because of clinical improvement in the placebo group. Gibbons, et al (1988) analyzed this
subsample, and they did not detect a significant treatment effect. However, from the
literature, there is no evidence suggesting that the effect of these anti-psychotic drugs have any interaction with age. There may be a couple of possible reasons for the discrepancy.

One interpretation would be that the power of the test is reduced by the relatively small sample size in subsample. If the effect size remains the same in the entire group (n = 436) as in the younger group (n = 95), by taking out about 341 subjects, the subsample could reduce the power to detect group difference by half ($\sqrt{95/436} = 0.47$).

There may be another reason for the non-significant results that is more relevant to this dissertation. In the younger group, the majority of incomplete subjects dropped out either due to treatment failure or clinical improvement. When treatment related dropouts occur, the distribution of the observations often resemble a truncated normal and therefore a method that ignores this aspect and assumes the usual normal model may lead to less accurate conclusions.

<table>
<thead>
<tr>
<th>Completion Status</th>
<th>Treatment (N=71)</th>
<th>Placebo (N=24)</th>
<th>Total (N=95)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Complete</td>
<td>61</td>
<td>86</td>
<td>18</td>
</tr>
<tr>
<td>Incomplete</td>
<td>10</td>
<td>14</td>
<td>6</td>
</tr>
<tr>
<td>Worsening</td>
<td>2</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>Improved</td>
<td>1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>MAR</td>
<td>7</td>
<td>70</td>
<td>1</td>
</tr>
</tbody>
</table>

**1.7 Prospectus**

This dissertation is motivated by this IMPS example. The main feature of this kind of data is that there is knowledge regarding why a missing value occurs. In some cases the patient dropped out is because his/her clinical condition either improved beyond a certain threshold or deteriorated beyond another threshold. Probability of nonresponse in this
situation could depend on the unobserved values of the outcome variable and hence are nonignorable (as defined in section 1.2). The observed values may be arising from a distribution truncated at a threshold and the missing data are missing due to truncation (MDT). In Chapter 2, a method (called as MDT method) specific for MDT data is summarized. In Chapter 3, the IMPS data are analyzed using the MDT method. Chapter 4 presents a simulation study to characterize the properties of the MDT method and compare the MDT with other commonly used imputation methods. Chapter 5 extends the MDT method to Rubin's multiple imputation method to account for the uncertainty about the correct value to impute. Chapter 6 concludes the dissertation with a summary comments and possible extensions to the work in this dissertation. Chapters 3 and 4 are written in journal article format and therefore contain some materials that also appear in other chapters.
Chapter 2

Missing Due to Truncation Method

2.1 Introduction

As discussed in section 1.6, when a subject’s clinical condition deteriorates or improves to such a degree that the subject has to drop out of the study, the observed data may be considered arising from a truncated normal distribution, and the missing data may be considered as missing due to truncation (MDT).

From the definition of MDT, it is clear that MDT is not independent of the unobserved data and therefore is nonignorable. Thus the traditional imputation methods including the multiple imputation available in existing software assuming missing as MAR may not be ideal. Ramakrishnan and Wang (2005) have proposed a method specific to MDT data using multivariate truncated normal distribution. It will be called as MDT method in this dissertation. The MDT method incorporates the information available from an individual’s observation prior to dropping out as well as the group level information at the time point the individual drops out. This is done using a multivariate truncated normal distribution. The MDT method is reviewed in this chapter. Although the material presented is a review of Ramakrishnan and Wang’ method, some of their notation has been changed to conform to the notation adopted in this dissertation.
2.2 The MDT data structure

Let \( n_t \) denote the number of individuals observed at time \( t \) and \( r_t \) denote the number of cases MDT (i.e., \( n_t + r_t = n \)). Consider first the case where the missing occurs at the last time point \( T \). Given that \( r_T \) observations are missing at time \( T \), the data matrix could be represented as,

\[
\begin{bmatrix}
  y_{11} & y_{12} & \cdots & y_{1T-1} \\
  y_{21} & y_{22} & \cdots & y_{2T-1} \\
  \vdots & \vdots & \ddots & \vdots \\
  y_{r_T} & y_{r_T} & \cdots & y_{r_T-1} \\
  y_{(r_T+1)1} & y_{(r_T+1)2} & \cdots & y_{(r_T+1)T-1} \\
  \vdots & \vdots & \ddots & \vdots \\
  y_{n1} & y_{n2} & \cdots & y_{nT-1} \\
\end{bmatrix}
= \begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_{r_T} \\
  y_{r_T+1} \\
  \vdots \\
  y_n \\
\end{bmatrix}
\]

The \( r_t, T-1 \) dimensional vectors \( y_1, y_2, \ldots, y_{r_T} \) correspond to the individuals with missing observations. They are independent and identically distributed \( T-1 \) variate random variables. The \( T \) dimensional vectors \( y_{r_T+1}^*, \ldots, y_n^* \) correspond to the individuals with complete data. They are independent and identically distributed \( T \) variate variables assumed to follow a truncated multivariate normal distribution. Specifically, the marginal distribution of the \( T \)th observation on the \( n - r_T \) individuals is considered to be from a truncated normal distribution, truncated at some threshold \( M \). Let \( \mu_t(\theta) \) denote a function representing the mean response vector of individuals at time \( t \), where \( \theta \) is an unknown, vector-valued parameter, which may be represented by \( X\beta + Z\gamma \), where \( X \) is a design
matrix for fixed effects and $Z$ is a design matrix for random effects, $\beta$ represents the fixed effects parameters and $\gamma$ represents the random effects parameters. Here, $\theta = (\beta, \Sigma_{\gamma})$, where $\Sigma_{\gamma}$ is the variance-covariance matrix of the random effect. The primary objective is to estimate and to test hypotheses of interest regarding the parameter $\theta$.

For the simplicity of presenting the method, consider first the case where the MDT is the only type of missing observations, and the MDT data are occurring at the right tail of the distribution at the last observed time point $T$. That is, the threshold, $M$, at which an individual drops out is assumed to be greater than the mean, $\mu_T(\theta)$, at time $T$.

2.3 Likelihood and Estimation

Since the missing data mechanism is nonignorable, the full likelihood needs to be maximized. Using the equation (1.1) for the non-ignorable case, the likelihood can be written as the joint distribution of $y_1, y_2, \ldots, y_{R_T}, y_{R_T + 1}^*, \ldots, y_m^*$ and missing indicators

$$I_1, I_2, \ldots, I_{R_T}, I_{R_T + 1}, \ldots, I_n$$

$$p(Y_{\text{obs}}, I | \theta, \xi) = p(y_1, y_2, \ldots, y_{R_T}, y_{R_T + 1}^*, \ldots, y_m^*, I_1, I_2, \ldots, I_{R_T}, I_{R_T + 1}, \ldots, I_n | \theta, \xi), \quad (2.2)$$

where

$$I_i = \begin{cases} 1 & \text{if } i \leq R_T \\ 0 & \text{if } i > R_T \end{cases}$$
Let $R_T = \sum_{i=1}^{n} I_i$ denote the random variable representing the number of individuals MDT at time $T$. Therefore $R_T$ is binomially distributed and is a sufficient statistic for $I_i$. Then (2.2) can be rewritten as the joint distribution of $Y_{obs} = (y_1', y_2', ..., y_{r_T}', y_{R_T}' + 1', ..., y_n')'$

and $R_T$ because the joint distribution of the $I_i$ is a function only of $R_T$. (This is similar to the likelihood for $p$ from Bernoulli trials can be written in the form of a binomial distribution.) To derive the likelihood one could write the joint probability distribution function as a product of conditional distributions as follows.

$$p(Y_{obs}, I | \theta, \xi) = p(y_1', y_2', ..., y_{R_T}', y_{R_T}' + 1', ..., y_n', R_T | \theta, \xi)$$

$$= p(y_{r_T}' + 1, T', ..., y_{nT} | y_1', y_2', ..., y_{R_T}', y_{R_T}' + 1', ..., y_n', R_T = r_T, \theta)$$

$$\times p(y_1', y_2', ..., y_{r_T}', y_{r_T}' + 1', ..., y_n' | R_T = r_T, \theta)P(R_T = r_T | \xi),$$

(2.3)

where $y_{r_T}' + 1, T', ..., y_{nT}'$ represent the $T$th observations of the random vectors $y_1', y_2', ..., y_n'$. When the $T$th observations from the random vectors $y_{r_T}' + 1', ..., y_n'$ are dropped, the vectors $y_{r_T}' + 1', ..., y_n'$ are also $T - 1$ dimensional and their distributions would be identical to the random vectors $y_1', y_2', ..., y_{r_T}'$. Further, the random vectors $y_{r_T}' + 1, T', ..., y_{nT}'$ and $y_1', y_2', ..., y_{r_T}'$ would be independent. Since, the individuals at any
given time point $t$ are independently distributed and since the joint distribution
of $y_1, y_2, \ldots, y_n$ is independent of $R_T$, the above joint distribution in equation (2.3) could
be further simplified as,

$$p(Y_{obs}, I | \theta, \xi) = p(y_1, y_2, \ldots, y_{r_T}, y_{r_T+1}, \ldots, y_n | R_T, \theta, \xi)$$

$$= p(y_{r_T+1}, \ldots, y_n, R_T = r_T, \theta) \times p(y_1, y_2, \ldots, y_n | \theta) P(R_T = r_T | \xi)$$

$$= \prod_{i=1}^{n-r_T} p(y_{r_T+i} | \theta) \prod_{i=1}^{n} p(y_i | \theta) P(R_T = r_T | \xi).$$

Now assume $y_1, y_2, \ldots, y_n$ are distributed as $T-1$ dimensional multivariate normal with
mean vector $\mu$ and variance-covariance matrix $\Sigma$. Further $I_i$'s are assumed to be Bernoulli
or equivalently $R_T$ is assumed to be binomially distributed. It will be shown that the
probability of missing (success) is a function of $\mu_T, \sigma^2_T$, and $M$, where $\mu_T$ and $\sigma^2_T$ are
the mean and the variance at time $T$ respectively and $M$ is the threshold beyond which a
patient will drop out. Let $\Sigma_T$ denote the covariance vector representing the covariances
between time point $T$ and the $T-1$ earlier observations. Then it can be shown that the
mean and the variance of the conditional truncated random variable $y_{r_T+i, T}$ conditional
on $y_{r_T+i}$ are

$$\mu^*_{T | i} = \mu_T | i - \sigma^2_T | i \xi_i(M)$$

(2.5)
and

\[
\sigma_T^* i^2 = \sigma_T^2 i \left[ 1 - \zeta_i(M) \left( \frac{M - \mu_T i}{\sigma_T i} + \zeta_i(M) \right) \right].
\]

where

\[
\mu_T i = \mu_T + \Sigma_T \Sigma^{-1}(y_{r_T} + i - \mu)
\]
\[
\sigma_T^2 i = \sigma_T^2 - \Sigma_T \Sigma^{-1} \Sigma_T
\]

and

\[
\zeta_i(M) = \frac{\phi \left( \frac{M - \mu_T i}{\sigma_T i} \right)}{\Phi \left( \frac{M - \mu_T i}{\sigma_T i} \right)}
\]

Here, \(\phi(\cdot)\) represents the standard normal pdf and \(\Phi(\cdot)\) represents the corresponding cdf.

Since \(\zeta_i(M)\) the normal density at \(M\) scaled by the cdf of normal distribution at \(M\) is positive, equation (2.5) implies that the mean at time point \(T\) would be \(\zeta_i(M)\) standard deviations larger than the mean of the truncated normal distribution. This shows that, when the data follow a truncated normal distribution, using the usual 'un-truncated' normal model could result in an underestimate of the mean at time point \(T\). This could lead to lack of power for detecting significant group differences.
The variable $R_T$, which represents the number MDT at time $T$, follows a binomial $(n, p_T)$ distribution with $p_T$, the probability that an individual will be MDT given by

$$p_T = 1 - \Phi \left( \frac{M - \mu_T}{\sigma_T} \right).$$

Thus, the likelihood function of equation (2.4) may be reduced to equation (2.8) in Ramakrishnan and Wang (2005). That is,

$$L(\mu_T, \sigma_T^2, \Sigma, \Sigma_T) \propto \frac{1}{\sigma_T^n} \frac{(n-r_T)^{n-r_T}}{\prod_{i=1}^{n-r_T} \Phi \left( \frac{M - \mu_T | i}{\sigma_T | i} \right)} \times \exp \left[ -\frac{1}{2\sigma_T^2} \sum_{i=1}^{n-r_T} \left( y_i - \mu_T | i \right) \left( y_i - \mu_T | i \right)^T \right] \times \frac{1}{|\Sigma|^{n/2}} \exp \left[ \frac{1}{2} \sum_{i=1}^{n-r_T} (y_i - \mu) \Sigma^{-1} (y_i - \mu) \right] \times \left( \begin{array}{c} n \\ r_T \end{array} \right) \left( 1 - \Phi \left( \frac{M - \mu_T}{\sigma_T} \right) \right)^{r_T} \left[ \Phi \left( \frac{M - \mu_T}{\sigma_T} \right) \right]^{n-r_T}.$$

where $\theta = (\mu_T, \sigma_T^2, \Sigma, \Sigma_T)$ and $\xi = (\mu_T, \sigma_T^2, M)$.

The maximum likelihood estimates (mle) could be obtained by taking the derivatives of the log likelihood with respect to the various parameters and equating to zero and solving simultaneously. However, since the likelihood involves cdf's of standard normal integrals the estimating equations would be non-linear and therefore will require iterative procedures such as the Newton-Raphson algorithm. The EM (Dempster, Laird and Rubin, 1977) algorithm is another alternative.
The EM algorithm simplifies the estimation procedure considerably. The EM algorithm is a general method of finding the maximum-likelihood estimate of the parameters from a given data set when the data is incomplete or have missing values (Knight, 1999, pp276). In the case of MDT, the observed data is the data matrix in (2.1), which will be denoted by \( S(y) \). It will be referred as incomplete data. The complete data would be obtained by adding \( y^* = (y_1^*, y_2^*, ..., y_r^*, T) \) to the observed sample. The joint density function of \( S(y) \) and \( y^* \) is

\[
p(S(y), y^* | \theta, \xi) = p(y^* | S(y), \theta, \xi) p(S(y) | \theta, \xi).
\]

In the EM algorithm the expected value of the complete data log-likelihood

\[
p(S(y), y^* | \theta, \xi) \text{ with respect to the unknown data } y^* \text{ given the observed data } S(y) \text{ and the current parameter estimates is first obtained. That is,}
\]

\[
E[\log p(S(y), y^* | \theta, \xi)] = E[Q(\theta, \xi; \theta^{i-1}, \xi^{i-1})],
\]

where \( \theta^{i-1} \) and \( \xi^{i-1} \) are the current parameter estimates that are used to evaluate the expectation and \( \theta \) and \( \xi \) are the new parameters that ultimately will be optimized in an attempt to maximize the likelihood. The evaluation of this expectation is the E-step of the EM algorithm.

The M-step of the EM algorithm is to maximize the log likelihood of the parameters given the 'complete' data obtained in the E-step. In general, for the \( i \)th iteration this may be written as,

\[
(\theta^i, \xi^i) = \max_{\theta, \xi} Q(\theta, \xi, \theta^{i-1}, \xi^{i-1}).
\]
These two steps are repeated until convergence. Each iteration is guaranteed to increase the log likelihood and the algorithm is guaranteed to converge to a local maximum of the likelihood function (Dempster, Laird and Rubin, 1977).

Given initial values of the various parameters, the expectation step for the \( i \)th MDT observation is achieved using the conditional truncated normal distribution defined as earlier. Notice that for the missing data the truncation is assumed to occur at \( M \) but at the opposite end of the distribution. (Please see Figure 2.1 for a graphical representation of this.) Thus,

\[
E\left[ y_{i,T} | y_i \right] = \mu_{0T|i} + \sigma_{0T|i} \bar{z}_{0i}(M),
\]

where \( \mu_{0T|i} \) and \( \sigma_{0T|i} \) are the initial values of

\[
\mu_{T|i} = \mu_T + \Sigma_T \Sigma^{-1}(y_i - \mu),
\]

\[
\sigma_T^2 = \sigma_T^2 - \Sigma_T \Sigma^{-1} \Sigma_T, \tag{2.9}
\]

and \( \bar{z}_{0i}(M) \) is the initial value of the expression

\[
\bar{z}_{i}(M) = \phi \left( \frac{M - \mu_{T|i}}{\sigma_{T|i}} \right) \quad \frac{1 - \Phi \left( \frac{M - \mu_{T|i}}{\sigma_{T|i}} \right)}{1 - \Phi \left( \frac{M - \mu_{T|i}}{\sigma_{T|i}} \right)} \tag{2.10}
\]
Once the observations are obtained from the E-step the M-step could be easily applied because the multivariate normal theory would apply and therefore explicit expressions for the means, variances and covariances would exist. This considerably improves on the Newton-Raphson procedure which will require computation of multivariate normal cdfs in every iteration. The EM algorithm, unlike the Newton-Raphson procedure does not provide an estimate of the variance-covariance matrix readily. If expressions for the second derivative of the log likelihood are available (which are needed for the Newton-Raphson procedure) the MLEs obtained from EM algorithm could be plugged into these expression to obtain the asymptotic estimate of the variance-covariance matrix from the observed information matrix.
Ramakishnan and Wang (2005) have proposed the initial estimate for the mean based on the middle part of the distribution. That is, the initial value could be,

\[ \mu_{0r} = \frac{1}{n - 2r_r} \sum_{i=r_r}^{n-r_r} y_{(i)r}, \]

where \( y_{(i)r} \) is the \( i \)th order statistic of the observed part of the sample at \( T \). Initial estimates for the variance and the covariance could be obtained using the sums of squares and products matrices based on the observed part of the data. Also, the initial estimate for \( M \) could be

\[ M_0 = \mu_{0r} + \sigma_{0r} \Phi^{-1} \left( 1 - \frac{r_r}{n} \right). \]

Better initial estimates for the mean and variance-covariance matrix could be based on repeated measures mixed model which treats MDT as missing at random. This will be further described in section 3.2.

2.4 Missing data at several time points

Suppose a subject reaches the threshold at time \( t \), the subject is presumed to be MDT for the remainder of the trial. Thus the MDT data are monotone missing. It allows imputing these observations sequentially. At \( t \) where the subject’s first MDT occurs, the procedure described in the previous section can be applied based on the observations at earlier time points from that subject and the available observations from other subjects up to time \( t \). Once the MDT at time \( t \) is imputed, treating the data as complete at \( t \), the same approach can be applied to impute the MDT at time \( t + 1 \). This can be continued until the MDT at last time point is imputed. Since at each time point, the numbers of MDT, \( r_2, ..., r_T \),
are not necessarily equal and the mean and variance may vary, the truncation values $M_2, M_3, \ldots, M_T$ may be different (equation (2.6)). Notice that the method is applicable only for occurrence of MDT from time points 2 and on.

2.5 Data missing due to truncation at both tails of the distribution

In reality, both tails of the distribution can be truncated, in which some subjects drop out due to exceeding a threshold value, and others due to falling below another threshold value. The general form of a normal distribution truncated in both tails is:

$$f_t(y) = \begin{cases} \phi(\frac{y - \mu_t}{\sigma_t}) & \text{if } M_{t,\text{low}} \leq y \leq M_{t,\text{high}} \\ \phi(\frac{M_{t,\text{high}} - \mu_t}{\sigma_t}) - \phi(\frac{M_{t,\text{low}} - \mu_t}{\sigma_t}) & \text{if } M_{t,\text{low}} > y \\ 0 & \text{otherwise} \end{cases}$$

Here, $M_{t,\text{low}}$ is the lower bound of the distribution at time $t$ and $M_{t,\text{high}}$ is the upper bound at time $t$. In the situation where data are missing due to truncation at both tails of the distribution, MDT method still can be applied. For each EM algorithm iteration, the estimation of the MDT has to be performed for both of the tails using the lower and upper truncation values separately. Suppose at time $t$, $r_{i,\text{low}}$ subjects are MDT at lower tail of the distribution, and $r_{i,\text{high}}$ subjects are MDT at upper tail. Arbitrarily choose one tail, say, the upper tail to start with. The missing portion is the set in which there are $r_{i,\text{high}}$ subjects whose measurements are higher than $M_{t,\text{high}}$. That is, the upper bound $M_{t,\text{high}}$ at time $t$ is given by

$$p(y_{it} > M_{t,\text{high}}) = 1 - \Phi(\frac{M_{t,\text{high}} - \mu_t}{\sigma_t}) = \frac{r_{i,\text{high}}}{N},$$
and the MDT at the upper tail of the distribution could be estimated as

\[ E \left[ y^*_{i,T} \mid y_i \right] = \mu_{T \mid i} + \sigma_{T \mid i} \xi_i(M_{t,\text{high}}), \]

where \( \xi_i(M_{t,\text{high}}) \) could be estimated from equation (2.10).

Once the MDT at the upper tail are estimated, the lower bound \( M_{t,\text{low}} \) at time \( t \) is given by

\[ p(y_{it} < M_{t,\text{low}}) = \Phi \left( \frac{M_{t,\text{low}} - \mu_i}{\sigma_i} \right) = \frac{r_{t,\text{low}}}{N}, \]

and MDT at the lower tail could be estimated as

\[ E \left[ y^*_{i,T} \mid y_i \right] = \mu_{T \mid i} - \sigma_{T \mid i} \xi_i(M_{t,\text{low}}), \]

where

\[ \xi_i(M_{t,\text{low}}) = \frac{\Phi \left( \frac{M_{t,\text{low}} - \mu_T \mid i}{\sigma_T \mid i} \right)}{\Phi \left( \frac{M_{t,\text{low}} - \mu_T \mid i}{\sigma_T \mid i} \right)} \]

### 2.6 MDT with MAR and MCAR

The generalization of MDT method to include MAR and MCAR are straightforward. Since repeated measures mixed models are commonly used to fit the data with MDT, MAR and MCAR cases automatically can be taken care of due to the specification of covariance structure in the repeated measure analysis. This will be discussed further in next section.
2.7 Other methods useful for longitudinal missing data

There are several widely accepted longitudinal imputation methods that are useful for dealing with the MDT situation. The goal of any imputation technique is to produce a complete data set, which can then be analyzed using complete-data inferential methods. Some methods are based on a single imputation such as last observation carried forward, individual regression prediction, and repeated measures mixed models. In addition to single imputation, Rubin’s multiple imputation method (1987) is becoming more widely accepted. Every imputation method implicitly or explicitly assumes a model for the missing data. These methods are briefly described here and further explored in Chapter 4.

2.7.1 Last observation carried forward method

It assigns the person’s last previous known observation to the missing value. In other words, for a subject with missing value of a particular variable at time \( t \), the missing value is imputed by his/her last observed value of that variable prior to drop out. The underlying assumption is that the observations at later time points won’t change after the subject drops out. This method is appropriate if the subject’s response tends to stabilize after a period of time. For example, suppose the response function over time is exponential. When \( t \) is large enough, the model \( y_{i,t} = y_{i,t-1} \) is a fair approximation. If the assumption that \( y_{i,t} \) is approximately equal to \( y_{i,t-1} \) doesn’t hold, the last observed value would be an inappropriate guess of the missing values. This method could lead to biased mean estimates and a biased low standard error (Shih, 2002).
2.7.2 Individual regression prediction method

This method fits a regression line between the outcome variable and time for each subject with missing value,

\[ y_{it} = \beta_{i0} + \beta_{i1} t + \epsilon_{it}, \]

where the parameters \( \beta_{i0} \) and \( \beta_{i1} \) vary over different individuals with missing data. The individual regression prediction method extrapolates missing observations based on the regression fit. This method assumes a linear response between the outcome variable and time. In practice, exploring the form of the response functions and obtaining its parameters is often one of the goals of clinical trials. Imposing an arbitrary linear response function could lead to estimation bias. For example, if the measurements over time follow a concave quadratic function, the mean of response at last time point would be overestimated. Moreover, this bias could be worse if few values are observed before the missing value.

2.7.3 Repeated measures mixed model

Repeated measures mixed model analysis assumes that missing data are MAR. Thus it ignores the information available when the data are MDT (Laird, 1988). Due to the specification of the covariance structure, observations at each time point influence parameter estimates at every other time point. Thus, repeated measures mixed model analysis uses all available data. The information from incomplete individuals whose observations are limited to early time points will be taken into account when estimating parameters at later time points. Further, since the repeated measures mixed model includes
the information from the incomplete data implicitly the imputation of the missing data is unnecessary.

2.7.4 Multiple Imputation

The advantage of single imputation is it allows most standard methods of analysis to be used. However, for single imputation, the variability due to the unknown missing values is not taken into consideration. Thus, quantities that depend on the variability of the variables such as correlations and covariance can be badly biased. Multiple imputation first was proposed by Rubin (1977) and then elaborated in his book (1987) as a way to address this issue associated with single imputation. Rubin’s multiple imputation method appears to be one of the most attractive methods for general purpose of handling missing data in multivariate analysis. Instead of imputing one value for each missing observation, this method suggests multiple (say \( m \), usually less than 10) imputation values be created to form multiple complete data sets. Then standard complete data analysis can be performed on each complete data set. In principle, the \( m \) imputations of the missing values are \( m \) random draws from the posterior predictive distribution of the missing values. The point estimate of the summary statistics from the \( m \) imputations is calculated as the average of the \( m \) imputations. The variance of the estimates are from two components, one is within imputation variance, calculated as the average variance of the \( m \) imputations, the other is between imputation variance, calculated as the difference between the summary statistic of each imputation and the average of the summary statistics of the \( m \) imputation. Thus the combined variance accounts for the uncertainty due to estimating the missing values. Multiple imputation method was criticized for its computing intensiveness
for imputing multiple data sets, testing models for each data set separately and
recombining the model results. With the development in the computation technology, the
computing time is no longer an issue. As for the repeated measures mixed model, multiple
imputation also assumes that missing data are MAR.

2.7.5. Treating missing as censored

One way to handle MDT data that has been proposed in the literature is to treat this
type of missing as censored data. The approach presented here is different from the
censoring approach. In general, there are two types of censoring (Klein & Moeschberger,
2003, pp63-72). Type I censoring assumes that if a subject is censored, the data (time to
event) for that subject have to be greater (left censoring) or less (right censoring) than a
prespecified censoring value. The MDT differs from this because time to event is not the
variable of interest. Also, the MDT method does not assume that the observations
necessarily increase or decrease monotonically over time. Once an individual’s
measurement passes the threshold, the only assumption made is that they remain beyond
the threshold. Moreover, when missing occurs at more than one time point, thresholds
may vary. For example, in the IMPS data, someone who drops out at week three may have
a lower tolerance than someone who drops out at week six.

In Type II censoring, n subjects are followed until the first r failures occur where r
is a predetermined integer (r < n), the remaining n - r subjects are considered censored at
that time point. In MDT data such as IMPS data this is obviously not the case. The
distinction made here between censoring and truncation is crucial when constructing the
likelihood. In some application where time to an event is the observation of interest,
censoring might be used (Scharfstein, 2005). However in the MDT case the truncation model would be the most appropriate.

In Chapter 3, the MDT method is applied to the IMPS dataset introduced in Section 1.6 to show how the analysis can be improved by incorporating the missing data mechanism. Chapter 4 will carry out a simulation study to compare the MDT method with last observation carried forward, individual regression prediction, and the repeated measures mixed model.
Chapter 3

Analysis of IMPS Data by MDT Method

3.1. Introduction

In this chapter, the IMPS data introduced in Section 1.6 are analyzed to illustrate the application of the MDT method. The IMPS data were collected in an NIMH schizophrenia collaborative study on treatment related changes in severity of illness. Specifically, Item 79 of the Inpatient Multidimensional Psychiatric Scale (IMPS; Lorr & Klett, 1966) was examined. Item 79 (severity of mental illness) was scored as 1 if normal, or not at all ill, 2 if borderline ill, 3 if mildly ill, 4 if moderately ill, 5 if markedly ill, 6 if severely ill and 7 if among the most extremely ill. In this study, patients were randomly assigned to receive one of four medications: placebo, chlorpromazine, fluphenazine, or thioridazine. Since the previous analyses revealed similar effects for the three anti-psychotic drugs (Gibbons & Hedeker, 1988), and the main goal of the study was to examine if the placebo effect was as good as any treatment, the three drug groups were combined to form the treatment group in the analysis presented here. The patients were followed up at four time points (baseline, weeks one, three and six) and the IMPS scores were collected at each time point.
A subset of this data set for the subjects whose ages range from 16-21 was of particular interest to the investigator. Since the full data set has been analyzed extensively in the literature (Gibbons et al, 1988, Gibbons & Hedeker, 1994) this subset is considered for additional scrutiny. The IMPS data was summarized in Table 1.2 and the original data are listed in Appendix A. As pointed out in Section 1.6, in the treatment group, 14% (10/71) subjects did not complete the study. Among the incomplete ones, 30% (3/10) of the missing subjects were non-ignorable missing of which two dropped out due to clinical deterioration (they would have had high IMPS scores if they had stayed in the study), and one had major improvement (he/she would have had low IMPS scores if observed). In the placebo group, 25% (6/24) subjects didn't complete the study. Among the incomplete ones, 83% (5/6) dropped out due to clinical deterioration and 17% (1/6) dropped out for some unknown reason and therefore will be treated as MAR. The observed data for the placebo and the treatment group are displayed in Figures 3.1 and 3.2. Complete cases are plotted as green lines, MAR cases as orange lines, deteriorating cases as red lines and improving case as blue line.

A histogram of the marginal distributions at weeks three and six are clearly skewed (see Appendix A) and the skewness perhaps is a result of the truncation caused by the non-ignorable missing patients. Any method that assumes multivariate normality and treats the missing observations as MAR could produce misleading results. In this chapter, the IMPS data will be reanalyzed by the MDT method under the appropriate multivariate truncated normal model. The main purpose of the analysis is to study, in this younger group,
whether or not the IMPS score decreases over time differently between the treatment group and placebo group.

A repeated measures mixed model with time, treatment group and time by treatment group interaction was fitted using the MDT method. In the application of the EM algorithm, the IMPS scores from each individual were assumed from a multivariate distribution. The MDT observations were estimated in the E-step using the procedure described in Section 2.3. The repeated measures mixed model was applied in the M-step. The advantage of the likelihood-based approach of mixed models is that it can accommodate data that are missing at random (MAR). Since the response function is likely to vary over treatment groups, the variance-covariance matrices for the two groups were allowed to vary.

Figure 3.1 Placebo group for ages 16-21 years
3.2 Application of the MDT method

The flowchart in Figure 3.3 describes the steps of the MDT method. The numbers in the boxes within parentheses represent the steps of the method. These steps are described below.

In the first step initial values of the parameters including mean and variance-covariance and truncation value $M$ for placebo and treatment group need to be provided. Although the EM algorithm is less sensitive to the initial values than most other iterative algorithms such as Newton-Raphson algorithm the convergence may be achieved faster by obtaining them from a repeated measures model treating all the missing data as MAR. For example, in SAS, one could use the following code (Program 1) to obtain the initial values.
Program 1. Initial Estimates

PROC MIXED DATA=WORK;
    CLASS TIME ID;
    MODEL IMPS=GROUP TIME GROUP*TIME OUTPM=PREDICTED S;
    REPEATED TIME TYPE=AR(1) SUB=ID R=1 95 GROUP=GROUP;
    ODS OUTPUT R=R;
    RUN;

The within subject variance-covariance structure has been set to an autoregressive covariance structure AR(1). However, PROC MIXED provides a variety of other covariance structures such as Compound Symmetry (CS), Unstructured (UN), and Spatial Power (SP(POW)). The GROUP = option allows the variance-covariance matrices in the treatment and placebo groups to be different. The ODS OUTPUT R = option produces a data set R containing variance-covariance estimates for the treatment and placebo group. The OUTPM=option produces a data set containing predicted means calculated from the equation

\[ \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 \text{Group} + \hat{\beta}_2 \text{Time} + \hat{\beta}_3 \text{Group} \times \text{Time}. \]

The initial value for the truncation threshold \( M \) is estimated for the treatment and the placebo groups separately as follows. In the placebo group (Table 3.1), five MDT cases are due to the IMPS score exceeding the truncation threshold. Therefore the initial value for \( M \) at week \( t \), for \( t = 3, 6 \), is estimated by,

\[ M_{P0t} = \mu_{P0t} + \sigma_{P0t} \Phi^{-1} \left( 1 - \frac{r_{Pt}}{n_p} \right), \]  

(3.1)

where \( r_{Pt} = 5 \), \( n_p = 24 \), \( \mu_{P0t} \) and \( \sigma_{P0t} \) are respectively the initial values of the IMPS mean and the standard deviation at time \( t \).
In the treatment group (Table 3.2), the MDT occurs at both tails of the distribution. Two sets of $M$ need to be estimated. For the MDT at the upper tail of the distribution, estimation of the truncation value is similar to the placebo group. For the MDT at the lower tail of the distribution (i.e., missing due to improvement), the $M$ are estimated by

$$M'_{T0t} = \mu_{T0t} - \sigma_{T0t} \Phi^{-1} \left( \frac{r_{Tt}}{n_T} \right),$$

where $r_{Tt} = 1$, $n_T = 71$, $\mu_{T0t}$ and $\sigma_{T0t}$ are respectively the initial values of the IMPS mean and the standard deviation at time $t$.

<table>
<thead>
<tr>
<th>Subject</th>
<th>ID</th>
<th>Weeks 0</th>
<th>1</th>
<th>3</th>
<th>6</th>
<th>Reasons for Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2106</td>
<td>6.00</td>
<td>5.00</td>
<td>.</td>
<td>.</td>
<td>Failure</td>
</tr>
<tr>
<td>2</td>
<td>2149</td>
<td>6.00</td>
<td>6.00</td>
<td>.</td>
<td>.</td>
<td>Failure</td>
</tr>
<tr>
<td>3</td>
<td>2320</td>
<td>6.50</td>
<td>7.00</td>
<td>.</td>
<td>.</td>
<td>Failure</td>
</tr>
<tr>
<td>4</td>
<td>6105</td>
<td>5.00</td>
<td>5.00</td>
<td>.</td>
<td>.</td>
<td>Failure</td>
</tr>
<tr>
<td>5</td>
<td>6116</td>
<td>6.00</td>
<td>5.00</td>
<td>.</td>
<td>.</td>
<td>Failure</td>
</tr>
<tr>
<td>6</td>
<td>3320</td>
<td>6.00</td>
<td>1.00</td>
<td>.</td>
<td>.</td>
<td>MAR</td>
</tr>
</tbody>
</table>

Table 3.2 Subjects with missing data in treatment group

<table>
<thead>
<tr>
<th>Subject</th>
<th>ID</th>
<th>Week 0</th>
<th>1</th>
<th>3</th>
<th>6</th>
<th>Reasons for Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2372</td>
<td>6.00</td>
<td>6.00</td>
<td>5.00</td>
<td>.</td>
<td>Failure</td>
</tr>
<tr>
<td>2</td>
<td>2121</td>
<td>5.00</td>
<td>5.00</td>
<td>5.00</td>
<td>.</td>
<td>Failure</td>
</tr>
<tr>
<td>3</td>
<td>4506</td>
<td>5.50</td>
<td>4.50</td>
<td>2.50</td>
<td>.</td>
<td>Improve</td>
</tr>
<tr>
<td>4</td>
<td>2113</td>
<td>5.0</td>
<td>6.00</td>
<td>.</td>
<td>.</td>
<td>MAR</td>
</tr>
<tr>
<td>5</td>
<td>2123</td>
<td>6.00</td>
<td>5.5</td>
<td>2.50</td>
<td>.</td>
<td>MAR</td>
</tr>
<tr>
<td>6</td>
<td>2331</td>
<td>4.5</td>
<td>4.5</td>
<td>.</td>
<td>.</td>
<td>MAR</td>
</tr>
<tr>
<td>7</td>
<td>3308</td>
<td>5.5</td>
<td>5.0</td>
<td>.</td>
<td>.</td>
<td>MAR</td>
</tr>
<tr>
<td>8</td>
<td>3314</td>
<td>5.5</td>
<td>5.0</td>
<td>.</td>
<td>.</td>
<td>MAR</td>
</tr>
<tr>
<td>9</td>
<td>4704</td>
<td>6.0</td>
<td>1.5</td>
<td>.</td>
<td>.</td>
<td>MAR</td>
</tr>
<tr>
<td>10</td>
<td>4711</td>
<td>6.5</td>
<td>2.5</td>
<td>.</td>
<td>.</td>
<td>MAR</td>
</tr>
</tbody>
</table>
Data with observations MDT and MAR

(1) Obtain initial values for means, covariance parameters from repeated measures model. Determine truncation threshold.

(2a) E-step
Estimate observations MDT in placebo group at each time point

(2b) E-step
Estimate observations MDT in treatment group at each time point sequentially.

(3) M-step
Update the parameters from repeated measures model using complete data.

(4) Is sum of difference of parameters between the iterations < tolerance
   No

   Yes

(5) Run repeated measures model to the complete data to estimate model parameters and to test hypothesis.

Figure 3.3 Flowchart of MDT method
Table 3.3 The initial values with AR(1) covariance structure

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Group</th>
<th>0</th>
<th>1</th>
<th>3</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Weeks</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.2472</td>
<td>4.6579</td>
<td>4.2264</td>
<td>3.8036</td>
</tr>
<tr>
<td>(\mu)</td>
<td>Placebo</td>
<td>5.2652</td>
<td>4.5231</td>
<td>3.9799</td>
<td>3.4475</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\Sigma)</td>
<td></td>
<td>1.4439</td>
<td>0.6216</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Given the initial values, the EM algorithm constructs an estimate of the complete data likelihood function (Figure 3.3 (2a) & (2b)) and then maximizes this likelihood to obtain new parameter estimates (Figure 3.3 (3)). This two-step iterative procedure is then repeated until convergence.

A SAS Macro in PROC IML (SAS 9.1) is used to implement the EM algorithm. The code for the E-step is written specifically for the MDT procedure. The M-step, however, is achieved by calling the PROC MIXED procedure from SAS.

The E-step for the MDT data is performed for the treatment and placebo groups separately. Starting from week three, where the first MDT values occur, the mean and variance for MDT conditioning on the individual’s observations at baseline and week one are estimated by

\[
\mu_{P3|i}^{(k)} = \mu_{P3}^{(k)} + \Sigma_{P3}^{(k)} \Sigma_{P}^{-1}(y_{P3|i}^{(k)} - \mu_{P3}^{(k)}),
\]

\[
\sigma_{P3|i}^{2} = \sigma_{P3}^{2} - \Sigma_{P3}^{(k)} \Sigma_{P}^{-1}(y_{P3|i}^{(k)} - \mu_{P3}^{(k)}).
\]

The MDT at week three for that subject is estimated by

\[
E\left[y_{P|3}^{(k)} \right] = \mu_{P3|3}^{(k)} + \Sigma_{P3|i}^{(k)} \Sigma_{P3}^{-1}(\mu_{P3}^{(k)}),
\]

(3.2)
where

\[
\phi \left( \frac{M^{(k)} - \mu^{(k)}_{P3}}{\sigma^{(k)}_{P3|i}} \right) \frac{M^{(k)}_{P3} - \mu^{(k)}_{P3|i}}{\sigma^{(k)}_{P3|i}} \right)

1 - \Phi \left( \frac{M^{(k)}_{P3} - \mu^{(k)}_{P3|i}}{\sigma^{(k)}_{P3|i}} \right)
\]

Once the five MDT cases at weeks three are estimated, data at weeks three is considered as complete. Since there were no new MDT cases at weeks six, the E-step need to be performed for the same five subjects with MDT at week three. The same procedure is applied by conditioning now on baseline to week three instead of baseline and week one.

The treatment group is more complicated than the placebo group because the MDT occurs on both tails. There are a total of three MDT cases in the treatment group (Table 3.2). Two are missing due to high IMPS score (deterioration) and the third is missing due to low IMPS score (improvement). For deterioration cases, the MDT is estimated similar to the placebo group. Once the estimation for the deterioration cases is completed, the estimation for the improvement case is estimated as

\[
E \left[ y^{(k)}_{T6|i} \right] = \mu^{(k)}_{T6|i} - \sigma^{(k)}_{T6|i} \phi^{(k)}_{T6|i} (M^{(k)}_{T6L}),
\]

(3.3)

where
The IMPS scores range from 1 to 7 (Gibbons, 1988), therefore, if the estimate of the observation MDT from equation (3.2) is larger than the upper bound 7 then it is set to be 7. Similarly, if the observation MDT from equation (3.3) is less than the lower bound 1, it is set to be 1.

In the M-step the repeated measures model is applied (Figure 3.3, (3)). In order to apply PROC MIXED, IMPS scores are strung out into one long response variable IMPS. That is, the dataset includes four variables: Subject ID, IMPS, Group and Time indicating the follow-up time. The updated \((k+1)\)th mean and variance-covariance structure are estimated by maximizing the restricted maximum likelihood (REML) for the treatment and placebo groups (Jennrich & Schluchter, 1986).

For the placebo group, the \((k + 1)\)th upper truncation value \(M_{pi}^{(k+1)}\) is estimated as

\[
M_{pi}^{(k+1)} = \mu_{pi}^{(k+1)} + \sigma_{pi}^{(k+1)} \Phi^{-1}\left(1 - \frac{r_p}{n_p}\right), \quad t = 3, 6,
\]

where \(r_p = 5\) and \(n_p = 24\).

For the treatment group, the \((k + 1)\)th upper truncation value \(M_{nu}^{(k+1)}\) is estimated as
where \( r_{TMU} = 1, r_{TUU} = 2 \) and \( n_r = 71 \).

The \((k+1)\text{th}\) lower truncation value is estimated as

\[
M_{Tn}^{(k+1)} = \mu_{Tn}^{(k+i)} + \sigma_{Tn}^{(k+i)}\Phi^{-1}\left(1 - \frac{r_{nU}}{n_T}\right), \quad t = 3, 6,
\]

where \( r_{nL} = 1 \) and \( n_r = 71 \).

For stopping the iterative procedure, take the absolute differences between the \(k\)th parameters and \((k+1)\)th parameters (Figure 3.3, (4)). If the sum of the differences is less than a tolerance value, say \(10^5\), the EM algorithm is stopped. When this is achieved, the model is refitted and the group by treatment effect is tested. Otherwise, the algorithm loops back to step 2a and 2b to estimate the MDTs with updated \((k+1)\)th parameter estimates. It took about 10-20 iterations to have the EM algorithm converge, under different models and variance-covariance structures, for the IMPS data.

Several model structures were fitted for the IMPS data. Different variance covariance structure such as CS, UN, AR(1) and SP(POW) were compared. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used to compare different models. The smaller the AIC and BIC are, the more appropriate the model is.

3.3 Analysis results by MDT method

First, the Time effect was treated as categorical variable (Program 2) and different variance-covariance structures were compared.
Program 2. Time as Categorical Variable

PROC MIXED DATA=COMPLETE;
   CLASS ID TIME GROUP;
   model IMPS=GROUP GROUP*TIME TIME/ddf=,254,254 s;
   repeated TIME/TYPE=AR(1) SUBJECT=ID GROUP=GROUP ;
RUN;

The DDF=option specifies the denominator degree of freedom (df) of the F-test for the Group*Time interaction. It is calculated by 268 – 14 = 254, where 268 is the df obtained from SAS PROC MIXED if there is no MDT, 14 is the estimated MDT observations (10 from placebo, 4 from treatment).

Covariance structure specification in PROC MIXED is important because the test statistics for the fixed effects are functions of it, and PROC MIXED can produce invalid results if the structure is misspecified (Wolfinger, 1993). Table 3.4 shows, both AIC (1085.7) and Schwarz's Bayesian Criterion (1095.9) are smaller for the AR(1) structure compared to the other structures (CS, UN and SP(POW)). This indicates that the model with AR(1) structure is most appropriate. Under the AR(1) structure the estimate of the Group by Time interaction is not significant (P value = 0.2084).

<table>
<thead>
<tr>
<th>Covariance structure</th>
<th>AIC(smaller is better)</th>
<th>BIC(smaller is better)</th>
<th>Den df</th>
<th>F value</th>
<th>P value</th>
<th># of iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP(POW)</td>
<td>1120.8</td>
<td>1131.0</td>
<td>254</td>
<td>1.29</td>
<td>0.2791</td>
<td>11</td>
</tr>
<tr>
<td>CS</td>
<td>1103.8</td>
<td>1114.0</td>
<td>254</td>
<td>3.02</td>
<td>0.0305</td>
<td>10</td>
</tr>
<tr>
<td>UN</td>
<td>1088.6</td>
<td>1139.7</td>
<td>254</td>
<td>1.91</td>
<td>0.1287</td>
<td>19</td>
</tr>
<tr>
<td>AR(1)</td>
<td>1085.7</td>
<td>1095.9</td>
<td>254</td>
<td>1.53</td>
<td>0.2084</td>
<td>11</td>
</tr>
<tr>
<td>MAR AR(1)</td>
<td></td>
<td></td>
<td>250</td>
<td>0.34</td>
<td>0.7978</td>
<td>NA</td>
</tr>
</tbody>
</table>

1Last row shows the results if all missing are treated as MAR.

Second, Hedeker and Gibbons (1997) studied the data extensively and showed that although the relationship of the IMPS score over time is not linear, the square root
transformation of time can linearize the relationship of the IMPS score over time. Thus the following model with the square root transformation of time as continuous independent variable was fitted and different variance-covariance structures were compared.

Program 3. Square Root of Time as Continuous Variable

```sas
PROC MIXED DATA=COMPLETE;
  CLASS ID TIME GROUP;
  MODEL IMPS=GROUP GROUP*STIME STIME/ddf=, 258, 258 s;
  REPEATED TIME/TYPE=AR(1) SUBJECT=ID GROUP=GROUP;
RUN;
```

where Stime is the square root transformation of Time. Different covariance structures were compared (Table 3.5). By the AIC and BIC criteria, the AR(1) structure fits the IMPS data the best. The Group by Time interaction in this case is marginally significant (P value=0.0572).

<table>
<thead>
<tr>
<th>Covariance structure</th>
<th>AIC(smaller is better)</th>
<th>BIC(smaller is better)</th>
<th>Den df</th>
<th>F value</th>
<th>P value</th>
<th># of iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP(POW)</td>
<td>1118.3</td>
<td>1128.6</td>
<td>258</td>
<td>3.06</td>
<td>0.0815</td>
<td>10</td>
</tr>
<tr>
<td>CS</td>
<td>1099.7</td>
<td>1109.9</td>
<td>258</td>
<td>8.25</td>
<td>0.0044</td>
<td>9</td>
</tr>
<tr>
<td>UN</td>
<td>1085.8</td>
<td>1136.9</td>
<td>258</td>
<td>3.79</td>
<td>0.0527</td>
<td>20</td>
</tr>
<tr>
<td><strong>AR(1)</strong></td>
<td><strong>1083.4</strong></td>
<td><strong>1093.6</strong></td>
<td><strong>258</strong></td>
<td><strong>3.65</strong></td>
<td><strong>0.0572</strong></td>
<td><strong>10</strong></td>
</tr>
<tr>
<td>MAR AR(1)</td>
<td></td>
<td></td>
<td>254</td>
<td>1.01</td>
<td>0.3163</td>
<td>NA</td>
</tr>
</tbody>
</table>

1last row shows the results if MDT is treated as MAR.
Some diagnostic plots such as residuals vs. predicted plot and normal quantile plot were conducted to check the model assumptions for the models under MDT method. No obvious pattern or significant departure from the model assumption was detected. This also provides a justification for treating IMPS score as continuous variable.

Compared to repeated measures analysis which treats all the missing as MAR, the significance level improved from 0.7978 to 0.2084 by using MDT method with categorical Time variable in the model and from 0.3163 to 0.0572 by using MDT method with the square root of Time in the model. Compared to the model with categorical Time variable, the model with continuous sqrt(Time) has smaller AIC (1083.4 vs. 1085.7) and
BIC(1093.6 vs. 1095.9), which indicates the continuous model is preferable. In this model, the Group by Stime interaction is marginally significant (P value=0.0572).

3.4 Discussion

Wang et al. (1995) analyzed the IMPS data using PROC GLM with the MANOVA statement (SAS, 6.0) where the time effect was characterized as a categorical variable. In Wang’s analysis, the MDT cases were estimated under the truncated multivariate model but it differs from the MDT method as follows. First, only the MDT cases were estimated in his analysis. Second, instead of using PROC MIXED, the PROC GLM was used and the PROC GLM with MANOVA option includes a subject only if the IMPS score at each time point is nonmissing for that subject. Therefore the subjects with MAR were discarded in his analysis. Third, the MDT cases were estimated by the moment estimators of the truncated normal distributions instead of the MLE. Ramakrishnan and Wang (2005) analyzed these data using the EM algorithm. Once again, PROC GLM with MNOVA statement was applied in their analysis.

In the analysis presented here, the MLEs, which have better asymptotic properties than the moment estimators, were obtained using the EM algorithm. Further, by fitting the repeated measures model using PROC MIXED in the M-step, instead of using PROC GLM with the MANOVA statement, the MAR cases were also adequately dealt with.

In summary, the MDT method was applied to the IMPS data. Two repeated measures models with AR(1) variance-covariance structure were selected. One treats Time as categorical variable, the other treats Time as continuous variable and its square root transformation was found to provide a better fit. Both of the models appropriately
incorporated the missing data mechanism into the analysis and the ability to detect the
treatment effect was improved. The underlying statistical conclusion of the analysis was
that the antipsychotic treatment is marginally significant in the age group 16-21 (p
value=0.0572).
Chapter 4
Simulation Study for MDT Method

4.1 Introduction

In longitudinal clinical trials, when a subject's clinical condition deteriorates or improves to such a degree that the subject drops out of the study, the observed data may be considered arising from a truncated normal distribution, and the missing data may be considered missing due to truncation (MDT). In an earlier article, Ramakrishnan and Wang (2005) proposed a method specific to MDT data using the multivariate truncated normal distribution. By estimating the parameters of interest from a likelihood appropriately defined under the MDT situation, the MDT method incorporates the information available from an individual's observations prior to dropping out as well as the group level information up to the time point the individual drops out.

In this chapter a simulation study was carried out primarily to study the properties of the MDT method and to compare the performance of other relevant methods. The factors chosen for the simulation study included, missing data mechanisms (MDT and MAR), forms of response functions (linear, concave and convex), missing time points (missing at last time point and missing at last two time points), dropout rates (5%, 10% and 20%), sample sizes (50, 100 and 200) and correlations (0.2, 0.4 and 0.8) in an AR(1)
covariance structure. These were performed under a four dimensional multivariate model with each dimension representing the observations from a different time point. The three other methods considered for comparison are last observation carried forward (LOCF), individual regression prediction (REG) and repeated measures mixed model (MIXED). These methods were compared in terms of bias and mean square error of the estimates of the parameters, namely mean, variance and correlation.

4.2 The MDT method

The MDT method is briefly restated in this section. For detailed description of MDT method, refer to Chapter 2 and Ramakrishnan and Wang’s article (2005).

At time \( t \) let the observations \( y_{it}, i = 1, 2, \ldots, n \) denote a sample of size \( N = n \) from a population with multivariate truncated normal distribution. Let \( r_t \) denote the number of cases MDT. Let \( M \) denote a threshold beyond which individuals would drop out. Once a subject's measurement passes the threshold \( M \), the subject is presumed MDT for the remainder of the trial. Assuming monotone missing, \( r_t \) will be greater than \( r_{t'} \) for \( t \) greater than \( t' \). Let \( \mu_t(\theta) \) denote a function representing the mean response vector of individuals at time \( t \), where \( \theta \) is an unknown, vector-valued parameter. The \( \mu_t(\theta) \) may represent a linear model of the form \( X\beta \), where \( X \) is a known design matrix and \( \beta \) is a fixed parameter vector. If random effects are present, it could also be of the form \( X\beta + Z\gamma \), where \( Z \) is a design matrix for random effects, and \( \gamma \) represents the random effects parameters. Here, \( \theta = (\beta, \Sigma_\gamma) \), where \( \Sigma_\gamma \) is the variance-covariance matrix of the random
effects. The primary objective is to estimate and to test hypotheses of interest regarding the parameter \( \theta \).

Without loss of generality assuming the threshold \( M \) at which an individual drops out is smaller than the mean, \( \mu_T(\theta) \), and the first \( r_T \) observations are missing at the last time point \( T \), the data matrix is written,

\[
\begin{bmatrix}
y_{11} & y_{12} & \cdots & y_{1T-1} \\
y_{21} & y_{22} & \cdots & y_{2T-1} \\
\vdots & \vdots & \ddots & \vdots \\
y_{\tau_1} & y_{\tau_2} & \cdots & y_{\tau T-1} \\
y_{(\tau+1)1} & y_{(\tau+1)2} & \cdots & y_{(\tau+1)T-1} \\
\vdots & \vdots & \ddots & \vdots \\
y_{n1} & y_{n2} & \cdots & y_{nT-1} \\
\end{bmatrix}
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_{\tau_T} \\
y_{\tau_T+1} \\
\vdots \\
y_n \\
\end{bmatrix}
\]

(4.1)

The \( T - 1 \) dimensional vectors \( y_1^*, y_2^*, \ldots, y_{r_T}^* \) are assumed independent identically distributed multivariate variables. The \( T \) dimensional vectors \( y_{r_T+1}^*, \ldots, y_n^* \) are independent identically distributed truncated multivariate normal variables, where the domain for the first \( T - 1 \) observations is \( (-\infty, \infty) \) and that for the \( T \)th observation is \( (M, \infty) \).

4.3 Likelihood and estimation

Using Little and Rubin's notation for nonignorable missing, the full likelihood can be written as the joint distribution of the vector valued random variable \( Y_{obs} = \left( y_1^*, y_2^*, \ldots, y_{r_T}^*, y_{r_T+1}^*, \ldots, y_n^* \right) \) and missing indicator vector \( I = \left( I_1, I_2, \ldots, I_{r_T+1}, \ldots, I_n \right) \).
\[
\left( I_1, I_2, \ldots, I_{R_T}, I_{R_T+1}, \ldots, I_n \right), \text{ where } I_j = 1 \text{ of the observation is missing and 0 otherwise.}
\]

The joint distribution of \( Y_{\text{obs}} \) and \( I \) is

\[
p(Y_{\text{obs}}, I | \theta, \xi) = p(y_1, y_2, \ldots, y_{R_T}, y_{R_T+1}, \ldots, y_n, I_1, I_2, \ldots, I_{R_T}, I_{R_T+1}, \ldots, I_n | \theta, \xi). \quad (4.2)
\]

Let \( R_T = \sum_{j=1}^{n} I_j \) denote the random variable representing the number of individuals MDT at time \( T \) so that \( R_T \) has binomial distribution and is a sufficient statistics for \( I_j \). Then equation (4.2) can be rewritten as the joint distribution of \( y_1, y_2, \ldots, y_{R_T}, y_{R_T+1}, \ldots, y_n \) and \( R_T \). To derive the likelihood one could write the joint probability distribution function as a product of conditional distributions as follows

\[
p(Y_{\text{obs}}, I | \theta, \xi) = p(y_1, y_2, \ldots, y_{R_T}, y_{R_T+1}, \ldots, y_n, R_T | \theta, \xi)
\]

\[
= p(y_{r_T+1}, y_{r_T+2}, \ldots, y_{nT} | y_1, y_2, \ldots, y_{R_T}, y_{R_T+1}, \ldots, y_n, R_T = r_T, \theta)
\]

\[
\times p(y_1, y_2, \ldots, y_{r_T}, y_{r_T+1}, \ldots, y_n | R_T = r_T, \theta)P(R_T = r_T | \xi),
\]

\[(4.3)\]

where \( y_{r_T+1}, y_{r_T+2}, \ldots, y_{nT} \) represent the \( T \)th observations of the random vectors \( y_{r_T+1}, \ldots, y_{nT} \).

\( y_n \). Once the \( T \)th observations are dropped, the random vectors \( y_{r_T+1}, \ldots, y_n \) are \( T-1 \) dimensional and their distributions are identical to the random vectors \( y_1, y_2, \ldots, y_{r_T} \).
Also, the two sets of random vectors \( y_{r_T+1,T}, \ldots, y_{n,T} \) and \( y_1, y_2, \ldots, y_{r_T} \) are independent. Since, the individuals at any given time point \( t \) are independently distributed and since the joint distribution of \( y_1, y_2, \ldots, y_n \) is independent of \( R_T \), the above joint distribution in equation (4.3) could be further simplified as,

\[
p(Y_{obs}, 1 | \theta, \xi) = p(y_{r_T+1,T}, \ldots, y_{n,T} | y_{r_T+1, \ldots, y_{n,T}}, R_T = r_T, \theta) \times p(y_1, y_2, \ldots, y_n | \theta) P(R_T = r_T | \xi) = \prod_{i=1}^{n-r_T} p(y_{r_T+i}^* | y_{r_T+i}, \theta) \prod_{i=1}^n p(y_i | \theta) P(R_T = r_T | \xi).
\]

Having the multivariate truncated normality and the distribution of \( R_T \) is binomial, the likelihood function reduces to equation (2.8) in Ramakrishnan and Wang (2005). That is,

\[
L(\mu_T, \sigma_T^2, \mu, \Sigma, \Sigma_T) \propto \frac{1}{\sigma_T | i} \left( \prod_{i=1}^{n-r_T} \left( 1 - \Phi \left( \frac{M - \mu_T | i}{\sigma_T | i} \right) \right) \times \exp \left( \frac{1}{2\sigma_T^2 | i} \sum_{i=1}^{n-r_T} \left( y_{r_T+i}^* - \mu_T | i \right)^2 \right) \times \frac{1}{|\Sigma|^{n/2}} \exp \left( \frac{1}{2} \sum_{i=1}^n \left( y_i - \mu \right)^T \Sigma^{-1} \left( y_i - \mu \right) \right) \times \left( n \atop r_T \right) \left( \Phi \left( \frac{M - \mu_T}{\sigma_T} \right) \right)^{r_T} \left[ 1 - \Phi \left( \frac{M - \mu_T}{\sigma_T} \right) \right]^{n-r_T}, \right.
\]

where \( \theta = (\mu_T, \sigma_T^2, \mu, \Sigma, \Sigma_T) \) and \( \xi = (\mu_T, \sigma_T^2, M) \).
and

\[ \mu_{T|i} = \mu_T + \Sigma_T \Sigma^{-1}(y_{r_T+i} - \mu), \]

\[ \sigma^2_{T|i} = \sigma_T^2 - \Sigma_T \Sigma^{-1}\Sigma_T'. \]

It can be shown that the mean and variance of the conditional truncated random variable 

\[ y_{r_T+i} \]

are

\[ \mu^*_{T|i} = \mu_{T|i} + \sigma_{T|i}\zeta_i(M), \]

\[ \sigma^*_{T|i} = \sigma_{T|i} \left[ 1 + \zeta_i(M) \left( \frac{M - \mu_{T|i}}{\sigma_{T|i}} - \zeta_i(M) \right) \right], \]

where

\[ \zeta_i(M) = \frac{\phi \left( \frac{M - \mu_{T|i}}{\sigma_{T|i}} \right)}{1 - \Phi \left( \frac{M - \mu_{T|i}}{\sigma_{T|i}} \right)} \]

The EM algorithm could simplify the estimation procedure considerably. The EM algorithm is a general method of finding the maximum-likelihood estimate of the parameters from a given data set when the data are incomplete or have missing values. In the case of MDT, the observed data are the data in equation (4.1), denoted by \( S(y) \). It
constitutes the incomplete data. The complete data would be obtained by adding
\[ y^* = (y_1^*, T, y_2^*, T, \ldots, y_r^*, T). \]
The joint density function of \( S(y) \) and \( y^* \) is
\[ p(S(y), y^* | \theta, \xi) = p(y^* | S(y), \theta, \xi) p(S(y) | \theta, \xi). \]

In the EM algorithm the expected value of the complete data log likelihood
\[ p(S(y), y^* | \theta, \xi) \] with respect to the unknown data \( y^* \) given the observed data \( S(y) \) and the current parameter estimates \( \theta^{k-1}, \xi^{k-1} \) is first obtained. That is,
\[ Q(\theta, \xi; \theta^{k-1}, \xi^{k-1}) = E \left[ \log p(S(y), y^* | \theta, \xi) | S(y) \theta^{k-1}, \xi^{k-1} \right], \]
where \( \theta^{k-1}, \xi^{k-1} \) are the current parameter estimates that is used to evaluate the missing \( y^* \). The \( \theta \) and \( \xi \) are the new parameters that ultimately will be optimized in an attempt to maximize the likelihood. The evaluation of this expectation is the E-step of the EM algorithm.

The second step (the M-step) of the EM algorithm is to maximize the likelihood of the parameters given the “complete” data obtained in the E-step. In general, for the \( k \)th iteration this may be written as,
\[ (\theta^k, \xi^k) = \max_{\theta, \xi} Q(\theta, \xi; \theta^{k-1}, \xi^{k-1}) \]
These two steps are repeated until convergence. Each iteration is guaranteed to increase the log likelihood and the algorithm is guaranteed to converge to a local maximum of the likelihood function (Dempster, Laird and Rubin, 1977).

Initial estimates for the mean and the variance-covariance parameters at time \( T \) could be based on repeated measures model treating MDT as MAR.
Given initial values of the various parameters, the expectation step to estimate the MDT observations is achieved using the conditional truncated normal distribution as discussed in Section 2.3 and in Ramakrishnan and Wang's article (2005).

Once the MDT observations are estimated from the E-step, the M-step could be easily applied because the multivariate normal theory would apply and thus explicit expressions for the means, variances and covariances would exist.

4.4 Other methods useful for longitudinal missing data

There are several widely acceptable longitudinal methods that are useful for dealing with the MDT situation. These methods are generally imputation methods, while MDT method basically estimates the model parameters under a better fitting distribution assumption. However, the imputed values may be obtained from the last E-step in the iteration procedure.

The goal of any imputation technique is to produce a complete data set, which can then be analyzed using complete-data inferential methods. Some methods are based on a single imputation such as last observation carried forward (LOCF), individual regression prediction, and a repeated measures mixed model. Although it may not be explicit, all the imputation methods assume a model for the missing data. Wang (1995) compared these methods with MDT applying MANOVA to the complete data. The MDT approach in his dissertation did not utilize the EM algorithm but instead estimated the missing observations using the moment estimators of the means, variances and covariances. In the next three sections the three methods compared in this dissertation with the MDT method are briefly described.
4.4.1 Last observation carried forward method (LOCF)

This method assigns the person’s last known observation prior to drop out to the missing value. In other words, for a subject dropping out at time \( t \), the missing value is imputed by his/her observed value of that variable at time \( t-1 \). The underlying assumption is that the observations at later time points will not change after the subject drops out. That is,

\[
p(y_{i,t'}^* = y_{i,t}) = 1,
\]

where \( t' \) denotes the time the subject was last observed and \( j = 1, ..., T - t' \).

4.4.2 Individual regression prediction method (REG)

This method fits a regression line between the outcome variable and time for each subject with missing value by estimating the conditional expectation, \( E(y_{i,t'+j}^* \mid y_{i1}, ..., y_{it}) \).

That is, \( \hat{y}_{it} = \hat{\beta}_{i0} + \hat{\beta}_{i1}t \), where \( t = 1, ..., t' \) and \((\hat{\beta}_{i0}, \hat{\beta}_{i1})\) is a least-square estimate of the regression parameter vector for subject \( i \). This method extrapolates the missing observations based on the regression fit. That is, the missing observation is essentially estimated by \( \hat{y}_{i,t'+j}^* = \hat{\beta}_{i0} + \hat{\beta}_{i1}(t' + j) \). The parameters \( \hat{\beta}_{i0} \) and \( \hat{\beta}_{i1} \) vary over different individuals. In other words, the conditional expectations \( E(y_{i,t'+j}^* \mid y_{i1}, ..., y_{it}) \) and \( E(y_{i',t'+j}^* \mid y_{i'1}, ..., y_{i't}) \) are independent, for every \( i, i' \) and \( j \), where \( i \neq i' \), \( i, i' = 1, 2, ..., n \) and \( j = 1, 2, ..., T - t' \).
4.4.3 Repeated measures mixed model method (MIXED)

The repeated measures mixed model treats the missing data as MAR. In the repeated measures model method, time $t$ was treated as a fixed effect and a common covariance structure between time points on the same subjects was specified. Due to the specification of covariance structure, observations at each time point could influence parameter estimates at every other time point. That is, the information from incomplete individuals whose observations are limited to early time points will be taken into account when estimating parameters at later time points. This is the consequence of treating the observations as MAR. Using the notation introduced in Chapter 1, this implies for subjects $i$ and $i'$, $E(y_{i,\text{mis}}) = E(y_{i',\text{mis}})$ if $y_{i,\text{obs}}$, $X_i$, and $y_{i',\text{obs}}$, $X_{i'}$ are identical. Further, since the repeated measures mixed model includes the information from the incomplete data implicitly as in the MDT case, the imputation of the missing data is not explicit.

4.5 Design of the simulation study

Three sets of simulations were performed in terms of the missing data mechanism and the form of response function. One set of data was simulated under MDT with linear response function, one set simulated under MDT with quadratic response function, and the third set simulated under MAR with linear response function. The purpose of the first two simulations was to compare the parameter estimates under two different types of response functions when the missing data are MDT. The third is to study the sensitivity to the misspecification of the missing data mechanism if the MDT method is applied to the data that are MAR. The four methods (LOCF, REG, MIXED and MDT) were compared in terms of the square of bias and mean square error of the parameter estimates.
4.5.1. Factors in the simulation study

The simulation parameters were chosen based on the factors that could have an influence on the estimation of the parameters of interest. Also, the magnitude of the mean, variance and covariance were chosen using the results from the IMPS data to reproduce the data sets that are close to known situations. The simulation number was chosen to be 100.

i) Sample size and dropout rate

Three sample sizes, 50, 100 and 200 were considered to study whether the different methods perform similarly for large samples. This also provides an opportunity to study the asymptotic properties of the estimates from the MDT method. The four time points used to simulate the data were \( t = 0, 1, 3 \) and 6 (to mimic the time points in the IMPS data). To represent the real situations adequately, drop out rates at the final time point \( t = 6 \), was chosen to be 5%, 10% and 20% respectively. If dropout occurs at time \( t = 3 \) as well as at time \( t = 6 \), the drop out rates at time \( t = 3 \) were selected to be 3%, 7% and 15% respectively and 5%, 15% and 20% at time \( t = 6 \) correspondingly.

ii) Variance-covariance matrix and correlation

For the within-subject variance-covariance matrix, a first-order autoregressive structure was used. The AR(1) correlation \( \rho \) was set at three different levels, namely 0.2, 0.4 and 0.8. Variance \( \sigma^2 \) at all time points was set at 2 (again similar to the IMPS data).

iii) The form of response function

All the imputation methods are likely to be influenced by the form of response function. Therefore, data were simulated under
a) Linear function: \( y = \beta_0 + \beta_1 t \), where \( \beta_0 \) is the baseline \((t = 0)\) measure and \( \beta_1 \) is the slope of the function.

b) Concave function: \( y = \beta_0 + \beta_1 t + \beta_2 t^2 \), where \( \beta_0 \) is the baseline \((t = 0)\) measure, \( \beta_1 \) is the linear slope, and \( \beta_2 \) is the measure of quadratic component with \( \beta_2 < 0 \).

c) Convex function: \( y = \beta_0 + \beta_1 t + \beta_2 t^2 \), where \( \beta_0 \), \( \beta_1 \) and \( \beta_2 \) are defined similarly as in b) but \( \beta_2 > 0 \).

The model coefficients were chosen so that the mean of the observation at the baseline is 5 \((\mu_0 = \beta_0 = 5)\), and the mean of the observations at the last time point \((t = 6)\) is \( \mu_4 = 2.6 \). In the linear case this turns out to be \( y = 5 - 0.4t \). The quadratic parameter \( \beta_2 \) (concave and convex) was set in such a way that at every time point, the concave and the convex functions are symmetric around the linear function (Figure 4.1).

![Figure 4.1 The three response functions simulated](image-url)
4.5.2. Generation of multivariate normal random samples with missing data

A random sample of \( y_i \) was generated using the SAS (SAS 9.1, 2002) multivariate normal random number generating function VNORMAL with the mean and variance-covariance matrix specified using the conditions in 4.5.1. The usage of the call function is

\[
\text{CALL VNORMAL } ((Y, \mu, \Sigma, N, seed));
\]

The \( \mu \) specifies a \( T \times 1 \) mean vector, where \( T = 4 \) is the number of time points. The \( \Sigma \) specifies a \( T \times T \) symmetric positive-definite covariance matrix. It has AR(1) structure with \( \sigma^2 = 2 \) and \( \rho = 0.2, 0.4 \) or 0.8. The \( N \) specifies the length of the series, namely the sample sizes (50, 100 and 200). The \( seed \) specifies the random number seed. The VNORMAL returns a multivariate normal random series \( Y \), which is an \( N \times T \) matrix that contains the generated normal random variables with mean \( \mu \) and covariance matrix \( \Sigma \).

The \( i \)th row of \( Y \) represents the observations from \( i \)th subject and the \( j \)th column represents the observations from \( j \)th time point.

Two kinds of missing data mechanism were generated, namely MDT and MAR.

i) Missing due to Truncation

Based on a given dropout rate at the last time point (denoted as \( \frac{r_A}{N} \)), assume the MDT occurs at the lower tail of the distribution, the threshold \( M \) is obtained using

\[
p(y_{i4} < M) = \frac{r_A}{N}, \tag{4.5}
\]
where $y_{i4} \sim N(\mu_4, \sigma)$. The part of the sample for which $y_{i4} < M$ were considered as MDT.

The missing data are generated assuming the monotone missing pattern. This is reasonable because one of the assumptions for the MDT method is that once the patient passes the dropout threshold, he/she will not come back; therefore the MDT observation is monotonically missing. When the MDT occurs at the fourth time point as well as at the third time point, the same $M$ is used to create the truncated distribution at the third time point. The mean $(\mu_3)$ at the third time point was recalculated so that the required drop out rates at the third point can be achieved.

ii) Missing at Random

The MAR data was generated by the code:

```sas
DO I=1 TO N;
    UNI[I, 1]=RANUNI(SEED);
    IF UNI[I, 1] > DROPOUT THEN
        Y[I, 4]= Y[I, 4];
    ELSE Y[I, 4]= .;
END;
```

Here, RANUNI is a SAS function which returns a random variate from a uniform (0, 1) distribution with seed controlled by the random number SEED. The DROPOUT is dropout rate at the last time point, which may be 0.05, 0.1, or 0.2 and $Y$ is a $N \times 4$ multivariate normal variable generated by the VNORMAL Call function.

4.5.3 Comparison of measures

The comparison of the four methods (MDT, LOCF, REG and MIXED) was based on the parameter estimates of interest, namely the mean estimates at missing time points, the correlation and variance estimates in AR(1) structure. The measures for comparison are listed below.
The average of mean estimates across the simulations for time = 6,

\[ \bar{\mu}_4 = \frac{1}{100} \sum_{i=1}^{100} \hat{\mu}_{4i}, \]

where \( \hat{\mu}_{4i} \) is the mean estimates at time = 6 from the model fitted in the \( i \)th simulation \( (i=1, 2, \ldots, 100) \).

For the linear response the MDT was allowed to occur at time = 6 as well as time = 3. The average of mean estimates at time = 3 is computed similarly. That is,

\[ \bar{\mu}_3 = \frac{1}{100} \sum_{i=1}^{100} \hat{\mu}_{3i}, \]

where \( \hat{\mu}_{3i}, i=1, 2, \ldots, 100, \) is mean estimates at time = 3. Similarly, the averages of the correlation and variance estimates were calculated.

Square of the biases and mean square errors (MSEs) of these quantities were also computed. The biases of the mean estimates were estimated as follows

\[ \text{Bias}(\hat{\mu}_4) = \bar{\mu}_4 - 2.6, \]
\[ \text{Bias}(\hat{\mu}_3) = \bar{\mu}_3 - \mu_3, \]

where \( \mu_3 = 2.934, 2.875 \) or 2.876. The biases of the variance and correlation estimates were estimated as

\[ \text{Bias}(\hat{\sigma}_i^2) = \bar{\sigma}^2 - 2, \]
\[ \text{Bias}(\hat{\rho}_i) = \bar{\rho} - \rho, \]

where \( \rho = 0.2, 0.4 \) or 0.8.

Mean square errors (MSEs) of the estimates were estimated as
When the bias is zero, these quantities are same as the variance of the mean estimates,

which will be around \( \frac{\sigma^2}{n} \), where \( \sigma^2 = 2 \) and \( n = 50, 100 \) or 200. The MSEs of the variance and correlation estimates were estimated as

\[
\text{MSE}(\hat{\mu}_j) = \frac{1}{100} \sum_{i=1}^{100} (\hat{\mu}_{ji} - \mu_j)^2.
\]

\[
\text{MSE}(\hat{\sigma}_j) = \frac{1}{100} \sum_{i=1}^{100} (\hat{\sigma}_i^2 - 2.0)^2.
\]

\[
\text{MSE}(\hat{\rho}) = \frac{1}{100} \sum_{i=1}^{100} (\hat{\rho}_i - \rho)^2.
\]

### 4.6 Simulation results

The results from the simulations are presented as plots and tables (Appendix B). The plots of MSE vs. square of bias of the parameter estimates are presented in Figures B.1 - B.15. The averages and standard errors of the parameter estimates are presented in Tables B.1 - B.12. The simulation results could be summarized as follows.

- When missing proportion is small all the methods perform reasonably well.
- Regression method estimates the means accurately for linear response function, but typically over estimates the variance and correlation especially when the correlation is low.
- The LOCF and regression are both sensitive to the form of response function.
• When the missing are not MAR, the estimates from MIXED method have large biases in most situation.

• When the data are missing due to truncation, MDT method performs best for all the parameters regardless of missing proportion and the forms of response function.

Linear response function was studied with MDT occurring at the last one or two time points and with different missing proportions (Table B.1-B.7 and Figure B.1-B.9). In this case the MDT performs best for all the parameters regardless of missing proportion. Regression prediction method performs well in terms of the mean estimates (Table B.1, B.4 & B.5, Figure B.1, B.4 & B.5). However, it overestimates the variance and correlation especially when the correlation is low (\(\rho = 0.2\) or 0.4) (Table B.2 - B.3, B.6 - B.7, Figure B.2 - B.3, B6 - B.7). For example, when correlation = 0.2, \(n = 100\) with 15% MDT at time = 3 and 20% MDT at time = 6, the mean estimates of \(\sigma^2\) is 7.548, which is about 2 times larger than the true value 2 (Tables B.6; Figures B.6).

Regression prediction and LOCF are sensitive to the form of the response function (Tables B.8-B.10, Figures B.10 – B.12). Regression prediction performs best when the data are generated by a linear model. LOCF performs the best when the data are generated by a convex model. The MDT method seems robust to the form of response function (Tables B.8-B.10, Figures B.10 – B.12).

From Table B.11- B.13 and Figures B.13 – B.15, for the cases where the missing values were simulated under MAR all the methods perform reasonably well when the missing proportion is small (5%). Increasing the missing to 10% or 20%, means of parameter estimates from MDT method are the most sensitive as compared to all the other
methods (Table B.11, Figure B.13). Also the MIXED performs the best as expected. Regression is as good as MIXED in terms of the mean estimates (Table B.11, Figure B.13) while it tends to overestimate the variance and correlation especially when the correlation is low ($\rho = 0.2$ or $0.4$) (Table B.12-B.13, Figure B.14-B.15).

4.7 Discussion

The comparison of the MDT method with other relevant methods was studied via simulation. In general when the missing proportion is small, the results show all the methods perform reasonably well suggesting the choice of a method for handling the missing data is not crucial in this case.

Although the regression method estimates the means accurately, it typically overestimates the variance and estimates the correlation with a large bias, especially when missing proportion is high. The reason for large bias in variability estimates is that the regression method utilizes the observations from only the missing subjects and not the group level information. Therefore, the regression method will more often fail to reject the null hypotheses on the fixed effects.

As expected the LOCF and Regression are both sensitive to the form of response function. The use of either of these methods is inappropriate when the large number of non-ignorable missing occurs and identifying the form of the response function is part of the analysis. The bias for the estimation for the MIXED method is large for most cases when the missing are not MAR.

When the data are missing due to truncation, the MDT method performs best for all the parameters regardless of missing proportion and the form of response function.
The robustness of the MDT method against the form of the response function is an
classification advantage since one of the primary interests in data analysis is identifying the form of
response function. If the imputation methods are sensitive to this form, it is likely that the
model fitted using the complete data will essentially reproduce the model used to impute
the data.

In practice, the choice of the method for dealing with the missing data is important especially when large proportion is missing. The MDT method should be used if the form
of the model is unknown and there is reason to believe the assumption of truncated normal
distribution is appropriate. Application of the other methods that do not assume truncated
normal distribution lead to unsatisfactory results. When the missing mechanism is
unknown, the application of MDT method is not recommended.
Chapter 5

MDT Method in conjunction with Multiple Imputation

5.1 Introduction

Multiple imputation is a technique first developed by Rubin (1977, 1978) to handle missing data in a variety of experiments and for a variety of missing data patterns. The technique essentially replaces each missing value with two or more acceptable values so that the uncertainty about the right value to impute could be measured and incorporated into the analysis. This technique also ensures the consistency and convergence properties of the estimators of interest (Little & Rubin, 2002).

The multiple imputation technique, briefly, is as follows: Create \( m (m \geq 2) \) complete data sets by replacing each missing value with \( m \) repeated random draws from a predictive distribution of the missing data. Analyze each of the \( m \) complete data sets using standard complete data procedures. Combine \( m \) sets of the point and variance estimates by 'Rubin's rule' (1987, pp76) (described later in section 5.3) to make valid inferences. Irrespective of which complete data analysis is used, the process of combining the point and variance estimates is essentially the same.

There are a few statistical packages available to implement the multiple imputation method, most of which are for ignorable missing. For example, SAS PROC MI procedure creates complete data sets for incomplete multivariate normal data. Another procedure, PROC MIANALYZE, is then used in conjunction with PROC MI to generate valid
statistical inferences about parameters by combining the predictive distribution from the $m$
complete data sets. Both PROC MI and PROC MIANALYZE assume ignorable missing
data mechanism. That is, the missing data are missing at random (MAR) and the parameter
$\theta$ of the data model and the parameter $\xi$ of the missing data indicator model are distinct
(equation (1.1)) in the sense that from a frequentist perspective, the joint parameter space
of $\theta$ and $\xi$ is the product of the parameter spaces of $\theta$ and $\xi$ and from a Bayesian
perspective, the priors of $\theta$ and $\xi$ has the form $p(\theta, \xi) = p(\theta)p(\xi)$.

For nonignorable missing situation, imputation assuming an ignorable response
mechanism will fail to correct the bias due to nonresponse adequately. Under ignorable
nonresponse, the conditioning on, whether $Y$ is missing or not, is irrelevant to estimation
of the posterior distribution of $Y$. Suppose there are two subjects with identical
covariates $X$, one of the subjects has a missing value at time $t$. In the case of ignorable
missing, the conditional distribution $f(Y | X, Y$ is missing) equals the conditional
distribution $f(Y | X, Y$ is not missing). However, if this is nonignorable,
$f(Y | X, Y$ is missing) $\neq f(Y | X, Y$ is not missing), and missing values could not be
imputed by the values of the other subjects whose $Y$ is observed with identical covariate $X$
at time $t$.

5.2 Review of methods for continuous repeated measures with nonignorable dropout

There are a few methods useful for dealing with continuous repeated measures with
nonignorable dropouts. Generally, these methods can be categorized into three types:
systematic difference from ignorable imputations, semi-parametric method and likelihood-based method.

Let \( \mathbf{y} \) be an \( N \times T \) data matrix representing \( N \) subjects measured at \( T \) time points. The \( \mathbf{y} = (y_1, \ldots, y_n)^T \), where \( y_i = (y_{i1}, \ldots, y_{iT})^T \) represents the set of repeated measures from subject \( i \) and a random sample from a \( T \) dimensional multivariate probability distribution \( f(\mathbf{y} | \theta) \) governed by parameter \( \theta \). If \( \mathbf{y} \) is not fully observed, following Little & Rubin’s (2002, pp12) notation, denote the observed portion of \( \mathbf{y} \) by \( \mathbf{y}_{\text{obs}} \) and the missing portion by \( \mathbf{y}_{\text{mis}} \). Let \( \mathbf{X} \) denote the fixed covariates such as treatment arms, gender, age or time points and let \( \mathbf{I} \) denote the \( N \times T \) missing indicator vector with \( I_i \) being the missing indicator variable for subject \( i \). The \( \mathbf{I} \) is subject to a probability distribution \( f(\mathbf{I} | \xi, \mathbf{y}) \) governed by parameter \( \xi \).

For the \( i \)th subject, denote the observed portion by \( \mathbf{y}_{i,\text{obs}} \), the missing portion by \( \mathbf{y}_{i,\text{mis}} \), the fixed covariates by \( \mathbf{X}_i \) and the missing indicator by \( I_{i,} \), where

\[
I_{i,:} = \begin{cases} 
1 & \text{if } y_{i,j} \text{ is missing} \\
0 & \text{if } y_{i,j} \text{ is observed} 
\end{cases}
\]

5.2.1 Systematic difference from ignorable imputations

i) Impute nonignorable \( \mathbf{y}_{i,\text{mis}} \) by a fixed transformation of ignorable imputed \( \mathbf{y}_{i,\text{mis}} \). For example, (nonignorable imputed \( \mathbf{y}_{i,\text{mis}} \)) \( = a \times (\text{ignorable imputed } \mathbf{y}_{i,\text{mis}}) \) or (nonignorable imputed \( \mathbf{y}_{i,\text{mis}} \)) \( = \exp \left[ a + b \times \log(\text{ignorable imputed } \mathbf{y}_{i,\text{mis}}) \right] \), where \( a \) and \( b \) are constants. These methods change the location, scale and shape of the ignorable
imputed values. The advantages of such transformation are easy to implement and to describe to non-statisticians.

ii) This is an extension of the above method. Only a certain percentage of the imputed data from ignorable imputation might be distorted by a fixed transformation. This method might be appropriate when there is a suspicion that missing data is from varying reasons, where only some of the nonrespondents are nonignorable and others may be ignorable.

5.2.2 Semiparametric nonresponse model

Rotnitzky et al. (1998) extend the generalized estimating equation (GEE) approach by proposing a class of augmented inverse probability of response weighted estimator to allow for nonignorable nonresponse in longitudinal studies. The proposed estimators don't require full specification of a parametric likelihood and their computation doesn't require numerical integration. They attempted to minimize the parametric assumptions by making limited use of the covariate information. This approach results in increased sensitivity of inference to the nonignorable component of the model and possibly leads to overly conservative inferences (Little & Rubin, 1999).

5.2.3 Likelihood-based methods

A number of model-based methods have been proposed for nonignorable dropout in longitudinal data analysis (Diggle & Kenward, 1994; Little 1993, 1994; Little and Wang, 1996; Wu & Bailey 1989; Schluchter, 1992). When data are incomplete, the distribution of the data is the joint probability model $f(Y_{\text{obs}}, Y_{\text{mis}}, I | X, \theta, \xi)$. Since $Y_{\text{mis}}$ is unknown, the likelihood function of $L(\theta, \xi | X, Y_{\text{obs}}, Y_{\text{mis}}, I)$ can not be evaluated. The
observed data likelihood of \( \theta \) and \( \xi \) are evaluated instead. The distribution of observed data is obtained by integrating \( Y_{mis} \) out of the joint density of \( Y = (Y_{obs}, Y_{mis}) \) and \( I \). That is,

\[
f(Y_{obs}, I \mid X, \theta, \xi) = \int f(Y_{obs}, Y_{mis}, I \mid X, \theta, \xi) \, dY_{mis}
\]

\[
= \int f(Y_{obs}, Y_{mis} \mid X, \theta) f(I \mid X, Y_{obs}, Y_{mis}, \xi) \, dY_{mis}
\]  

(5.1)

The observed data likelihood of \( \theta \) and \( \xi \) is any function of \( \theta \) and \( \xi \) proportional to (5.1)

\[
L(\theta, \xi \mid X, Y_{obs}, I) \propto f(Y_{obs}, I \mid X, \theta, \xi)
\]  

(5.2)

Under nonignorable missing assumption, ML estimation of \( \theta \) requires models for the missing data mechanism (i.e. \( f(I \mid X, Y_{obs}, Y_{mis}, \xi) \)) and maximization of the full likelihood as opposed to the ignorable missing assumption in which only the likelihood of the observed data distribution needs to be maximized as discussed in section 1.3. The ML estimates could be obtained through an iterative procedure such as Newton-Raphson or the EM algorithm. A large sample covariance matrix for the parameters \( \theta \) can be estimated either directly by the information matrix obtained by differentiating the log likelihood twice with respect to \( \theta \) and \( \xi \) if closed form expressions are available or by bootstrap sampling.

Little and Rubin (2002, pp 313) distinguished two basic approaches based on the full likelihood to model the nonignorable nonresponse: selection model approach and pattern-mixture model approach.

i) Selection model approach

Under the assumption that the subjects are modeled as independent,
\[ f(y, I \mid X, \theta, \xi) = \prod_{i=1}^{n} f(y_i, I_i \mid X_i, \theta, \xi). \]

Specify \( f(y_i, I_i \mid X_i, \theta, \xi) = f(y_i \mid X_i, \theta) f(I_i \mid X_i, y_i, \xi), \) where the first factor characterizes the distribution of the population data defined by \( \theta \) and the second factor models the distribution of response mechanism characterized by \( \xi \). This approach has been called the selection modeling approach because of the specification for the response mechanism that selects subjects to be respondents.

The random coefficient selection model is proposed for the analysis of repeated measures data with nonignorable missing (Little, 1995). It specifies random coefficients \( \beta_i \) that vary across the subjects. The complete data likelihood for subject \( i \) is based on a model for joint distribution of \( y_i, I_i \) and \( \beta_i \) conditioning on covariates \( X_i \) and fixed parameters and may be factored as:

\[
\begin{align*}
f(y_i, I_i, \beta_i \mid X_i, \theta, \xi, \phi) &= f(y_i \mid X_i, \beta_i, \theta) f(\beta_i \mid X_i, \phi) f(I_i \mid X_i, y_i, \beta_i, \xi) \\
&= f(y_i \mid X_i, \beta_i, \theta) f(\beta_i \mid X_i, \phi) f(I_i \mid X_i, y_i, \beta_i, \xi, \phi, \xi). \quad (5.3)
\end{align*}
\]

where \( \beta_i \) is subject to a probability distribution \( f(\beta_i \mid X_i, \phi) \) governed by parameter \( \phi \).

The first two factors define the joint distribution of \( y_i \) and \( \beta_i \), representing the complete likelihood if there is no missing data. The third factor models the probability of missing at a particular time as a function of \( X_i, y_i \) and random effect \( \beta_i \).

Little (1995) distinguished the random coefficient selection model for repeated measures data by nonignorable outcome-based dropout and nonignorable random-coefficient-based dropout.

For nonignorable outcome-based dropout, the last expression in (5.3) is
Diggle and Kenward (1994) used the term “informative drop-out” for this mechanism where the dropout depends on the current and previous values of $Y$.

For nonignorable random-coefficient-based dropout, the last expression in (5.3) is

$$f(I_i | X_i, Y_{i,obs}, Y_{i,mis}, \beta_i, \xi) = f(I_i | X_i, Y_{i,obs}, Y_{i,mis})$$

where missing depends on underlying random coefficients $\beta_i$. For example, one of the random coefficients may represent a slope. Random-coefficient-based dropout indicates that dropout depends on this underlying, unobserved slope. In other words, the dropout depends on past, current and future values of $y_i$.

ii) Pattern-mixture model approach

The joint distribution of $y_i$ and $I_i$ can be factored as,

$$f(y_i, I_i | X_i, \theta, \xi) = f(y_i | X_i, I_i, \theta) f(I_i | X_i, \xi),$$

where the first factor specifies the distribution of population data $y_i$ in the strata defined by conditioning on different patterns of missing data $I_i$. The second factor models the distribution of missing data pattern $I_i$ parameterized by $\xi$. The term “pattern-mixture” reflects the facts that the resulting marginal distribution of complete data is a mixture of respondents and nonrespondents stratified over the missing patterns.

The full likelihood for repeated measures with nonignorable missing by random coefficient pattern-mixture models is

$$f(y_i, I_i, \beta_i | X_i, \theta, \xi, \phi) = f(y_i | X_i, \beta_i, I_i, \theta) f(\beta_i | X_i, I_i, \phi) f(I_i | X_i, \xi)$$  

(5.4)
where the first term specifies the distribution of $y_i$ in the strata defined by conditioning on $\beta_i$ and $I_i$. The second term is the distribution of the random effects conditioned on $I_i$.

Similar to selection model, Little (1995) distinguished the random coefficient pattern-mixture models into a) outcome-dependent dropout model, where random parameter $\beta_i$ has the same distribution across the dropout pattern. Here the equation (5.4) can be rewritten as

$$f(y_i, I_i, \beta_i | X_i, \theta, \xi, \phi) = f(y_i | X_i, \beta_i, I_i, \theta) f(\beta_i | X_i, \phi) f(I_i | X_i, \xi).$$

And b) random-coefficient dependent dropout, where the dropout depends on $X_i$ and the random coefficient $\beta_i$, here the equation (5.4) can be rewritten as

$$f(y_i, I_i, \beta_i | X_i, \theta, \xi, \phi) = f(y_i | X_i, \beta_i, \theta) f(\beta_i | X_i, I_i, \phi) f(I_i | X_i, \xi).$$

iii) Missing Due to Truncation (MDT) approach

This method was described in chapter 2 and Ramakrishnan and Wang's paper (2005). It can be rewritten in the above pattern-mixture framework to apply the multiple imputation technique. The equation (5.4) in MDT situation may be written,

$$f(y_i, I_i, \beta_i | X_i, \theta, \xi, \phi) = f(y_i | X_i, \beta_i, I_i, \theta) f(\beta_i | X_i, I_i, \phi) f(I_i | X_i, \xi) \quad (5.5)$$

Notice that the main difference between equation (5.5) and equation (5.3) is that the conditional distributions of $f(y_i | X_i, \beta_i, I_i, \theta)$ and $f(\beta_i | X_i, I_i, \phi)$ don’t depend on $I_i$ in equation (5.3). Comparing (5.5) and (5.4), the MDT and the pattern-mixture model seem identical, and therefore the MDT follows the framework of pattern-mixture model. However, there exists a fundamental difference in modeling the conditional
distribution \( f(y_i \mid X_i, \beta_i, I_i, \theta) \). Pattern-mixture, in general, models the distribution \( f(y_i \mid X_i, \beta_i, I_i, \theta) \) for both \( I_y = 0 \) and \( 1 \) using the same family of distribution (say, multivariate normal distribution), although the mean and variance are allowed to change between \( I_y = 0 \) and \( 1 \). In the MDT case, the conditional distributions of \( y_i \) given \( I_y \) are allowed to come from a different family of distribution. This is not explicitly or implicitly described under pattern-mixture model. In the MDT case, if \( I_y = 0 \), it is assumed that the observation follows a multivariate normal distribution, if \( I_y = 1 \), it follows a truncated multivariate normal distribution.

In Chapter 2, a simplified form of this equation (Equation 2.7) where the method is used for a general mean \( \mu(\theta) \) was presented. The covariates \( X_i \) and the random effect \( \beta_i \) are inherent in \( \mu(\theta) \). The missing pattern under MDT is known and it is monotone. Therefore the joint distribution of the \( I_i \)'s only depends on the distribution of the number of missing (Bernoulli vs. binomial). That is, a binomial model will be adequate to define \( f(I_i \mid X_i, \xi) \).

### 5.3 MDT method in conjunction with Multiple Imputation (MI)

Although majority of MI procedures involve the use of ignorable missing models, MI can also be used with nonignorable missing data (Rubin, 2003). The real issue with the use of nonignorable models in MI is that without external information, the modeling assumption of the missing mechanism is rarely justifiable at the time of imputation. However, when external information is available, MI could be applied. For example, MI
A multiple imputation method for MDT data is proposed in this section. In brief, the MDT method in combination with bootstrap sampling is used to generate multiple imputed data sets. Then SAS PROC MIANALYZE is applied to combine the inference results across the $m$ imputations as in the MAR case. The justification for the approach follows from Little & Rubin (2002, pp. 216) and is described below.

One of the methods Little & Rubin (2002, pp216) suggested is based on the asymptotic distribution of ML estimates. For the multiple imputation, if the ML estimate $\hat{\theta}$ of the parameter of the model denoted by $\theta$ and a consistent estimate of its large sample covariance matrix $C(\hat{\theta})$ are available, one could draw $\tilde{\theta}^{(d)}$ from its asymptotic normal posterior distribution, then draw the missing values $Y_{mis}^{(d)}$ from its posterior predictive distribution. That is, for $d = 1, \ldots, m$,

$$Y_{mis}^{(d)} \sim f(Y_{mis} | Y_{obs}, \tilde{\theta}^{(d)}),$$

where

$$\tilde{\theta}^{(d)} \sim N(\hat{\theta}, C(\hat{\theta})).$$

The MDT method described in Chapter 2 uses the EM algorithm to obtain the ML estimate $\hat{\theta}$ of $\theta$. The large-sample covariance matrix of the parameters, $C(\hat{\theta})$, is not readily available. Therefore, the large sample covariance is not completely specified. To
overcome this, bootstrap sampling similar to the one suggested by Little & Rubin (2002, pp216) is implemented to estimate the posterior distribution of the model parameter \( \theta \). The bootstrap is a technique proposed by Efron (1977) in which the sampling distribution of a statistic is created by resampling from a set of observed data. Bootstrap estimates \((\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(B)})\) can be computed as follows:

For \( b = 1, \ldots, B \),

1) Generate a bootstrap sample \( Y_{obs}^{(boot,b)} \) from the original data set, with replacement, of the same size as the observed sample.

2) Estimate the missing values in \( Y_{obs}^{(boot,b)} \) by applying the MDT procedure to the bootstrap sample \( Y_{obs}^{(boot,b)} \). Notice that the number of missing values may not be the same in each bootstrap sample.

3) Get the bootstrap parameter estimate \( \hat{\theta}^{(b)} \) from each complete bootstrap sample, Then the bootstrap mean and variance can be calculate as,

\[
\hat{\theta}_{boot} = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}^{(b)},
\]

\[
V_{boot} = \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\theta}^{(b)} - \hat{\theta}_{boot})^2.
\]

This bootstrap process is appropriate in Rubin’s theory. That is, the bootstrap samples are asymptotically equivalent to a sample from the posterior distribution of \( \theta \) (Little & Rubin, 2002, pp216).
Once the bootstrap variance $V_{\text{boot}}$ are obtained, $m$ complete data sets can be generated from the joint posterior predictive distribution. That is, for $d = 1, \ldots, m$, the missing values could be drawn as,

$$Y_{\text{mis}}^{(d)} \sim p(Y_{\text{mis}} \mid Y_{\text{obs}}, \tilde{\theta}^{(d)})$$

where $\tilde{\theta}^{(d)}$ is a random draw from multivariate normal distribution with mean $\hat{\theta}$ and variance covariance $V_{\text{boot}}$, where $\hat{\theta}$ is ML estimate from the original data set using the MDT method.

The $m$ complete data sets are analyzed using standard procedures such as the mixed model procedure and regression procedure and so on.

The results from the $m$ complete data sets are combined for valid statistical inferences. SAS PROC MIANALYZE is used to combine the inference from imputed data sets.

The $m$ different sets of the point and variance estimates for a parameter $\theta$ are computed from the $m$ complete data sets. Let $\hat{\theta}_i$ and $\bar{U}_i$ be the point and variance estimates from the $i$th imputed data set $i = 1, 2, \ldots, m$. Rubin (1987, pp76) gives the following rules for combining them. The point estimate for $\theta$ from multiple imputations is simply the average of the $m$ complete-data estimates:

$$\bar{\theta} = \frac{1}{m} \sum_{i=1}^{m} \hat{\theta}_i$$

The within-imputation variance, $\bar{U}$ is the average of the $m$ complete-data variance estimates,
The between-imputation variance, $W$ is the variance of the $m$ complete-data estimates

$$W = \frac{1}{m-1} \sum_{i=1}^{m} (\hat{\theta}_i - \bar{\theta})^2,$$

Then the variance estimate associated with $\bar{\theta}$ is the total variance of $(\theta - \bar{\theta})$,

$$E = \bar{U} + \left(1 + \frac{1}{m}\right)B.$$ 

The statistic $(\theta - \bar{\theta})T^{-\frac{1}{2}}$ is approximately distributed as a student $t$ distribution

$$(\theta - \bar{\theta})E^{-\frac{1}{2}} \sim t_{\nu_m}.$$ 

where $\nu_m$ is the degrees of freedom given by

$$\nu_m = (m-1) \left[ 1 + \frac{\bar{U}}{(1+\frac{1}{m})W} \right]^2. \quad (5.6)$$

The relative increase in variance due to nonresponse is calculated as.

$$r = \frac{(1+\frac{1}{m})W}{\bar{U}}$$

The fraction of missing information about $\theta$ is calculated as

$$\hat{\lambda} = \frac{r + 2/(\nu_m + 3)}{r + 1}$$

In applications, calculation of $r$ and $\hat{\lambda}$ is highly recommended for assessing how the missing data contribute to inferential uncertainty about $\theta$. However, for a small number of
imputations the estimates $r$ and $\hat{\lambda}$ may vary considerably for different seed values. But the inferences regarding the model parameters are often not as sensitive.

Barnard and Rubin (1999) recommended an improved expression for the degrees of freedom for small data sets when the complete data degree of freedom is small and the between imputation variance is small:

$$v^*_m = \left[ \frac{1}{v_m} + \frac{1}{v_{obs}} \right]^{-1},$$

where

$$v_{obs} = (1-r) \left( \frac{v_{com}+1}{v_{com}+3} \right) v_{com}.$$

If the fraction of missing information is modest (e.g. <30%), as few as five multiple imputations (or even three in some cases) is adequate under each model for nonresponse (Rubin, 1996). It can be much less than the acceptable number of simulations for the inference based on the empirical distribution of the draws. For example, in bootstrap or jackknife simulation, hundreds or thousands are often needed to obtain an acceptable level of accuracy. There are two reasons for the validity of a very small imputation number. First, the simulation is only being used to solve the missing data aspect of the problem, with reliance for handling the rest of the information left to the complete data method. Let $\lambda$ be the fraction of missing information about a scalar estimator, the relative efficiency (on the variance scale) of using finite imputation estimator relative to the infinite imputation estimator is $[1 + (\lambda/m)]^{-1/2}$, which is close to one with a realistic fraction of missing information and modest $m$ (Rubin 1987, Table4.1). Second, the rules for
combining the \( m \) complete data analysis are calibrated for the simulation error. Both the variance \( E \) of estimate and degree of freedom \( \nu_m \) contain the predictive amount of error due to finite imputation. Rubin and Schenker (1986) report that multiple-imputation interval are properly adjusted to have at least the nominal coverage in a variety of scenarios even for \( m \) as small as 2.

5.4 Analysis of IMPS data using the MDT method in conjunction with MI

The flowchart in Figure 5.1 describes the 6-step procedure for the MDT method in conjunction with the MI procedure. The numbers in the boxes within parentheses represent the steps of the method described more in detail below.

The IMPS data was used in Chapter 3 to demonstrate the MDT method. A repeated measures mixed model was used to fit the IMPS data within the MDT method. The model includes treatment group, the square root transformation of time (sqrt(time)) and interaction of group by sqrt(time) were used as covariates. Based on the AIC and BIC criteria, first-order autoregression AR(1) was chosen to be the best covariance structure. The treatment effect over time (group by sqrt(time) interaction) was found to be marginally significant (P value = 0.0572). Adjustment was made to the degrees of freedom to account for the inherent missing data estimation. Here, the MI approach provides an alternative to account for the uncertainty in imputation.

In summary, Steps 1-4 are used to generate \( m \) imputed datasets. The missing values \( Y_{mis} \) are imputed by randomly drawing the predictive distributions of the parameters and observations. The missing values may include missing due to truncation as well as missing at random. They are imputed simultaneously. In Step 5 the \( m \) imputed data sets
are analyzed by calling the PROC MIXED procedure from SAS to fit the model described above. In Step 6, PROC MIANALYZE is used to combining results from the \( m \) imputed data sets to generate valid statistical inferences about treatment effect.

In Step 1, the original data set is analyzed by MDT method to get \( \hat{\theta} \), the estimate of the parameter vector of the multivariate truncated normal distribution. The parameter vector includes the parameters of the linear model as well as the variance covariance parameters (under AR(1) structure) \( \sigma^2 \) and \( \rho \), and the truncation threshold \( M \) for the treatment and placebo groups. As pointed out in Section 5.3, the estimation of the variance covariance matrix of \( \hat{\theta} \) is not tractable. Therefore in Step 2, bootstrap sampling is conducted to get estimates of variance covariance matrix of the parameter estimates. Although Efron (1993, pp52) recommends the number of bootstrap samples, \( B \), in the range 50-200, here \( B=1000 \) bootstrap samples were used for higher precision. To preserve the same subject ratio between the treatment and placebo groups as in the original data, the bootstrap samples were obtained by re-sampling the subjects separately for the two groups. Then, the MDT method was applied to each sample to get the parameter estimates \( \hat{\theta}^{(b)} \), \( b=1,\ldots,1000 \) and bootstrap variance covariance estimates of the parameters was calculated

\[
V_{boot} = V(\hat{\theta}^{(b)}) = \frac{1}{1000-1} \sum_{b=1}^{1000} (\hat{\theta}^{(b)} - \hat{\theta}_{boot}) (\hat{\theta}^{(b)} - \hat{\theta}_{boot})',
\]

Where

\[
\hat{\theta}_{boot} = \sum_{b=1}^{1000} \frac{1}{1000} \hat{\theta}^{(b)}.
\]
In Step 3, the $\tilde{\Theta}^{(d)}$, $d = 1, \ldots, m$, was randomly drawn from the $p$-variate normal, $N_p(\tilde{\Theta}, V_{\text{boot}})$, where $p$ is the total number of parameters including means as well as variances and covariances. The number of imputations $m$ is arbitrary and was chosen to be 10. Here the multivariate normality was verified by appropriate diagnostics. However, at this stage the properties of the bootstrap distribution could be studied using histogram and other multivariate tests of normality. Transformations can be made if the normal assumption doesn’t hold. For example, the sampling distribution of the correlation is often negatively skewed. Fisher’s $r$ to $z$ transformation, $\hat{r} = 0.5 \times \log(\frac{1 + \hat{\rho}}{1 - \hat{\rho}})$, where $\hat{\rho}$ is the correlation coefficient estimate, could be used for the correlation $\rho$. The statistic $\hat{r}$ is approximately normally distributed around $r = \frac{1}{2} \log(\frac{1 + \rho}{1 - \rho})$, with a constant standard deviation of $\frac{1}{\sqrt{n-3}}$ (Tong, 1990, pp18). The $\tilde{\rho}^{(d)}$ is obtained by first random drawing $\tilde{r}^{(d)}$ from the multivariate normal distribution with other parameters, then applying the inverse transformation $\tilde{\rho}^{(d)} = \frac{e^{2\tilde{r}^{(d)}} - 1}{e^{2\tilde{r}^{(d)}} + 1}$. In addition, the log transformation is commonly used to normalize variance $\sigma^2$.

In Step 4 missing observations $Y_{\text{mis}}^{(d)}$ in $d$th dataset are drawn from $f(Y_{\text{mis}} \mid Y_{\text{obs}}, \tilde{\Theta}^{(d)})$ to create $d$th imputed dataset and a variable is added to each set to indicate imputation number. Here $f(Y_{\text{mis}} \mid Y_{\text{obs}}, \tilde{\Theta}^{(d)})$ represents a conditional truncated normal distribution giving parameter $\tilde{\Theta}^{(d)}$ and the observed data. Take for example, suppose $Y_{t,\text{mis}}$ is missing
because it is greater than the truncation value $M$ at time $T$. The truncated normal variate $Y_{i,\text{mis}}^{(d)}$ could be generated from the appropriately truncated uniform distribution and then use the inverse of the cdf to obtain the truncated normal variate. In other words, first the conditional mean and variance of $Y_{i,\text{mis}}^{(d)}$ are calculated using:

$$
\tilde{\mu}_{T/i}^{(d)} = \mu_T^{(d)} + \tilde{\Sigma}_T^{(d)}\tilde{\Sigma}_T^{(d)-1}(y_i - \tilde{\mu}^{(d)}),
$$

$$
\tilde{\sigma}_{T/i}^{(d)2} = \sigma_T^{(d)2} - \tilde{\Sigma}_T^{(d)}\tilde{\Sigma}_T^{(d)-1}\tilde{\Sigma}_T^{(d)'},
$$

(5.7)

where $y_i$ are the previous $T-1$ observations for subject $i$.

Second, considering the domain for $Y_{i,\text{mis}}^{(d)}$ is $(\tilde{M}^{(d)}, \infty)$, transfer the lower bound $\tilde{M}^{(d)}$ to z-scores as,

$$
\tilde{z}_{\text{lower}}^{(d)} = (\tilde{M}^{(d)} - \tilde{\mu}_{T/i}^{(d)}) / \tilde{\sigma}_{T/i}^{(d)}.
$$

Third, generate the truncated uniform random variable $\tilde{u}^{(d)}$ in $(\Phi(\tilde{z}_{\text{lower}}^{(d)}), 1)$ by

$$
\tilde{u}^{(d)} = (1 - \Phi(\tilde{z}_{\text{lower}}^{(d)})) \times U + \Phi(\tilde{z}_{\text{lower}}^{(d)}),
$$

Here, $U$ is the uniform random number on $(0, 1)$.

Finally, the normal random variate $Y_{i,\text{mis}}^{(d)}$ is generated as,

$$
Y_{i,\text{mis}}^{(d)} = \tilde{\mu}_{T/i}^{(d)} + \tilde{\sigma}_{T/i}^{(d)} \times \Phi^{-1}(\tilde{u}^{(d)}).
$$

In SAS, $\Phi(\cdot)$, $U$ and $\Phi^{-1}(\cdot)$ are given by function cdf('normal', \cdot), ranuni(seed) and quantile('NORMAL', \cdot) respectively.

For the MAR case, the missing values $Y_{i,\text{mis}}^{(d)}$ are imputed by randomly drawing from the normal distribution with mean and variance as in equation (5.7).
A SAS Macro in PROC IML (SAS 9.1) is used to impute the missing values. The other way to create the imputed dataset is to impute the MDT first and then call for SAS PROC MI procedure to impute the MAR cases.

After creating the $m$ imputed datasets, (step 5) generate the parameter estimates for each of the $m$ imputed datasets using the repeated measures mixed model:

```
PROC MIXED DATA=MIIMPS;
   CLASS TIME;
   MODEL IMPS= GROUP STIME GROUP*STIME/SOLUTION COVB;
   REPEATED TIME/TYPE=AR(1) SUBJECT=ID GROUP=GROUP;
   ODS OUTPUT SOLUTIONF=MIXPAREMS CovB=MIXCOVB;
   BY _IMPUTATION_;
RUN;
```

The output data sets mixparms and mixcovb contain the model parameter estimates, and the covariance matrices associated with these parameter estimates, respectively. In the final step, these two datasets are used to combine the analysis results by PROC MIANALYZE:

```
PROC MIANALYZE PARMS=MIXPAREMS EDF=283
   COVB(EFFECTVAR=ROWCOL)=MIXCOVB;
   MODELEFFECTS INTERCEPT GROUP STIME GROUP*STIME;
RUN;
```

The EDF= option specifies the complete data degrees of freedom for the parameter estimates. The complete data degree of freedom is obtained from the output of the PROC MIXED for each imputation. For the IMPS data it was 283 in all the imputations.
Figure 5.1 Flowchart of MDT method in conjunction with MI
The analysis shows the relative increase in variance due to missing values for treatment effect (i.e. group by sqrt(time) interaction) is 0.055, and the fraction of missing information about treatment effect is 0.053. That is, by combining multiple imputation, about 5.5% more variation is appropriately included in the inference comparing to MDT method alone and about 5.3% missing information is incorporated in the analysis.

(Although, as mentioned in section 5.3., these estimates are seed dependent, in the MDT case, perhaps due to the fact that it is model based rather than random as in the MAR case, when calculated for different seeds these numbers were comparable.) Table 5.2 shows that IMPS score decreases over time (P value = 0.0023) and hypothesis test for treatment effect has P value = 0.0655. Comparing this with MDT method alone, where treatment effect has P value = 0.0572, the treatment effect has become less significant. This demonstrates that, a simple adjustment of the degrees of freedom to account for the imputation of the missing data does not incorporate the uncertainty in the imputation method adequately.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Variance</th>
<th>Relative Increase in Variance</th>
<th>DF</th>
<th>Fraction Missing Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.00002</td>
<td>0.073</td>
<td>0.073</td>
<td>280.93</td>
</tr>
<tr>
<td>group</td>
<td>0.00002</td>
<td>0.094</td>
<td>0.094</td>
<td>280.95</td>
</tr>
<tr>
<td>stime</td>
<td>0.00111</td>
<td>0.021</td>
<td>0.023</td>
<td>245.12</td>
</tr>
<tr>
<td>group*stime</td>
<td>0.00148</td>
<td>0.029</td>
<td>0.031</td>
<td>246.24</td>
</tr>
</tbody>
</table>
Table 5.2 Parameter estimates using MDT in conjunction with MI

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std Error</th>
<th>DF</th>
<th>Minimum</th>
<th>Maximum</th>
<th>t for H₀:θ=0</th>
<th>Pr &gt;</th>
<th>t</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.194</td>
<td>0.270</td>
<td>280.93</td>
<td>5.188</td>
<td>5.203</td>
<td>19.23</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>group</td>
<td>0.037</td>
<td>0.307</td>
<td>280.95</td>
<td>0.030</td>
<td>0.043</td>
<td>0.12</td>
<td>0.9045</td>
<td></td>
<td></td>
</tr>
<tr>
<td>stime</td>
<td>-0.465</td>
<td>0.151</td>
<td>245.12</td>
<td>-0.513</td>
<td>-0.385</td>
<td>-3.08</td>
<td>0.0023</td>
<td></td>
<td></td>
</tr>
<tr>
<td>group*stime</td>
<td>-0.325</td>
<td>0.176</td>
<td>246.24</td>
<td>-0.419</td>
<td>-0.284</td>
<td>-1.85</td>
<td>0.0655</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.5 Comparison with MI procedure treating all missing as MAR

To compare how the MDT method in conjunction with multiple imputation improves upon the commonly used multiple imputation method assuming missing are MAR, the IMPS data were analyzed ignoring the missing mechanism and assuming all the missing data are MAR. SAS PROC MI and PROC MIANALYZE procedures were implemented for the analysis. In this case, (for the specific seed used) the relative increase in variance due to missing values for treatment effect (i.e. group by sqrt(time) interaction) is 0.037, and the fraction of missing information about treatment effect is 0.036 (Table 5.3). That is, by combining multiple imputation, about 3.7% more variation is appropriately included in the inference comparing repeated measures mixed model and about 3.6% missing information is incorporated in the analysis. (However, when seed was changed these numbers varied considerably.)

Table 5.4 shows that IMSP score decreases over time (P value < 0.001) and hypothesis test for treatment effect is not significant (P value = 0.2972). Compared to the results from multiple imputation treating all the missing as MAR, the MDT in conjunction with multiple imputation has a higher power to detect significant treatment effect for IMPS.
data (P value = 0.0655 vs. 0.2972) by incorporating the missing data mechanism into the analysis.

Table 5.3 Multiple imputation variance information using MI assuming MAR

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Variance</th>
<th>Relative Increase in Variance</th>
<th>Fraction Missing Information</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Between</td>
<td>Within</td>
<td>Total</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.00009</td>
<td>0.080</td>
<td>0.080</td>
</tr>
<tr>
<td>Group</td>
<td>0.00007</td>
<td>0.100</td>
<td>0.020</td>
</tr>
<tr>
<td>Stime</td>
<td>0.00041</td>
<td>0.017</td>
<td>0.017</td>
</tr>
<tr>
<td>group*stime</td>
<td>0.00068</td>
<td>0.022</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Table 5.4 Multiple Imputation parameter Estimates using MI assuming MAR

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Std Error</th>
<th>DF</th>
<th>Minimum</th>
<th>Maximum</th>
<th>0</th>
<th>Pr &gt;</th>
<th>t for H0: parameter=</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.252</td>
<td>0.283</td>
<td>280.71</td>
<td>5.245</td>
<td>5.265</td>
<td>18.56</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>group</td>
<td>0.016</td>
<td>0.316</td>
<td>280.78</td>
<td>0.007</td>
<td>0.027</td>
<td>0.05</td>
<td>0.9605</td>
<td></td>
</tr>
<tr>
<td>Stime</td>
<td>-0.589</td>
<td>0.132</td>
<td>258.49</td>
<td>-0.611</td>
<td>-0.564</td>
<td>-4.47</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>group*stime</td>
<td>-0.159</td>
<td>0.152</td>
<td>249.89</td>
<td>-0.199</td>
<td>-0.130</td>
<td>-1.04</td>
<td>0.2972</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 6
Summary and Extensions

6.1 Dissertation Summary

Missing values are a problem in many data sets and seem especially common in the medical and social sciences. Recent years have seen a growing interest in statistical methods that properly account for incomplete data. The choice of appropriate methods to handle missing data depends on the missing data mechanism (Rubin, 1976; Little & Rubin, 1987, 1989; Schafer, 1997). A number of methods that provide for a nonignorable dropout process have been proposed (Diggle & Kenward 1994; Little & Wang, 1996, Little & Raghunathan, 1999, Albert, 2000; Wu & Albert, 2001). These methods make various assumptions about the mechanism of drop-out, and their utility in practice depends on whether these assumptions apply in a specific application (Little, 1994).

In this dissertation, a method proposed to handle nonignorable missing data, especially, missing data due to its value beyond some unobserved threshold was investigated. Clinical trials comparing treatments that follow patients over a period of time often suffer from dropouts. In some cases these dropouts are treatment related. It may be due to clinical improvement or deterioration. When treatment related dropouts occur the distribution of the observations often resemble a truncated normal. A missing due to truncation method (MDT method) was proposed by Ramakrishna and Wang (2005) to
analyze such data under a truncated multivariate normal distribution. They also proposed an EM algorithm to simplify the estimation of the truncated normal likelihood and to utilize standard software (such as SAS) for the analysis. However, the application of the EM algorithm was not formalized and the properties of the MDT method were not formally studied. Further, they did not explore the relationship of their method to the more widely accepted multiple imputation technique.

Therefore, the first objective of this dissertation was to apply EM algorithm to the MDT method using IMPS data set from a collaborative study conducted by the National Institute of Mental Health, and to conduct a simulation study to compare MDT method with some other widely used imputation methods. The second objective of the dissertation is to develop a method that combines the MDT method with Rubin’s multiple imputation to handling nonignorable missing due to truncation data.

In chapter 3, A SAS Macro in PROC IML (SAS 9.1) was used to implement the EM algorithm. The code for the E-step was entirely written specifically for the MDT procedure. The M-step, however, was achieved by calling the PROC MIXED procedure from SAS. The analysis results from the IMPS data showed MDT method improved precision and increased power to detect a marginally significant treatment effect by appropriately incorporating the missing data mechanism.

In Chapter 4, a simulation study was carried out to compare the performance of other relevant methods with the MDT method. In Chapter 5, IMPS data were used to demonstrate how the MDT method can be combined with multiple imputation procedure to incorporate the uncertainty due to imputation.
6.2 Extension and suggestions for future research

The research presented in this dissertation could be extended further. A few possible areas of interest are listed below.

6.2.1 Extend MDT method to multivariate outcomes

The MDT method described in Chapter 2 was developed for longitudinal studies with one outcome variable. Multivariate outcomes are frequently measured in longitudinal clinical studies, and missing data due to patients dropping out of the study also is a common problem in these studies (Tilley and Marler et al, 1996; Daskalakis and Laird et al 2002; Roy and Lin, 2002). Accordingly, the MDT method could be extended to analyze multivariate outcomes with data missing due to truncation. In this case, once a subject drops out of the study, all the outcomes from that subject will be missing from that time point. If the dropout is due to all the outcomes passing the thresholds, the missing outcomes could be assumed from a multivariate truncated distribution where all the variables are truncated. If the dropout is due to some of the outcomes passing the thresholds, only these outcomes should be treated as MDT, others should be MAR. In both situation, the estimation of missing outcomes on the same subjects at a certain time point should be simultaneously performed and therefore requires multivariate version of equation (2.8) - (2.10). After estimation of the missing values, the parameter estimation and hypotheses test could be done by SAS PROC MIXED as in the univariate outcome case. For instance, in order to apply the SAS PROC MIXED procedure for multivariate outcomes, first arrange all the outcome variables in a vector form (instead of a matrix as in the case of PROC GLM). In addition, define a new variable to indicate which outcome it
is for each subject and outcome. Add the outcome indicator variable along with its interactions with other variables in the MODEL statement. Then use two distinct repeated effects in REPEATED statement (SAS PROC MIXED) to specify the repeated cases across time as well as across outcomes. The covariance structure needs to be specified as direct (Kronecker) product structures designed for multivariate repeated measures (Galecki, 1994). Currently, the available direct product covariance structures in SAS PROC MIXED are UN@AR(1), UN@CS, UN@UN. The first factor of Kronecker product models the covariance across the multivariate observations and the second factor models the covariance across time.

For example, suppose observed data consist of heights, weights and systolic blood pressure of several children measured over successive years. The input data set then contains: \( Y \), all of the heights, weights and blood pressures, YEAR, indicating the year of measurement and ID, indicating each child on which the measurement was taken. In addition, define a variable, say VAR, indicating whether the measurement is a height, a weight or a blood pressure (by a number or character). Then the PROC MIXED code for a Kronecker AR(1) structure across years would be

```plaintext
PROC MIXED;
   CLASS VAR YEAR ID;
   MODEL Y= VAR YEAR VAR*YEAR;
   REPEATED VAR YEAR/TYPE= UN@AR SUBJECT=ID;
RUN;
```

In Section 1.6.1, a multivariate outcome data set on methadone treatment was introduced. A method for analyzing these data was proposed by Roy and Lin (2000 & 2002). A briefly summary of their method follows.
Assume these three outcome variables measure an underlying latent variable, $U$, a treatment practice effectiveness score with errors that might be correlated over time. Let $y_{ijk}$ be the $j$th outcome ($j=1, \ldots, J$) measured at the $k$th time point ($k=1, \ldots, K$) for unit $i$ ($i=1, \ldots, n$). And let $U_{ik}$ be the unobserved latent variable measured at the $k$th time point for unit $i$. The three outcomes are related to the latent variable by the longitudinal latent variable model,

$$y_{ijk} = \beta_{0j} + \beta_{1j}U_{ik} + b_{ij} + e_{ijk},$$

where $\beta_j = (\beta_{0j}, \beta_{1j})^T$ is a vector of regression coefficients, the $b_{ij}$ models the correlation between the measurement on the same outcome over time within each unit $i$, distributed as $N(0, \xi_j)$, the $e_{ijk}$ are error term distributed as $N(0, \tau^2_j)$.

To study the covariates effects, $U_{ik}$ is regressed on the covariates $X_{ik}$ (e.g., unit and client characteristics). A linear mixed model (Lair & Ware, 1982) is specified for the dependence of $U_{ik}$ on $X_{ik}$. More specifically, a random intercept model is used to account for the underlying longitudinal measurements on $U_{ik}$. That is,

$$U_{ik} = X_{ik}'\alpha + Z_{ik}'a_i + c_{ik},$$

where $\alpha$ is a $p \times 1$ vector of regression coefficients, $Z_{ik}$ is $q \times 1$ covariate vector, $a_i$ is a random effect vector distributed as $N(0, D(\theta))$, $D$ is variance covariance matrix parameterized by a vector of variance components $\theta$ and $c_{ik}$ are residuals distributed as $N(0,1)$. 
A selection model is used to model nonignorable dropouts, where the dropout probability depends on the historical data. That is, the probability of dropout depends on the latent treatment practice effectiveness score at the dropout time point and at the previous time point. For example the selection model may be,

$$\log \operatorname{it}(p_{ik}) = \psi_0 + \psi_1 U_{i,k-1} + \psi_2 U_{ik},$$

where $\psi_2 = -1, -0.5, 0, 0.5, 1$ with $\psi_2 = 0$ indicating ignorable missing.

A transition model of these covariates is developed to model missing time-varying covariates (such as percentage of staff that was ex-addicted, percentage of African-American clients) at the time of dropout. The covariate vector $X_{ik} (p \times 1)$ could be partitioned into two parts: a complete covariate vector, $S_{ik} (p_1 \times 1)$, and an incomplete covariate vector, $T_{ik} (p_2 \times 1)$. Let $T_{ikl}$ be the value of the $l$th covariate of $T_{ik}$ ($l = 1, \ldots, p_2$) at the $k$th time point for unit $i$. The transition model is assumed as $T_{ikl} = \lambda_{q_{il}} + \lambda_{il} T_{i,k-1,l} + d_{ikl}$ for each $T_{ik}$ ($l = 1, \ldots, p_2$), where $\lambda_{q_{il}}$ and $\lambda_{il}$ are regression coefficients and $d_{ikl}$ is independently distributed as $N(0, \sigma_i^2)$. This transition model allows for correlation within covariates over time as well as the cross-sectional correlation among different covariates $T_{ikl}$ ($l = 1, \ldots, p_2$) at the same time point.

The EM algorithm is developed to estimate the model parameters by simplifying the multidimensional integration in full-likelihood estimation.
6.2.2. Allow for the threshold to be random

The proposed MDT method in Chapter 2 assumes the truncation occurs at the same threshold for all the individuals in same treatment group. Even though this is reasonable in most applications there may be some cases where thresholds may vary among subjects. For example, some patients may have lower tolerance limit to the treatment effects than others. That is, they are more likely to drop out if the treatment effect is not significant enough. One possible approach to accommodate this may be to estimate the unobservable subject-specific thresholds assuming they are randomly distributed among individuals. If thresholds are known to be in a certain range based on clinical knowledge; one could assume that the threshold for each individual is a random observation from a uniform distribution in that range. Another approach may be to treat the threshold of dropping out depending on the patient’s previous observed responses. For example, if clinical improvement between follow-ups does not reach certain level, say \( b \), the patient will drop out. Estimation of the subject specific threshold could incorporate this information by specifying the threshold to be the subject’s observation prior to dropout + \( b \).

There are some other possible extensions to this dissertation. When the MDT occurs at multiple time points they were estimated sequentially starting from first occurrence of MDT. The estimation of MDT at subsequent time points were conditioned on the subject’s observations prior to dropout including the imputed MDT. A more efficient way perhaps would be to estimate the multiple MDT simultaneously using a multivariate truncated normal distribution. However the estimation will involve multivariate normal cdfs which requires numerical method such as a dimension reduction
method, a Monte Carlo method or a quadrature method in numerical analysis (Tong, 1990, pp187-191). Genz (1993) compared several numerical computation methods for multivariate normal probability such as Deak’s methods using a transformation to a spherical coordinate system, Genz’s methods using a transformation of the original integration region to the unit hypercube and Schervish’s methods using a locally adaptive numerical integration algorithm. Genz concluded that multivariate normal probabilities can be robustly and reliably computed as low to moderate accuracy levels for problems with up to ten dimensions. High accuracy or high dimension problem can require long computation times for these methods and it is not clear what is the best method for this type of problem.
Literature Cited


SAS OlineDoc®, version9.1(2002). SAS Institute Inc. Cary, NC. SAS® and all other SAS Institute Inc. product and service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.


Wang Z.Y. (1995) Imputation of Data Missing Due to Truncation----Maximum Likelihood


APPENDIX A

IMPS Data Listing and Histogram Plots: Chapter 3
## APPENDIX A

### IMPS Data Listing and Histogram Plots: Chapter 3

Table A.1 IMPS data listing

<table>
<thead>
<tr>
<th>ID</th>
<th>Reason of Missing</th>
<th>Weeks 0</th>
<th>Weeks 1</th>
<th>Weeks 3</th>
<th>Weeks 6</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1129</td>
<td>Completed</td>
<td>6.0</td>
<td>4.5</td>
<td>3.50</td>
<td>2.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>1306</td>
<td>Completed</td>
<td>4.5</td>
<td>5.0</td>
<td>1.50</td>
<td>2.5</td>
<td>Placebo</td>
</tr>
<tr>
<td>1317</td>
<td>Completed</td>
<td>6.0</td>
<td>5.5</td>
<td>6.00</td>
<td>5.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>2153</td>
<td>Completed</td>
<td>5.0</td>
<td>3.5</td>
<td>3.00</td>
<td>4.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>2302</td>
<td>Completed</td>
<td>3.0</td>
<td>4.0</td>
<td>4.33</td>
<td>4.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>2308</td>
<td>Completed</td>
<td>5.0</td>
<td>3.5</td>
<td>4.00</td>
<td>3.5</td>
<td>Placebo</td>
</tr>
<tr>
<td>3114</td>
<td>Completed</td>
<td>6.0</td>
<td>6.0</td>
<td>6.00</td>
<td>5.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>3302</td>
<td>Completed</td>
<td>4.5</td>
<td>4.0</td>
<td>4.00</td>
<td>1.5</td>
<td>Placebo</td>
</tr>
<tr>
<td>5319</td>
<td>Completed</td>
<td>5.0</td>
<td>5.0</td>
<td>5.00</td>
<td>5.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>6301</td>
<td>Completed</td>
<td>4.5</td>
<td>4.0</td>
<td>4.50</td>
<td>4.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>7104</td>
<td>Completed</td>
<td>4.5</td>
<td>4.0</td>
<td>4.00</td>
<td>3.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>7105</td>
<td>Completed</td>
<td>6.0</td>
<td>5.0</td>
<td>5.50</td>
<td>5.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>7109</td>
<td>Completed</td>
<td>6.0</td>
<td>6.0</td>
<td>6.00</td>
<td>6.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>7114</td>
<td>Completed</td>
<td>4.0</td>
<td>3.5</td>
<td>2.50</td>
<td>2.5</td>
<td>Placebo</td>
</tr>
<tr>
<td>8112</td>
<td>Completed</td>
<td>4.0</td>
<td>4.5</td>
<td>2.50</td>
<td>2.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>8302</td>
<td>Completed</td>
<td>4.5</td>
<td>5.0</td>
<td>5.00</td>
<td>3.5</td>
<td>Placebo</td>
</tr>
<tr>
<td>9106</td>
<td>Completed</td>
<td>5.0</td>
<td>4.5</td>
<td>4.50</td>
<td>3.0</td>
<td>Placebo</td>
</tr>
<tr>
<td>9309</td>
<td>Completed</td>
<td>6.5</td>
<td>5.5</td>
<td>6.00</td>
<td>5.50</td>
<td>Placebo</td>
</tr>
<tr>
<td>3320</td>
<td>Other</td>
<td>6.0</td>
<td>1.0</td>
<td>.</td>
<td>.</td>
<td>Placebo</td>
</tr>
<tr>
<td>ID</td>
<td>Reason of Missing</td>
<td>Weeks</td>
<td>Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------------------------</td>
<td>-------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2106</td>
<td>Treatment Failure</td>
<td>6.0</td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2149</td>
<td>Treatment Failure</td>
<td>6.0</td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2320</td>
<td>Treatment Failure</td>
<td>6.5</td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6105</td>
<td>Treatment Failure</td>
<td>5.0</td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6116</td>
<td>Treatment Failure</td>
<td>6.0</td>
<td>Placebo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1109</td>
<td>Completed</td>
<td>4.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1113</td>
<td>Completed</td>
<td>4.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1114</td>
<td>Completed</td>
<td>6.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1308</td>
<td>Completed</td>
<td>6.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1314</td>
<td>Completed</td>
<td>6.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1318</td>
<td>Completed</td>
<td>3.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2105</td>
<td>Completed</td>
<td>4.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2108</td>
<td>Completed</td>
<td>6.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2112</td>
<td>Completed</td>
<td>2.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2126</td>
<td>Completed</td>
<td>4.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2132</td>
<td>Completed</td>
<td>5.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2136</td>
<td>Completed</td>
<td>6.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2142</td>
<td>Completed</td>
<td>5.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2159</td>
<td>Completed</td>
<td>5.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2163</td>
<td>Completed</td>
<td>5.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2312</td>
<td>Completed</td>
<td>6.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2316</td>
<td>Completed</td>
<td>5.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2336</td>
<td>Completed</td>
<td>5.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3102</td>
<td>Completed</td>
<td>6.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3103</td>
<td>Completed</td>
<td>5.5</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3108</td>
<td>Completed</td>
<td>6.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3303</td>
<td>Completed</td>
<td>5.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3307</td>
<td>Completed</td>
<td>5.0</td>
<td>Treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Reason of Missing</td>
<td>Weeks</td>
<td></td>
<td></td>
<td>Group</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------------------</td>
<td>-------</td>
<td>----------------</td>
<td>----------------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>3309</td>
<td>Completed</td>
<td>6.0</td>
<td>6.0</td>
<td>5.50</td>
<td>5.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>4103</td>
<td>Completed</td>
<td>5.5</td>
<td>4.5</td>
<td>5.00</td>
<td>3.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>4302</td>
<td>Completed</td>
<td>5.0</td>
<td>5.0</td>
<td>5.00</td>
<td>4.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>4504</td>
<td>Completed</td>
<td>6.0</td>
<td>5.5</td>
<td>5.50</td>
<td>4.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>4518</td>
<td>Completed</td>
<td>6.5</td>
<td>5.0</td>
<td>5.00</td>
<td>4.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>4522</td>
<td>Completed</td>
<td>3.5</td>
<td>5.5</td>
<td>5.00</td>
<td>4.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>4702</td>
<td>Completed</td>
<td>6.5</td>
<td>5.5</td>
<td>6.00</td>
<td>5.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>4718</td>
<td>Completed</td>
<td>6.5</td>
<td>4.0</td>
<td>5.50</td>
<td>3.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>5101</td>
<td>Completed</td>
<td>5.0</td>
<td>4.5</td>
<td>2.50</td>
<td>2.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>5107</td>
<td>Completed</td>
<td>6.0</td>
<td>5.5</td>
<td>5.50</td>
<td>4.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>5113</td>
<td>Completed</td>
<td>5.0</td>
<td>4.0</td>
<td>4.00</td>
<td>3.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>5124</td>
<td>Completed</td>
<td>5.0</td>
<td>4.0</td>
<td>3.00</td>
<td>2.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>5126</td>
<td>Completed</td>
<td>5.0</td>
<td>5.0</td>
<td>3.00</td>
<td>3.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>5306</td>
<td>Completed</td>
<td>4.0</td>
<td>4.0</td>
<td>2.00</td>
<td>2.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>5314</td>
<td>Completed</td>
<td>3.5</td>
<td>4.0</td>
<td>2.00</td>
<td>4.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>5328</td>
<td>Completed</td>
<td>6.0</td>
<td>5.0</td>
<td>5.00</td>
<td>5.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>6103</td>
<td>Completed</td>
<td>7.0</td>
<td>7.0</td>
<td>6.00</td>
<td>5.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>6109</td>
<td>Completed</td>
<td>5.0</td>
<td>4.0</td>
<td>3.00</td>
<td>3.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>6113</td>
<td>Completed</td>
<td>4.0</td>
<td>4.0</td>
<td>3.00</td>
<td>3.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>6310</td>
<td>Completed</td>
<td>6.0</td>
<td>6.0</td>
<td>5.00</td>
<td>5.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>6315</td>
<td>Completed</td>
<td>5.5</td>
<td>3.5</td>
<td>4.00</td>
<td>3.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>7111</td>
<td>Completed</td>
<td>4.5</td>
<td>5.5</td>
<td>5.00</td>
<td>4.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>7119</td>
<td>Completed</td>
<td>6.0</td>
<td>5.5</td>
<td>5.50</td>
<td>5.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>7120</td>
<td>Completed</td>
<td>4.5</td>
<td>4.5</td>
<td>4.00</td>
<td>3.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>7310</td>
<td>Completed</td>
<td>4.5</td>
<td>3.5</td>
<td>2.50</td>
<td>2.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>7319</td>
<td>Completed</td>
<td>5.5</td>
<td>4.5</td>
<td>4.00</td>
<td>3.50</td>
<td>Treatment</td>
</tr>
<tr>
<td>8105</td>
<td>Completed</td>
<td>4.5</td>
<td>5.5</td>
<td>4.00</td>
<td>4.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>8107</td>
<td>Completed</td>
<td>4.0</td>
<td>3.0</td>
<td>2.00</td>
<td>2.00</td>
<td>Treatment</td>
</tr>
<tr>
<td>ID</td>
<td>Reason of Missing</td>
<td>Weeks</td>
<td>Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------------------</td>
<td>-------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8108</td>
<td>Completed</td>
<td>4.0</td>
<td>3.00 2.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8110</td>
<td>Completed</td>
<td>6.0</td>
<td>5.50 5.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8116</td>
<td>Completed</td>
<td>6.0</td>
<td>5.50 3.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8126</td>
<td>Completed</td>
<td>5.0</td>
<td>4.50 4.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8305</td>
<td>Completed</td>
<td>5.0</td>
<td>4.50 2.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8306</td>
<td>Completed</td>
<td>4.0</td>
<td>4.00 2.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8313</td>
<td>Completed</td>
<td>6.0</td>
<td>5.00 3.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9111</td>
<td>Completed</td>
<td>6.0</td>
<td>5.00 6.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9312</td>
<td>Completed</td>
<td>5.5</td>
<td>2.50 3.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9314</td>
<td>Completed</td>
<td>5.0</td>
<td>4.00 4.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2113</td>
<td>Other</td>
<td>5.0</td>
<td>6.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2123</td>
<td>Other</td>
<td>6.0</td>
<td>5.50 2.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2331</td>
<td>Other</td>
<td>4.5</td>
<td>4.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3308</td>
<td>Other</td>
<td>5.5</td>
<td>5.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3314</td>
<td>Other</td>
<td>5.5</td>
<td>5.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4704</td>
<td>Other</td>
<td>6.0</td>
<td>1.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4711</td>
<td>Other</td>
<td>6.5</td>
<td>2.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4506</td>
<td>Treatment Improvement</td>
<td>5.5</td>
<td>4.50 2.50</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2121</td>
<td>Treatment Failure</td>
<td>5.0</td>
<td>5.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2372</td>
<td>Treatment Failure</td>
<td>6.0</td>
<td>6.00</td>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A. Three weeks

![Histogram of IMPS scores for placebo group at 3 weeks](image)

B. Six weeks

![Histogram of IMPS scores for placebo group at 6 weeks](image)

Figure A.1 Histogram of IMPS scores for placebo group at 3 weeks and 6 weeks
A. Three weeks

![Histogram of IMPS scores for treatment group at 3 weeks and 6 weeks](image)

B. Six weeks

![Histogram of IMPS scores for treatment group at 3 weeks and 6 weeks](image)

Figure A.2 Histogram of IMPS scores for treatment group at 3 weeks and 6 weeks
APPENDIX B

Simulation Results: Chapter 4
# APPENDIX B

## Simulation Results: Chapter 4

Table B.1 Mean ($\mu_4 = 2.6$) estimates (s.d) from different methods for linear response and MDT at last time point

<table>
<thead>
<tr>
<th>Method</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.652(0.188)</td>
<td>2.685(0.189)</td>
<td>2.747(0.197)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.030(1.359)</td>
<td>5.570(1.289)</td>
<td>8.190(1.426)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.790(0.171)</td>
<td>2.938(0.170)</td>
<td>3.178(0.185)</td>
</tr>
<tr>
<td>REG</td>
<td>2.695(0.200)</td>
<td>2.764(0.218)</td>
<td>2.846(0.285)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.760(0.181)</td>
<td>2.881(0.181)</td>
<td>3.102(0.191)</td>
</tr>
</tbody>
</table>

For $\rho = 0.2$

<table>
<thead>
<tr>
<th>Method</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.639(0.191)</td>
<td>2.663(0.192)</td>
<td>2.708(0.195)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.120(1.647)</td>
<td>5.770(1.413)</td>
<td>8.410(1.450)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.760(0.169)</td>
<td>2.879(0.167)</td>
<td>3.096(0.155)</td>
</tr>
<tr>
<td>REG</td>
<td>2.648(0.194)</td>
<td>2.666(0.219)</td>
<td>2.710(0.244)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.741(0.178)</td>
<td>2.846(0.182)</td>
<td>3.042(0.176)</td>
</tr>
</tbody>
</table>

For $\rho = 0.4$

<table>
<thead>
<tr>
<th>Method</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.592(0.201)</td>
<td>2.582(0.205)</td>
<td>2.570(0.213)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.440(1.647)</td>
<td>6.290(1.066)</td>
<td>8.850(1.403)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.692(0.193)</td>
<td>2.771(0.180)</td>
<td>2.918(0.175)</td>
</tr>
<tr>
<td>REG</td>
<td>2.592(0.199)</td>
<td>2.584(0.206)</td>
<td>2.573(0.218)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.659(0.195)</td>
<td>2.704(0.188)</td>
<td>2.788(0.192)</td>
</tr>
</tbody>
</table>
Table B.2 Variance ($\sigma^2=2$) estimates (s.d) from different methods for linear response and MDT at last time point

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>1.949(0.222)</td>
<td>1.916(0.213)</td>
<td>1.864(0.210)</td>
<td>1.948(0.140)</td>
<td>1.916(0.135)</td>
<td>1.868(0.127)</td>
<td>1.947(0.100)</td>
<td>1.916(0.098)</td>
<td>1.867(0.094)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.03(1.359)</td>
<td>5.57(1.289)</td>
<td>8.19(1.426)</td>
<td>4.36(0.927)</td>
<td>5.77(0.839)</td>
<td>8.50(1.000)</td>
<td>4.57(0.700)</td>
<td>6.07(0.624)</td>
<td>8.19(1.426)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.915(0.219)</td>
<td>1.877(0.213)</td>
<td>1.828(0.214)</td>
<td>1.916(0.136)</td>
<td>1.877(0.135)</td>
<td>1.833(0.135)</td>
<td>1.914(0.099)</td>
<td>1.877(0.101)</td>
<td>1.835(0.102)</td>
</tr>
<tr>
<td>REG</td>
<td>2.038(0.316)</td>
<td>2.113(0.355)</td>
<td>2.309(0.420)</td>
<td>2.027(0.209)</td>
<td>2.098(0.228)</td>
<td>2.295(0.275)</td>
<td>2.028(0.140)</td>
<td>2.103(0.170)</td>
<td>2.310(0.220)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.915(0.220)</td>
<td>1.872(0.217)</td>
<td>1.824(0.218)</td>
<td>1.914(0.060)</td>
<td>1.873(0.137)</td>
<td>1.828(0.133)</td>
<td>1.911(0.100)</td>
<td>1.871(0.100)</td>
<td>1.827(0.099)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>1.950(0.084)</td>
<td>1.924(0.236)</td>
<td>1.880(0.230)</td>
<td>1.953(0.151)</td>
<td>1.927(0.147)</td>
<td>1.884(0.143)</td>
<td>1.953(0.107)</td>
<td>1.928(0.105)</td>
<td>1.886(0.103)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.12(1.647)</td>
<td>5.77(1.413)</td>
<td>8.41(1.450)</td>
<td>4.45(0.914)</td>
<td>5.98(1.034)</td>
<td>8.82(1.067)</td>
<td>4.80(0.696)</td>
<td>6.38(0.722)</td>
<td>8.41(1.450)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.898(0.231)</td>
<td>1.850(0.232)</td>
<td>1.799(0.227)</td>
<td>1.901(0.142)</td>
<td>1.855(0.142)</td>
<td>1.807(0.144)</td>
<td>1.899(0.103)</td>
<td>1.855(0.104)</td>
<td>1.807(0.104)</td>
</tr>
<tr>
<td>REG</td>
<td>2.054(0.339)</td>
<td>2.148(0.384)</td>
<td>2.339(0.438)</td>
<td>2.045(0.213)</td>
<td>2.134(0.251)</td>
<td>2.344(0.293)</td>
<td>2.046(0.155)</td>
<td>2.144(0.187)</td>
<td>2.346(0.220)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.905(0.234)</td>
<td>1.860(0.234)</td>
<td>1.813(0.231)</td>
<td>1.906(0.145)</td>
<td>1.863(0.145)</td>
<td>1.817(0.144)</td>
<td>1.904(0.104)</td>
<td>1.862(0.104)</td>
<td>1.817(0.104)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>1.993(0.355)</td>
<td>2.006(0.345)</td>
<td>2.004(0.340)</td>
<td>1.993(0.221)</td>
<td>2.003(0.216)</td>
<td>2.004(0.215)</td>
<td>1.999(0.160)</td>
<td>2.006(0.160)</td>
<td>2.006(0.158)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.44(1.647)</td>
<td>6.29(1.066)</td>
<td>8.85(1.403)</td>
<td>4.90(1.159)</td>
<td>6.70(0.893)</td>
<td>9.31(1.152)</td>
<td>5.29(0.715)</td>
<td>7.16(0.678)</td>
<td>8.85(1.403)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.861(0.314)</td>
<td>1.795(0.309)</td>
<td>1.720(0.301)</td>
<td>1.863(0.030)</td>
<td>1.796(0.193)</td>
<td>1.722(0.186)</td>
<td>1.865(0.142)</td>
<td>1.793(0.140)</td>
<td>1.720(0.136)</td>
</tr>
<tr>
<td>REG</td>
<td>2.050(0.406)</td>
<td>2.103(0.409)</td>
<td>2.209(0.417)</td>
<td>2.053(0.260)</td>
<td>2.110(0.266)</td>
<td>2.221(0.285)</td>
<td>2.053(0.170)</td>
<td>2.121(0.176)</td>
<td>2.234(0.189)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.903(0.046)</td>
<td>1.870(0.330)</td>
<td>1.832(0.327)</td>
<td>1.903(0.202)</td>
<td>1.867(0.204)</td>
<td>1.832(0.206)</td>
<td>1.905(0.148)</td>
<td>1.867(0.151)</td>
<td>1.834(0.151)</td>
</tr>
</tbody>
</table>
Table B.3 Correlation estimates (s.d) from different methods for linear response and MDT at last time point

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>$n = 50$</th>
<th>$n = 100$</th>
<th>$n = 200$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>0.186(0.090)</td>
<td>0.186(0.090)</td>
<td>0.186(0.089)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.030(1.359)</td>
<td>5.570(1.289)</td>
<td>8.190(1.426)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.189(0.085)</td>
<td>0.200(0.084)</td>
<td>0.221(0.081)</td>
</tr>
<tr>
<td>REG</td>
<td>0.208(0.091)</td>
<td>0.235(0.091)</td>
<td>0.283(0.087)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.179(0.088)</td>
<td>0.177(0.087)</td>
<td>0.171(0.085)</td>
</tr>
<tr>
<td></td>
<td>$\rho = 0.2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>0.384(0.084)</td>
<td>0.384(0.085)</td>
<td>0.384(0.084)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.120(1.647)</td>
<td>5.770(1.413)</td>
<td>8.410(1.450)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.378(0.084)</td>
<td>0.378(0.081)</td>
<td>0.389(0.080)</td>
</tr>
<tr>
<td>REG</td>
<td>0.400(0.085)</td>
<td>0.416(0.085)</td>
<td>0.448(0.081)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.372(0.085)</td>
<td>0.364(0.082)</td>
<td>0.358(0.083)</td>
</tr>
<tr>
<td></td>
<td>$\rho = 0.4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>0.790(0.045)</td>
<td>0.792(0.045)</td>
<td>0.792(0.046)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.440(1.647)</td>
<td>6.290(1.066)</td>
<td>8.850(1.403)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.777(0.046)</td>
<td>0.768(0.048)</td>
<td>0.760(0.048)</td>
</tr>
<tr>
<td>REG</td>
<td>0.792(0.045)</td>
<td>0.793(0.045)</td>
<td>0.797(0.043)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.782(0.331)</td>
<td>0.778(0.048)</td>
<td>0.774(0.049)</td>
</tr>
</tbody>
</table>
### Table B.4 Mean ($\mu_3$) estimates (s.d) from different methods for linear response and MDT at last two time points

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.958(0.186)</td>
<td>2.925(0.183)</td>
<td>2.972(0.188)</td>
<td>2.960(0.123)</td>
<td>2.927(0.115)</td>
<td>2.974(0.124)</td>
<td>2.955(0.095)</td>
<td>2.924(0.095)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.230(1.469)</td>
<td>7.490(1.446)</td>
<td>11.99(1.987)</td>
<td>5.660(1.130)</td>
<td>7.850(1.158)</td>
<td>12.39(1.530)</td>
<td>5.870(0.761)</td>
<td>8.220(0.799)</td>
</tr>
<tr>
<td>LOCF</td>
<td>3.021(0.174)</td>
<td>3.040(0.173)</td>
<td>3.170(0.180)</td>
<td>3.027(0.112)</td>
<td>3.043(0.111)</td>
<td>3.167(0.122)</td>
<td>3.018(0.093)</td>
<td>3.040(0.091)</td>
</tr>
<tr>
<td>REG</td>
<td>2.983(0.198)</td>
<td>2.980(0.242)</td>
<td>3.016(0.319)</td>
<td>2.977(0.142)</td>
<td>2.966(0.166)</td>
<td>3.001(0.216)</td>
<td>2.967(0.111)</td>
<td>2.955(0.123)</td>
</tr>
<tr>
<td>MIXED</td>
<td>3.029(0.180)</td>
<td>3.063(0.176)</td>
<td>3.246(0.182)</td>
<td>3.040(0.114)</td>
<td>3.073(0.111)</td>
<td>3.252(0.117)</td>
<td>3.031(0.094)</td>
<td>3.070(0.094)</td>
</tr>
</tbody>
</table>

### $\rho = 0.2$

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.947(0.187)</td>
<td>2.906(0.182)</td>
<td>2.936(0.184)</td>
<td>2.949(0.116)</td>
<td>2.909(0.115)</td>
<td>2.941(0.116)</td>
<td>2.947(0.099)</td>
<td>2.907(0.099)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.410(1.688)</td>
<td>7.670(1.664)</td>
<td>11.93(1.945)</td>
<td>5.790(1.175)</td>
<td>8.050(1.337)</td>
<td>12.40(1.494)</td>
<td>6.150(0.687)</td>
<td>8.460(0.909)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.995(0.186)</td>
<td>3.000(0.169)</td>
<td>3.101(0.169)</td>
<td>3.001(0.124)</td>
<td>3.006(0.112)</td>
<td>3.107(0.120)</td>
<td>2.997(0.098)</td>
<td>3.001(0.094)</td>
</tr>
<tr>
<td>REG</td>
<td>2.944(0.213)</td>
<td>2.914(0.244)</td>
<td>2.913(0.271)</td>
<td>2.940(0.155)</td>
<td>2.907(0.163)</td>
<td>2.898(0.194)</td>
<td>2.935(0.112)</td>
<td>2.892(0.124)</td>
</tr>
<tr>
<td>MIXED</td>
<td>3.012(0.186)</td>
<td>3.039(0.168)</td>
<td>3.198(0.169)</td>
<td>3.020(0.122)</td>
<td>3.049(0.110)</td>
<td>3.211(0.112)</td>
<td>3.017(0.098)</td>
<td>3.047(0.094)</td>
</tr>
</tbody>
</table>

### $\rho = 0.4$

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.921(0.195)</td>
<td>2.851(0.197)</td>
<td>2.833(0.200)</td>
<td>2.925(0.135)</td>
<td>2.856(0.137)</td>
<td>2.839(0.140)</td>
<td>2.925(0.104)</td>
<td>2.857(0.107)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.570(1.736)</td>
<td>8.010(1.396)</td>
<td>11.78(1.784)</td>
<td>6.160(1.195)</td>
<td>8.570(1.249)</td>
<td>12.26(1.461)</td>
<td>6.440(0.857)</td>
<td>9.180(0.957)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.952(0.188)</td>
<td>2.915(0.189)</td>
<td>2.958(0.180)</td>
<td>2.960(0.142)</td>
<td>2.927(0.138)</td>
<td>2.968(0.129)</td>
<td>2.959(0.102)</td>
<td>2.927(0.097)</td>
</tr>
<tr>
<td>REG</td>
<td>2.924(0.200)</td>
<td>2.857(0.216)</td>
<td>2.837(0.226)</td>
<td>2.926(0.155)</td>
<td>2.875(0.160)</td>
<td>2.836(0.170)</td>
<td>2.922(0.111)</td>
<td>2.851(0.113)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.961(0.186)</td>
<td>2.938(0.185)</td>
<td>3.008(0.173)</td>
<td>2.970(0.141)</td>
<td>2.951(0.135)</td>
<td>3.019(0.123)</td>
<td>2.969(0.102)</td>
<td>2.951(0.096)</td>
</tr>
</tbody>
</table>

1. $\mu_3 = 2.934$ for $P(T_3) = 3\%$, $P(T_4) = 5\%$, $\mu_3 = 2.875$ for $P(T_3) = 7\%$, $P(T_4) = 10\%$, $\mu_3 = 2.876$ for $P(T_3) = 15\%$, $P(T_4) = 20\%$. 
Table B.5 Mean ($\mu_4=2.6$) estimates (s.d) from different methods for linear response and MDT at last two time points

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.614(0.196)</td>
<td>2.602(0.205)</td>
<td>2.595(0.238)</td>
<td>2.603(0.140)</td>
<td>2.594(0.137)</td>
<td>2.585(0.146)</td>
<td>2.599(0.091)</td>
<td>2.590(0.090)</td>
<td>2.576(0.094)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.230(1.469)</td>
<td>7.490(1.446)</td>
<td>11.99(1.987)</td>
<td>5.660(1.130)</td>
<td>7.850(1.158)</td>
<td>12.39(1.530)</td>
<td>5.870(0.761)</td>
<td>8.220(0.799)</td>
<td>12.98(1.163)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.771(0.173)</td>
<td>2.880(0.180)</td>
<td>3.061(0.193)</td>
<td>2.759(0.117)</td>
<td>2.867(0.115)</td>
<td>3.041(0.128)</td>
<td>2.749(0.083)</td>
<td>2.859(0.083)</td>
<td>3.033(0.084)</td>
</tr>
<tr>
<td>REG</td>
<td>2.642(0.301)</td>
<td>2.689(0.439)</td>
<td>2.655(0.668)</td>
<td>2.603(0.241)</td>
<td>2.633(0.310)</td>
<td>2.594(0.440)</td>
<td>2.589(0.173)</td>
<td>2.606(0.212)</td>
<td>2.589(0.294)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.780(0.190)</td>
<td>2.907(0.198)</td>
<td>3.147(0.219)</td>
<td>2.771(0.128)</td>
<td>2.899(0.125)</td>
<td>3.129(0.140)</td>
<td>2.763(0.086)</td>
<td>2.891(0.085)</td>
<td>3.119(0.092)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\rho = 0.2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.608(0.197)</td>
<td>2.596(0.210)</td>
<td>2.587(0.222)</td>
<td>2.599(0.138)</td>
<td>2.589(0.137)</td>
<td>2.580(0.139)</td>
<td>2.597(0.089)</td>
<td>2.588(0.090)</td>
<td>2.571(0.093)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.410(1.688)</td>
<td>7.670(1.664)</td>
<td>11.93(1.945)</td>
<td>5.790(1.175)</td>
<td>8.050(1.337)</td>
<td>12.40(1.494)</td>
<td>6.150(0.687)</td>
<td>8.460(0.909)</td>
<td>13.41(1.092)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.735(0.177)</td>
<td>2.820(0.182)</td>
<td>2.975(0.167)</td>
<td>2.725(0.118)</td>
<td>2.812(0.111)</td>
<td>2.969(0.108)</td>
<td>2.720(0.079)</td>
<td>2.806(0.076)</td>
<td>2.957(0.079)</td>
</tr>
<tr>
<td>REG</td>
<td>2.560(0.285)</td>
<td>2.539(0.432)</td>
<td>2.443(0.522)</td>
<td>2.531(0.232)</td>
<td>2.510(0.281)</td>
<td>2.388(0.359)</td>
<td>2.518(0.157)</td>
<td>2.474(0.195)</td>
<td>2.377(0.250)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.771(0.184)</td>
<td>2.889(0.199)</td>
<td>3.122(0.208)</td>
<td>2.758(0.121)</td>
<td>2.882(0.121)</td>
<td>3.115(0.130)</td>
<td>2.754(0.080)</td>
<td>2.879(0.083)</td>
<td>3.105(0.085)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\rho = 0.4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.578(0.204)</td>
<td>2.551(0.210)</td>
<td>2.599(0.229)</td>
<td>2.580(0.150)</td>
<td>2.553(0.154)</td>
<td>2.512(0.160)</td>
<td>2.578(0.096)</td>
<td>2.552(0.101)</td>
<td>2.511(0.105)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.570(1.736)</td>
<td>8.010(1.396)</td>
<td>11.78(1.784)</td>
<td>6.160(1.195)</td>
<td>8.570(1.249)</td>
<td>12.26(1.461)</td>
<td>6.440(0.857)</td>
<td>9.180(0.957)</td>
<td>13.04(1.161)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.660(0.191)</td>
<td>2.700(0.187)</td>
<td>2.783(0.186)</td>
<td>2.663(0.135)</td>
<td>2.706(0.128)</td>
<td>2.788(0.122)</td>
<td>2.659(0.090)</td>
<td>2.706(0.082)</td>
<td>2.784(0.079)</td>
</tr>
<tr>
<td>REG</td>
<td>2.555(0.230)</td>
<td>2.509(0.288)</td>
<td>2.423(0.333)</td>
<td>2.546(0.186)</td>
<td>2.488(0.216)</td>
<td>2.405(0.250)</td>
<td>2.535(0.120)</td>
<td>2.473(0.145)</td>
<td>2.386(0.168)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.677(0.193)</td>
<td>2.741(0.190)</td>
<td>2.861(0.197)</td>
<td>2.680(0.137)</td>
<td>2.747(0.130)</td>
<td>2.864(0.127)</td>
<td>2.676(0.091)</td>
<td>2.747(0.084)</td>
<td>2.861(0.083)</td>
</tr>
</tbody>
</table>
Table B.6 Variance ($\sigma^2=2$) estimates (s.d) from different methods for linear response and MDT at last two time points

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>$n=50$</th>
<th></th>
<th></th>
<th>$n=100$</th>
<th></th>
<th></th>
<th>$n=200$</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>1.917(0.213)</td>
<td>1.861(0.196)</td>
<td>1.767(0.186)</td>
<td>1.916(0.128)</td>
<td>1.859(0.122)</td>
<td>1.767(0.113)</td>
<td>1.919(0.093)</td>
<td>1.861(0.091)</td>
<td>1.769(0.085)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.230(1.469)</td>
<td>7.490(1.446)</td>
<td>11.99(1.987)</td>
<td>5.660(1.130)</td>
<td>7.850(1.158)</td>
<td>12.39(1.530)</td>
<td>5.870(0.761)</td>
<td>8.220(0.799)</td>
<td>12.98(1.163)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.838(0.220)</td>
<td>1.762(0.212)</td>
<td>1.676(0.217)</td>
<td>1.837(0.133)</td>
<td>1.757(0.133)</td>
<td>1.675(0.139)</td>
<td>1.840(0.095)</td>
<td>1.761(0.098)</td>
<td>1.680(0.100)</td>
</tr>
<tr>
<td>REG</td>
<td>3.027(1.376)</td>
<td>4.453(2.431)</td>
<td>7.562(3.704)</td>
<td>3.169(1.297)</td>
<td>3.169(1.297)</td>
<td>7.548(2.633)</td>
<td>3.176(0.916)</td>
<td>4.477(1.163)</td>
<td>7.552(1.827)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.843(0.210)</td>
<td>1.763(0.203)</td>
<td>1.669(0.204)</td>
<td>1.838(0.127)</td>
<td>1.758(0.126)</td>
<td>1.667(0.124)</td>
<td>1.840(0.092)</td>
<td>1.758(0.095)</td>
<td>1.668(0.094)</td>
</tr>
<tr>
<td>$\rho = 0.2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>1.938(0.232)</td>
<td>1.901(0.226)</td>
<td>1.836(0.220)</td>
<td>1.941(0.142)</td>
<td>1.902(0.137)</td>
<td>1.834(0.133)</td>
<td>1.943(0.104)</td>
<td>1.904(0.102)</td>
<td>1.837(0.101)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.410(1.688)</td>
<td>7.670(1.664)</td>
<td>11.93(1.945)</td>
<td>5.790(1.175)</td>
<td>8.050(1.337)</td>
<td>12.40(1.494)</td>
<td>6.150(0.687)</td>
<td>8.460(0.909)</td>
<td>13.41(1.092)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.846(0.235)</td>
<td>1.769(0.240)</td>
<td>1.696(0.247)</td>
<td>1.846(0.142)</td>
<td>1.771(0.147)</td>
<td>1.698(0.153)</td>
<td>1.845(0.101)</td>
<td>1.773(0.104)</td>
<td>1.703(0.107)</td>
</tr>
<tr>
<td>REG</td>
<td>2.877(1.309)</td>
<td>3.928(2.117)</td>
<td>6.516(3.078)</td>
<td>2.984(1.089)</td>
<td>2.984(1.089)</td>
<td>6.552(1.999)</td>
<td>3.014(0.749)</td>
<td>4.162(1.059)</td>
<td>6.514(1.333)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.840(0.220)</td>
<td>1.759(0.220)</td>
<td>1.670(0.222)</td>
<td>1.838(0.134)</td>
<td>1.756(0.132)</td>
<td>1.665(0.132)</td>
<td>1.837(0.098)</td>
<td>1.756(0.099)</td>
<td>1.667(0.102)</td>
</tr>
<tr>
<td>$\rho = 0.4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.018(0.356)</td>
<td>2.053(0.354)</td>
<td>2.073(0.351)</td>
<td>2.019(0.223)</td>
<td>2.050(0.225)</td>
<td>2.068(0.220)</td>
<td>2.028(0.165)</td>
<td>2.055(0.168)</td>
<td>2.071(0.170)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.570(1.736)</td>
<td>8.010(1.396)</td>
<td>11.78(1.784)</td>
<td>6.160(1.195)</td>
<td>8.570(1.249)</td>
<td>12.26(1.461)</td>
<td>6.440(0.857)</td>
<td>9.180(0.957)</td>
<td>13.04(1.081)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.877(0.323)</td>
<td>1.828(0.331)</td>
<td>1.773(0.339)</td>
<td>1.875(0.200)</td>
<td>1.825(0.206)</td>
<td>1.770(0.206)</td>
<td>1.879(0.148)</td>
<td>1.827(0.154)</td>
<td>1.770(0.157)</td>
</tr>
<tr>
<td>REG</td>
<td>2.389(0.803)</td>
<td>2.830(1.008)</td>
<td>3.716(1.302)</td>
<td>2.924(0.704)</td>
<td>2.924(0.704)</td>
<td>3.795(0.874)</td>
<td>2.465(0.384)</td>
<td>3.000(0.493)</td>
<td>3.870(0.558)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.866(0.316)</td>
<td>1.806(0.322)</td>
<td>1.742(0.329)</td>
<td>1.862(0.197)</td>
<td>1.799(0.200)</td>
<td>1.734(0.201)</td>
<td>1.866(0.147)</td>
<td>1.800(0.152)</td>
<td>1.735(0.158)</td>
</tr>
</tbody>
</table>
Table B.7 Correlation estimates (s.d) from different methods for linear response and MDT at last two time points

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>0.204(0.092)</td>
<td>0.222(0.091)</td>
<td>0.245(0.093)</td>
<td>0.213(0.062)</td>
<td>0.228(0.062)</td>
<td>0.250(0.064)</td>
<td>0.216(0.042)</td>
<td>0.231(0.043)</td>
<td>0.253(0.043)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.230(1.469)</td>
<td>7.490(1.446)</td>
<td>11.99(1.987)</td>
<td>5.660(1.130)</td>
<td>7.850(1.158)</td>
<td>12.39(1.530)</td>
<td>5.870(0.761)</td>
<td>8.220(0.799)</td>
<td>12.98(1.163)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.202(0.088)</td>
<td>0.236(0.089)</td>
<td>0.294(0.092)</td>
<td>0.210(0.063)</td>
<td>0.239(0.064)</td>
<td>0.295(0.064)</td>
<td>0.213(0.042)</td>
<td>0.241(0.042)</td>
<td>0.301(0.040)</td>
</tr>
<tr>
<td>REG</td>
<td>0.340(0.158)</td>
<td>0.472(0.156)</td>
<td>0.623(0.109)</td>
<td>0.376(0.131)</td>
<td>0.375(0.131)</td>
<td>0.643(0.067)</td>
<td>0.394(0.097)</td>
<td>0.522(0.078)</td>
<td>0.656(0.040)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.175(0.087)</td>
<td>0.175(0.084)</td>
<td>0.164(0.092)</td>
<td>0.180(0.061)</td>
<td>0.177(0.060)</td>
<td>0.166(0.063)</td>
<td>0.182(0.041)</td>
<td>0.177(0.042)</td>
<td>0.168(0.042)</td>
</tr>
</tbody>
</table>

\[ \rho = 0.2 \]

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>0.400(0.086)</td>
<td>0.415(0.084)</td>
<td>0.435(0.085)</td>
<td>0.407(0.057)</td>
<td>0.421(0.059)</td>
<td>0.440(0.059)</td>
<td>0.411(0.040)</td>
<td>0.423(0.040)</td>
<td>0.444(0.041)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.410(1.688)</td>
<td>7.670(1.664)</td>
<td>11.93(1.945)</td>
<td>5.790(1.175)</td>
<td>8.050(1.337)</td>
<td>12.40(1.494)</td>
<td>6.150(0.687)</td>
<td>8.460(0.909)</td>
<td>13.41(1.092)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.392(0.085)</td>
<td>0.410(0.083)</td>
<td>0.451(0.085)</td>
<td>0.398(0.056)</td>
<td>0.416(0.061)</td>
<td>0.458(0.061)</td>
<td>0.400(0.039)</td>
<td>0.417(0.041)</td>
<td>0.463(0.040)</td>
</tr>
<tr>
<td>REG</td>
<td>0.481(0.122)</td>
<td>0.557(0.119)</td>
<td>0.668(0.080)</td>
<td>0.506(0.096)</td>
<td>0.506(0.096)</td>
<td>0.686(0.053)</td>
<td>0.522(0.067)</td>
<td>0.605(0.058)</td>
<td>0.695(0.032)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.368(0.084)</td>
<td>0.359(0.083)</td>
<td>0.344(0.091)</td>
<td>0.372(0.056)</td>
<td>0.363(0.059)</td>
<td>0.349(0.062)</td>
<td>0.374(0.040)</td>
<td>0.363(0.041)</td>
<td>0.351(0.042)</td>
</tr>
</tbody>
</table>

\[ \rho = 0.4 \]

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th></th>
<th></th>
<th>n = 100</th>
<th></th>
<th></th>
<th>n = 200</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>0.795(0.045)</td>
<td>0.800(0.044)</td>
<td>0.807(0.043)</td>
<td>0.799(0.030)</td>
<td>0.803(0.029)</td>
<td>0.810(0.029)</td>
<td>0.801(0.020)</td>
<td>0.806(0.020)</td>
<td>0.813(0.019)</td>
</tr>
<tr>
<td># of itera</td>
<td>5.570(1.736)</td>
<td>8.010(1.396)</td>
<td>11.78(1.784)</td>
<td>6.160(1.195)</td>
<td>8.570(1.249)</td>
<td>12.26(1.461)</td>
<td>6.440(0.857)</td>
<td>9.180(0.957)</td>
<td>13.04(1.081)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.788(0.046)</td>
<td>0.791(0.047)</td>
<td>0.801(0.046)</td>
<td>0.791(0.030)</td>
<td>0.795(0.031)</td>
<td>0.806(0.031)</td>
<td>0.794(0.020)</td>
<td>0.798(0.021)</td>
<td>0.808(0.021)</td>
</tr>
<tr>
<td>REG</td>
<td>0.799(0.045)</td>
<td>0.805(0.045)</td>
<td>0.816(0.037)</td>
<td>0.814(0.029)</td>
<td>0.814(0.029)</td>
<td>0.823(0.023)</td>
<td>0.811(0.019)</td>
<td>0.819(0.019)</td>
<td>0.826(0.014)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.780(0.047)</td>
<td>0.773(0.050)</td>
<td>0.766(0.052)</td>
<td>0.783(0.031)</td>
<td>0.777(0.033)</td>
<td>0.771(0.035)</td>
<td>0.786(0.021)</td>
<td>0.779(0.023)</td>
<td>0.773(0.025)</td>
</tr>
</tbody>
</table>
Table B.8 Mean ($\mu_4=2.6$) estimates (s.d) from different form of response function\(^1\) and 20% MDT at last time point

<table>
<thead>
<tr>
<th>Form Method</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$n = 50$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.765(0.197)</td>
<td>2.747(0.197)</td>
<td>2.747(0.197)</td>
<td>2.756(0.133)</td>
<td>2.738(0.133)</td>
<td>2.738(0.133)</td>
<td>2.750(0.089)</td>
<td>2.731(0.090)</td>
<td>2.731(0.089)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.190(1.426)</td>
<td>8.190(1.423)</td>
<td>8.190(1.426)</td>
<td>8.500(1.000)</td>
<td>8.500(1.000)</td>
<td>8.500(1.000)</td>
<td>8.920(0.774)</td>
<td>8.920(0.774)</td>
<td>8.190(1.426)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.905(0.186)</td>
<td>3.468(0.212)</td>
<td>3.178(0.185)</td>
<td>2.895(0.129)</td>
<td>3.434(0.141)</td>
<td>3.161(0.125)</td>
<td>2.888(0.088)</td>
<td>3.436(0.098)</td>
<td>3.1540.086)</td>
</tr>
<tr>
<td>REG</td>
<td>2.284(0.351)</td>
<td>3.425(0.288)</td>
<td>2.846(0.285)</td>
<td>2.280(0.258)</td>
<td>3.390.198)</td>
<td>2.828(0.206)</td>
<td>2.274(0.176)</td>
<td>3.386(0.145)</td>
<td>2.822(0.144)</td>
</tr>
<tr>
<td>MIXED</td>
<td>3.120(0.191)</td>
<td>3.102(0.191)</td>
<td>3.102(0.191)</td>
<td>3.104(0.129)</td>
<td>3.086(0.129)</td>
<td>3.084(0.129)</td>
<td>3.096(0.085)</td>
<td>3.078(0.085)</td>
<td>3.078(0.085)</td>
</tr>
<tr>
<td><strong>$n = 100$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.727(0.195)</td>
<td>2.708(0.195)</td>
<td>2.708(0.195)</td>
<td>2.718(0.131)</td>
<td>2.700(0.131)</td>
<td>2.700(0.131)</td>
<td>2.713(0.086)</td>
<td>2.695(0.086)</td>
<td>2.695(0.086)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.410(1.450)</td>
<td>8.410(1.450)</td>
<td>8.410(1.450)</td>
<td>8.820(1.067)</td>
<td>8.820(1.067)</td>
<td>8.820(1.067)</td>
<td>9.180(0.809)</td>
<td>9.180(0.809)</td>
<td>8.410(1.450)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.825(0.170)</td>
<td>3.382(0.176)</td>
<td>3.096(0.155)</td>
<td>2.817(0.124)</td>
<td>3.377(0.121)</td>
<td>3.085(0.111)</td>
<td>2.810(0.083)</td>
<td>3.358(0.078)</td>
<td>3.076(0.073)</td>
</tr>
<tr>
<td>REG</td>
<td>2.154(0.348)</td>
<td>3.283(0.211)</td>
<td>2.710(0.244)</td>
<td>2.143(0.259)</td>
<td>3.259(0.161)</td>
<td>2.692(0.188)</td>
<td>2.137(0.176)</td>
<td>3.249(0.116)</td>
<td>2.684(0.130)</td>
</tr>
<tr>
<td>MIXED</td>
<td>3.060(0.178)</td>
<td>3.042(0.178)</td>
<td>3.042(0.176)</td>
<td>3.050(0.120)</td>
<td>3.032(0.120)</td>
<td>3.031(0.120)</td>
<td>3.040(0.076)</td>
<td>3.022(0.076)</td>
<td>3.022(0.076)</td>
</tr>
<tr>
<td><strong>$n = 200$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.588(0.213)</td>
<td>2.570(0.213)</td>
<td>2.570(0.213)</td>
<td>2.589(0.150)</td>
<td>2.571(0.150)</td>
<td>2.571(0.150)</td>
<td>2.588(0.098)</td>
<td>2.570(0.098)</td>
<td>2.570(0.098)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.850(1.403)</td>
<td>8.850(1.403)</td>
<td>8.850(1.403)</td>
<td>9.310(1.152)</td>
<td>9.310(1.152)</td>
<td>9.310(1.152)</td>
<td>9.930(0.868)</td>
<td>9.930(0.868)</td>
<td>8.850(1.403)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.660(0.197)</td>
<td>3.192(0.183)</td>
<td>2.918(0.175)</td>
<td>2.661(0.142)</td>
<td>3.187(0.120)</td>
<td>2.916(0.118)</td>
<td>2.655(0.091)</td>
<td>3.191(0.075)</td>
<td>2.915(0.073)</td>
</tr>
<tr>
<td>REG</td>
<td>2.041(0.324)</td>
<td>3.121(0.184)</td>
<td>2.573(0.218)</td>
<td>2.046(0.262)</td>
<td>3.112(0.140)</td>
<td>2.570(0.176)</td>
<td>2.027(0.171)</td>
<td>3.113(0.090)</td>
<td>2.561(0.111)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.806(0.192)</td>
<td>2.788(0.193)</td>
<td>2.788(0.192)</td>
<td>2.805(0.131)</td>
<td>2.788(0.132)</td>
<td>2.787(0.131)</td>
<td>2.802(0.084)</td>
<td>2.784(0.083)</td>
<td>2.784(0.083)</td>
</tr>
</tbody>
</table>

\(^{1}\text{Convex response function: } Y = 5 - 1.333t + 0.156t^2; \text{ Concave response function: } Y = 5 + 0.5336t - 0.1556t^2; \text{ Linear response function: } Y = 5 - 0.4t ,
<table>
<thead>
<tr>
<th>Form Method</th>
<th>n = 50</th>
<th>n = 100</th>
<th>n = 200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Convex</td>
<td>Concave</td>
<td>Linear</td>
</tr>
<tr>
<td>MDT</td>
<td>1.864(0.210)</td>
<td>1.864(0.210)</td>
<td>1.864(0.210)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.190(1.426)</td>
<td>8.190(1.426)</td>
<td>8.190(1.426)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.879(0.241)</td>
<td>1.945(0.206)</td>
<td>1.828(0.214)</td>
</tr>
<tr>
<td>REG</td>
<td>2.974(0.679)</td>
<td>2.322(0.369)</td>
<td>2.309(0.420)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.824(0.218)</td>
<td>1.824(0.218)</td>
<td>1.824(0.218)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>1.880(0.230)</td>
<td>1.880(0.230)</td>
<td>1.880(0.230)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.410(1.450)</td>
<td>8.410(1.450)</td>
<td>8.410(1.450)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.898(0.264)</td>
<td>1.877(0.210)</td>
<td>1.799(0.227)</td>
</tr>
<tr>
<td>REG</td>
<td>3.206(0.723)</td>
<td>2.188(0.332)</td>
<td>2.339(0.438)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.813(0.231)</td>
<td>1.813(0.231)</td>
<td>1.813(0.231)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.004(0.340)</td>
<td>2.004(0.340)</td>
<td>2.004(0.340)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.850(1.403)</td>
<td>8.850(1.403)</td>
<td>8.850(1.403)</td>
</tr>
<tr>
<td>LOCF</td>
<td>1.966(0.354)</td>
<td>1.717(0.274)</td>
<td>1.720(0.301)</td>
</tr>
<tr>
<td>REG</td>
<td>3.478(0.695)</td>
<td>1.828(0.306)</td>
<td>2.209(0.417)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.832(0.327)</td>
<td>1.832(0.327)</td>
<td>1.832(0.327)</td>
</tr>
</tbody>
</table>

1Convex response function: \( Y = 5 - 1.333t + 0.156t^2 \); Concave response function: \( Y = 5 + 0.5336t - 0.1556t^2 \); Linear response function: \( Y = 5 - 0.4t \),
### Table B.10 Correlation estimates (s.d) from different form of response function and 20% MDT at last time point

<table>
<thead>
<tr>
<th>Form Method</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 50</td>
<td>n = 100</td>
<td>n = 200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>0.186(0.089)</td>
<td>0.186(0.089)</td>
<td>0.221(0.081)</td>
<td>0.192(0.060)</td>
<td>0.192(0.060)</td>
<td>0.224(0.058)</td>
<td>0.195(0.040)</td>
<td>0.195(0.040)</td>
<td>0.229(0.037)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.190(1.426)</td>
<td>8.190(1.423)</td>
<td>8.190(1.426)</td>
<td>8.500(1.000)</td>
<td>8.500(1.000)</td>
<td>8.500(1.000)</td>
<td>8.920(0.774)</td>
<td>8.920(0.774)</td>
<td>8.190(1.426)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.238(0.086)</td>
<td>0.199(0.077)</td>
<td>0.283(0.087)</td>
<td>0.242(0.060)</td>
<td>0.201(0.056)</td>
<td>0.286(0.063)</td>
<td>0.246(0.038)</td>
<td>0.205(0.037)</td>
<td>0.293(0.040)</td>
</tr>
<tr>
<td>REG</td>
<td>0.285(0.091)</td>
<td>0.249(0.078)</td>
<td>0.171(0.085)</td>
<td>0.290(0.064)</td>
<td>0.250(0.059)</td>
<td>0.176(0.058)</td>
<td>0.298(0.042)</td>
<td>0.256(0.038)</td>
<td>0.179(0.039)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.171(0.085)</td>
<td>0.171(0.085)</td>
<td>0.186(0.089)</td>
<td>0.176(0.059)</td>
<td>0.176(0.058)</td>
<td>0.192(0.133)</td>
<td>0.179(0.039)</td>
<td>0.179(0.039)</td>
<td>0.195(0.040)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Form Method</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 50</td>
<td>n = 100</td>
<td>n = 200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>0.384(0.084)</td>
<td>0.384(0.084)</td>
<td>0.389(0.080)</td>
<td>0.390(0.056)</td>
<td>0.390(0.056)</td>
<td>0.395(0.056)</td>
<td>0.394(0.038)</td>
<td>0.394(0.038)</td>
<td>0.399(0.036)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.410(1.450)</td>
<td>8.410(1.450)</td>
<td>8.410(1.450)</td>
<td>8.820(1.067)</td>
<td>8.820(1.067)</td>
<td>8.820(1.067)</td>
<td>9.180(0.809)</td>
<td>9.180(0.809)</td>
<td>8.410(1.450)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.419(0.083)</td>
<td>0.347(0.077)</td>
<td>0.448(0.081)</td>
<td>0.427(0.057)</td>
<td>0.353(0.055)</td>
<td>0.457(0.056)</td>
<td>0.430(0.036)</td>
<td>0.355(0.036)</td>
<td>0.461(0.035)</td>
</tr>
<tr>
<td>REG</td>
<td>0.453(0.080)</td>
<td>0.391(0.078)</td>
<td>0.358(0.083)</td>
<td>0.462(0.055)</td>
<td>0.398(0.054)</td>
<td>0.363(0.057)</td>
<td>0.467(0.034)</td>
<td>0.401(0.034)</td>
<td>0.366(0.038)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.358(0.083)</td>
<td>0.358(0.083)</td>
<td>0.384(0.084)</td>
<td>0.363(0.057)</td>
<td>0.363(0.057)</td>
<td>0.390(0.056)</td>
<td>0.366(0.038)</td>
<td>0.366(0.038)</td>
<td>0.394(0.038)</td>
</tr>
</tbody>
</table>

\( \rho = 0.2 \)

<table>
<thead>
<tr>
<th>Form Method</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
<th>Convex</th>
<th>Concave</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 50</td>
<td>n = 100</td>
<td>n = 200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>0.792(0.046)</td>
<td>0.792(0.046)</td>
<td>0.760(0.048)</td>
<td>0.796(0.030)</td>
<td>0.796(0.029)</td>
<td>0.765(0.031)</td>
<td>0.799(0.020)</td>
<td>0.799(0.020)</td>
<td>0.766(0.021)</td>
</tr>
<tr>
<td># of itera</td>
<td>8.950(1.403)</td>
<td>8.950(1.403)</td>
<td>8.950(1.403)</td>
<td>9.310(1.152)</td>
<td>9.310(1.152)</td>
<td>9.310(1.152)</td>
<td>9.930(0.868)</td>
<td>9.930(0.868)</td>
<td>8.850(1.403)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.802(0.044)</td>
<td>0.685(0.055)</td>
<td>0.797(0.043)</td>
<td>0.806(0.029)</td>
<td>0.690(0.036)</td>
<td>0.802(0.028)</td>
<td>0.808(0.020)</td>
<td>0.689(0.025)</td>
<td>0.805(0.018)</td>
</tr>
<tr>
<td>REG</td>
<td>0.776(0.042)</td>
<td>0.701(0.054)</td>
<td>0.774(0.049)</td>
<td>0.780(0.026)</td>
<td>0.708(0.034)</td>
<td>0.778(0.032)</td>
<td>0.782(0.017)</td>
<td>0.709(0.023)</td>
<td>0.780(0.022)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.774(0.049)</td>
<td>0.774(0.049)</td>
<td>0.792(0.046)</td>
<td>0.778(0.032)</td>
<td>0.778(0.032)</td>
<td>0.796(0.029)</td>
<td>0.780(0.022)</td>
<td>0.780(0.151)</td>
<td>0.799(0.020)</td>
</tr>
</tbody>
</table>

\( \rho = 0.8 \)

1 Convex response function: \( Y = 5 - 1.333t + 0.156t^2 \); Concave response function; \( Y = 5 + 0.5336t - 0.1556t^2 \); Linear response function: \( Y = 5 - 0.4t \).
Table B.11 Mean ($\mu_4=2.6$) estimates (s.d) from linear response function and MAR at last time point

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th>n = 100</th>
<th>n = 200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>2.498(0.208)</td>
<td>2.414(0.216)</td>
<td>2.252(0.228)</td>
</tr>
<tr>
<td># of itera</td>
<td>3.930(1.166)</td>
<td>5.330(1.240)</td>
<td>7.820(1.559)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.671(0.202)</td>
<td>2.727(0.194)</td>
<td>2.825(0.211)</td>
</tr>
<tr>
<td>REG</td>
<td>2.605(0.219)</td>
<td>2.600(0.207)</td>
<td>2.560(0.253)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.607(0.200)</td>
<td>2.606(0.198)</td>
<td>2.598(0.211)</td>
</tr>
<tr>
<td>MDT</td>
<td>2.500(0.209)</td>
<td>2.405(0.225)</td>
<td>2.246(0.239)</td>
</tr>
<tr>
<td># of itera</td>
<td>3.930(1.225)</td>
<td>5.410(1.164)</td>
<td>8.010(1.453)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.666(0.216)</td>
<td>2.727(0.213)</td>
<td>2.846(0.205)</td>
</tr>
<tr>
<td>REG</td>
<td>2.586(0.217)</td>
<td>2.592(0.241)</td>
<td>2.606(0.276)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.608(0.212)</td>
<td>2.608(0.213)</td>
<td>2.616(0.232)</td>
</tr>
</tbody>
</table>

\[ \rho = 0.2 \]

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th>n = 100</th>
<th>n = 200</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDT</td>
<td>2.440(0.218)</td>
<td>2.326(0.225)</td>
<td>2.140(0.224)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.090(1.264)</td>
<td>5.210(1.140)</td>
<td>7.310(1.195)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.659(0.198)</td>
<td>2.727(0.204)</td>
<td>2.845(0.208)</td>
</tr>
<tr>
<td>REG</td>
<td>2.593(0.202)</td>
<td>2.592(0.209)</td>
<td>2.596(0.223)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.595(0.198)</td>
<td>2.598(0.203)</td>
<td>2.595(0.206)</td>
</tr>
</tbody>
</table>

\[ \rho = 0.8 \]
<table>
<thead>
<tr>
<th>Method</th>
<th>n = 50</th>
<th>n = 100</th>
<th>n = 200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td><strong>ρ = 0.2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.043(0.248)</td>
<td>2.048(0.239)</td>
<td>2.038(0.232)</td>
</tr>
<tr>
<td># of itera</td>
<td>3.930(1.166)</td>
<td>5.330(1.240)</td>
<td>7.820(1.559)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.033(0.256)</td>
<td>2.046(0.261)</td>
<td>2.063(0.262)</td>
</tr>
<tr>
<td>REG</td>
<td>2.117(0.314)</td>
<td>2.217(0.383)</td>
<td>2.440(0.439)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.013(0.249)</td>
<td>2.012(0.245)</td>
<td>2.011(0.245)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ρ = 0.4</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.044(0.272)</td>
<td>2.054(0.269)</td>
<td>2.041(0.262)</td>
</tr>
<tr>
<td># of itera</td>
<td>3.930(1.225)</td>
<td>5.410(1.164)</td>
<td>8.010(1.453)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.016(0.264)</td>
<td>2.040(0.278)</td>
<td>2.076(0.285)</td>
</tr>
<tr>
<td>REG</td>
<td>2.080(0.288)</td>
<td>2.201(0.328)</td>
<td>2.385(0.383)</td>
</tr>
<tr>
<td>MIXED</td>
<td>2.004(0.269)</td>
<td>2.006(0.274)</td>
<td>2.006(0.276)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ρ = 0.8</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>2.102(0.367)</td>
<td>2.133(0.366)</td>
<td>2.123(0.363)</td>
</tr>
<tr>
<td># of itera</td>
<td>4.090(1.264)</td>
<td>5.210(1.140)</td>
<td>7.310(1.195)</td>
</tr>
<tr>
<td>LOCF</td>
<td>2.018(0.351)</td>
<td>2.044(0.349)</td>
<td>2.081(0.352)</td>
</tr>
<tr>
<td>REG</td>
<td>2.037(0.359)</td>
<td>2.101(0.366)</td>
<td>2.199(0.378)</td>
</tr>
<tr>
<td>MIXED</td>
<td>1.988(0.350)</td>
<td>1.990(0.349)</td>
<td>1.989(0.352)</td>
</tr>
</tbody>
</table>
Table B.13 Correlation estimates (s.d) from linear response function and MAR at last time point

<table>
<thead>
<tr>
<th>Missing Method</th>
<th>n = 50</th>
<th>n = 100</th>
<th>n = 200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5%</td>
<td>10%</td>
<td>20%</td>
</tr>
<tr>
<td>MDT</td>
<td>0.184(0.092)</td>
<td>0.182(0.092)</td>
<td>0.185(0.092)</td>
</tr>
<tr>
<td># of iter</td>
<td>3.930(1.166)</td>
<td>5.330(1.240)</td>
<td>7.820(1.559)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.202(0.090)</td>
<td>0.213(0.090)</td>
<td>0.238(0.084)</td>
</tr>
<tr>
<td>REG</td>
<td>0.214(0.093)</td>
<td>0.237(0.093)</td>
<td>0.288(0.085)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.188(0.091)</td>
<td>0.188(0.091)</td>
<td>0.189(0.091)</td>
</tr>
<tr>
<td></td>
<td>ρ = 0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDT</td>
<td>0.380(0.085)</td>
<td>0.375(0.082)</td>
<td>0.364(0.086)</td>
</tr>
<tr>
<td># of iter</td>
<td>3.930(1.225)</td>
<td>5.410(1.164)</td>
<td>8.010(1.453)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.392(0.082)</td>
<td>0.403(0.083)</td>
<td>0.424(0.079)</td>
</tr>
<tr>
<td>REG</td>
<td>0.400(0.082)</td>
<td>0.421(0.084)</td>
<td>0.452(0.084)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.385(0.084)</td>
<td>0.385(0.084)</td>
<td>0.383(0.086)</td>
</tr>
<tr>
<td>MDT</td>
<td>0.760(0.053)</td>
<td>0.745(0.053)</td>
<td>0.725(0.056)</td>
</tr>
<tr>
<td># of iter</td>
<td>4.090(1.264)</td>
<td>5.210(1.140)</td>
<td>7.310(1.195)</td>
</tr>
<tr>
<td>LOCF</td>
<td>0.790(0.044)</td>
<td>0.792(0.043)</td>
<td>0.793(0.042)</td>
</tr>
<tr>
<td>REG</td>
<td>0.792(0.044)</td>
<td>0.796(0.043)</td>
<td>0.800(0.041)</td>
</tr>
<tr>
<td>MIXED</td>
<td>0.790(0.045)</td>
<td>0.790(0.045)</td>
<td>0.790(0.046)</td>
</tr>
</tbody>
</table>
Figure B.1 Mean ($\mu_4$) estimates from different methods for linear response, MDT at last time point and AR(1)=0.2, 0.4 and 0.8.
Figure B.2 Variance estimates from different methods for linear response, MDT at last time point and AR(1)=0.2, 0.4 and 0.8.
Figure B.3 Correlation estimates from different methods for linear response, MDT at last time point and AR(1)=0.2, 0.4 and 0.8.
Figure B.4 Mean($\mu_3$) estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8.
Figure B.5 Mean($\mu_4$) estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8.
Figure B.6 Variance estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8.
Figure B.7 Correlation estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8.
Figure B.8 Variance estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8. Enlarge the left corners of plots in Figure B.6.
Figure B.9 Correlation estimates from different methods for linear response, MDT at last two time points and AR(1)=0.2, 0.4 and 0.8. Enlarge the left corners of plots in Figure B.7.
Figure B.10 Mean ($\mu_4$) estimates from different forms of response function, 20% MDT at last time point and AR(1)=0.2, 0.4 and 0.8.
Figure B.11 Variance estimates from different forms of response function, 20% MDT at last time point and AR(1)=0.2, 0.4 and 0.8.
Figure B.12 Correlation estimates from different forms of response function, 20% MDT at last time point, and AR(1)=0.2, 0.4 and 0.8.
Figure B.13 Mean ($\mu_4$) estimates from different methods with MAR at last time point and AR(1)=0.2, 0.4 and 0.8.
Figure B.14 Variance estimates from different methods with 20% MAR at last time point and AR(1)=0.2, 0.4 and 0.8.
Figure B.15 Correlation estimates from different methods with 20% MAR at last time point and AR(1)=0.2, 0.4 and 0.8.
APPENDIX C

Programs for Application of MDT Method to IMPS data
APPENDIX C

Programs for Application of MDT Method to IMPS data

C.1 program for iterating between E-step and M-step

******************************************************************************

This code is written for IMPS dataset. However, it could be easily adjusted to the data set with more than two groups and/or more than four repeated measures.

Input data sets:
impute.univarpt: univariate format of IMPS data of two-groups clinical trial study
4 variables are contained: id, time, imps, group
impute.imps0: multivariate format of IMPS data with placebo group only.
sorted by "flag" indicating missing mechanism (NM MAR Failure Improve). Six variables are contained: id time1-time&_t and flag
impute.imps1: multivariate format for treatment group
sorted by "flag" indicating missing mechanism (NM MAR Failure Improve). Six variables are contained: id time1-time&_t and flag

Define macro variables:
libname: define the location of input datasets
data: univariate format of IMPS data
data0 multivariate format for placebo group
data1 multivariate format for treatment group
&_pi=constant('pi'): define constant π
&_n: define total sample size
&_t: define number of repeated measures
&diff: define the convergence criterion

missup03-missup0&_t: define number of missing values at each time point for placebo group at higher tail
missup13-missup1&_t: define number of missing values at each time point for treatment group at higher tail
misslw03-missup0&_t: define number of missing values at each time point for placebo group at lower tail
Output data sets
work.imps11: final imputation for placebo group (multivariate format)
work.imps10: final imputation for treatment group (multivariate format)
work.complete0 final imputation for placebo group (univariate format)
work.complete1 final imputation for treatment group (univariate format)
work.complete final imputation with two groups (univariate format)

LIBNAME IMPUTE "E:\IMPUTATION";
FILENAME MYMACRO "E:\IMPUTATION"; /*AUTOCALL MACRO LIBRARY */
OPTIONS MPRINT MLOGIC SYMBOLGEN YEARCUTOFF=1950;
OPTIONS MAUTOSOURCE
SASAUTOS=(MYMACRO,SASAUTOS);
/*PROC PRINTTO LOG='E:/IMPUTATION/IMPUTE FINAL/LOG.TXT' NEW;RUN;*/
PROC PRINTTO LOG=LOG;RUN;
TITLE;
%LET _PI=CONSTANT('PI');
%GLOBAL _PI;

DATA UNIVARPT;
SET IMPUTE.UNIVARPT;
RESPONSE=IMPS;
DROP IMPS;
RUN;

PROC PRINT DATA=IMPUTE.UNIVARPT;RUN;
PROC PRINT DATA=IMPUTE.IMPS0;RUN;
PROC PRINT DATA=IMPUTE.IMPS1;RUN;

MODIFY A STYLE TEMPLATE

PROC TEMPLATE;
DEFINE STYLE STYLES.NEWRTF;
PARENT = STYLES.RTF;
REPLACE COLOR_LIST /
"BG"=WHITE
"FG"=BLACK
"BGH"=WHITE
"LW"=BLUE;
END;
RUN;

OPEN AN RTF FILE
%MACRO EMALGORITHM/ParmBuff;
/*DEFINE MACRO VARIABLES FORM &SYSPBUFF*/
%PUT SYSPBUFF CONTAINS: &SYSPBUFF;
%LET I=%SCAN(&SYSPBUFF,1);
%LET DIFF=%SCAN(&SYSPBUFF,2);
%LET _T=%SCAN(&SYSPBUFF,3);
%LET N=%SCAN(&SYSPBUFF,4);
%LET LIBNAME=%SCAN(&SYSPBUFF,5);
%LET DATA=%SCAN(&SYSPBUFF,6);
%LET DATA0=%SCAN(&SYSPBUFF,7);
%LET DATA1=%SCAN(&SYSPBUFF,8);
%LET Y=%SCAN(&SYSPBUFF,9);
/*%LET TYPE=%QSCAN(&SYSPBUFF,10);*/
%LET NUM=10;
%LET DSNAME=%SCAN(&SYSPBUFF,%EVAL(&NUM));
%LET T=3;
%DO %WHILE(&DSNAME NE);
%LET MISSUPO%EVAL(&T)=%SCAN(&SYSPBUFF,%EVAL(&NUM));
%LET MISSUP1%EVAL(&T)=%SCAN(&SYSPBUFF,%EVAL(&NUM+1));
%LET MISSLWO%EVAL(&T)=%SCAN(&SYSPBUFF,%EVAL(&NUM+2));
%LET MISSLW1%EVAL(&T)=%SCAN(&SYSPBUFF,%EVAL(&NUM+3));
%LET NUM=%EVAL(&NUM+4);
%LET T=%EVAL(&T+1);
%LET DSNAME=%SCAN(&SYSPBUFF, &NUM);
%END;

/*INITIAL INPUTATION*/
DATA Work(DROP=&Y);
SET &DATA;
TIME1=TIME;
*RESPONSE=&Y;
RUN;

PROC PRINT DATA=&DATA;RUN;
%MSTEP(Work); /*FIT REPEATED MEASURE MODEL AND GET THE MEAN AND VARIANCE-COVARIANCE STRUCTURE*/

DATA IMPS10(RENAME=(ID=TIME0 WEEK0=TIME1 WEEK1=TIME2 WEEK3=TIME3 WEEK6=TIME4)); /*NEED TO BE ADJUSTED ACCORDING TO # OF TIME POINTS*/
SET IMPUTE.&DATA0; /*UNIVARIATE FOR PLACEBO GROUP*/

DATA IMPS11(RENAME=(ID=TIME0 WEEK0=TIME1 WEEK1=TIME2 WEEK3=TIME3 WEEK6=TIME4));
SET IMPUTE.&DATA1;
RUN; /*UNIVARIATE FOR TREATMENT GROUP*/

%ESTEP; /*IMPUTE THE MISSING VALUES*/

TITLE ",&I TH IMPUTATION FOR PLACEBO GROUP ";
PROC PRINT DATA=IMPS10 (OBS=6); RUN;
TITLE "&I TH IMPUTATION FOR TREATMENT GROUP ";
PROC PRINT DATA=IMPS11 (OBS=10); RUN;

TITLE;
/*INTERATIVE IMPUTATION UNTILL CONVERGENCE*/
%MACRO SIMULATION;
%DO %WHILE ((&I<=30) AND (%SYSEVALF(&DIFF>10E-5)=1));
   %LET J=&I;
   %LET I=%EVAL(&I+1);
   %DO G=0 %TO 1;
   PROC IML;
     USE IMPS1&G;
     READ ALL VAR(”TIME0”:”TIME&_T”) INTO IMPS1;
     CLOSE;
     TIME={0,1,3,6};
     NRIMPS=NROW(IMPS1);
     NRC=NRIMPS*%EVAL(&-T);
     COMPLETE=J(NRC,4,0);
     DO I=1 TO NRIMPS;
        DO J=1 TO &-T;
          H=(I-l)*%EVAL(&-T)+J;
          K=J+l;
          COMPLETE[H,l]=IMPS1[I,l];
          COMPLETE[H,2]=TIME[J,l];
          COMPLETE[H,3]=IMPS1[I,K];
          COMPLETE[H,4]=%EVAL(&G);
         END;
     END;
     CREATE COMPLETE&G VAR{ID TIME RESPONSE GROUP};
     APPEND FROM COMPLETE;
     QUIT;
   %END;
   DATA COMPLETE;
   SET COMPLETE0 COMPLETE1 ;
   TIME1=TIME;
   RUN;
   TITLE "&J TH IMPUTED PROC MIXED PROCEDURE FOR &G GROUP ";
%MSTEP(COMPLETE);
%ESTEP;

PROC IML; *MSTEP;
   %DO L=0 %TO 1;
   %DO R=0 %TO 1;
      USE IMPS&L&R;
      READ ALL VAR("TIME0":"TIME&_T") INTO IMPS&L&R;
      CLOSE;
   %END;
%END;
/* PRINT IMPS01 IMPS11 IMPS00 IMPS10; */
DIFF0=ABS(IMPS11-IMPS01);
DIFF1=ABS(IMPS10-IMPS00);
DIFF=0;
DO D=2 TO %EVAL(&_T+1);
  DIFF=DIFF0[+,D]+DIFF1[+,D]+DIFF;
END;
CREATE DIFF VAR{DIFF};
APPEND FROM DIFF;
TITLE "DIFF BETWEEN &I AND &J TH IMPUTATION";
PRINT DIFF;
QUIT;
DATA DIFF;
SET DIFF;
CALL SYMPUT("DIFF", DIFF);
STOP;
RUN;
TITLE "&I TH IMPUTATION FOR PLACEBO GROUP ";
PROC PRINT DATA=IMPS10 (OBS=6); RUN;
TITLE "&I TH IMPUTATION FOR TREATMENT GROUP ";
PROC PRINT DATA=IMPS11 (OBS=10); RUN;
%MEND;
%MEND SIMULATION;
%MEND;
%MEND SIMULATION
RUN;
%MEND;
%EMALGORITHM(1,100,4,95,IMPUTE,UNIVARPT,IMPS0,IMPS1,IMPS,5,1,0, 0,5,2,0,1)

DATA COMPLETE;
SET COMPLETE;
IF -1<RESPONSE<1 THEN RESPONSE=1;
IF RESPONSE>7 THEN RESPONSE=7;
STIME=SQRT(TIME);
RUN;

PROC PRINT DATA=COMPLETE;
WHERE -1<RESPONSE<=1 OR RESPONSE>=7;
RUN;

PROC PRINT DATA=COMPLETE;RUN;

PROC MIXED DATA=COMPLETE;
CLASS ID TIME1 GROUP;
MODEL RESPONSE=GROUP GROUP*STIME STIME/DDF=.254.254 S RESIDUAL
OUTPM=PREDICTEDS;
REPEATED TIME1/TYPE=AR(1) SUBJECT=ID GROUP=GROUP;
RUN;

ODS RTF CLOSE;
C.2 Program for E-Step of EM algorithm to estimate the MDT

******************************

Input data sets:

imps1: multivariate format for treatment group
    variables: id time1-time&_t group
    group is categorical variable, others are numerical variables

imps0: multivariate format for placebo group
    variables: id time1-time&_t group
    group is categorical variable, others are numerical variables

mean: mean estimates for treatment and placebo group

r1: variance-covariance estimates for treatment group

r0: variance-covariance estimates for placebo group

Define macro variables:

missup03-missup0&_t: define number of missing values at each time point for placebo group at higher tail

missup13-missup1&_t: define number of missing values at each time point for treatment group at higher tail

misslw03-misslw0&_t: define number of missing values at each time point for placebo group at lower tail

misslw13-misslw1&_t: define number of missing values at each time point for treatment group at lower tail

_&_t: define number of repeated measures

Output data sets:

imps10: complete data for placebo group

imps11: complete data for treatment group

*MACRO ESTEP;
*CREATE ESTEP MACRO TO IMPUTE MISSING VALUES ACCORDING TO MEAN AND VARIANCE;
%DO G=0 %TO 1;
    PROC IML;
    USE MEAN;
    READ ALL VAR{MEAN} WHERE(GROUP=&G)INTO MEAN;
    MEAN=MEAN';
    CLOSE;
    USE IMPS1&G;
    READ ALL VAR("TIME0":"TIME&_T") INTO IMPS1;
    CLOSE;
IF &I>1 THEN IMPS0=IMPS1;
N&G=NROW(IMPS1);

******CREATE VAR-COV FOR CONDITIONAL MNORMAL****;
%DO H=2 %TO %EVAL(&_T-1); **IMPUTE FROM 3RD TIMEPOINT;
USE R&G;
READ ALL VAR _ALL_ INTO COV;
CLOSE;
K=%EVAL(&H);
VAR&H=COV[1:K,1:K];
VAR%EVAL(&H+1)=COV[K+1,K+1];
SD1%EVAL(&H+1)=SQRT(VAR%EVAL(&H+1));
COV%EVAL(&H+1)=COV[K+1,1:K];
SIGMA%EVAL(&H+1)=VAR%EVAL(&H+1)-
COV%EVAL(&H+1)\*INV(VAR1&H)*COV%EVAL(&H+1)\';
SD%EVAL(&H+1)=SQRT(SIGMA%EVAL(&H+1));
**VARIANCES OF CONDITIONAL NORMAL DISTRIBUTION;
%LET MISS=&G%EVAL(&H+1);
MU%EVAL(&H+1)=J(%EVAL(&&MISSUP&MISS+&&MISSLW&MISS),1,0);
IF %EVAL(&&MISSUP&MISS)>0 THEN DO;
MUP%EVAL(&H+1)=MEAN[1,%EVAL(&H+1)]+SQRT(VAR%EVAL(&H+1))\*PROBIT(1-
&&MISSUP&MISS/N&G);
***MEAN[1,4]FROM ELSEWHERE;
**THRESHOLDS AT &H+ TIMEPT;
TITLE1 " &I TH IMPUTATION &G GROUP TRUNCATED PTS AND IMPUTED DATA";
PRINT MUP%EVAL(&H+1);
DO l=1 TO &MISSUP&MISS;
MV%EVAL&H+1)[I,1]=MEAN[1,%EVAL(&H+1)]+COV%EVAL(&H+1)\*INV(VAR1&H)*(I
MPS1[1,2:K+1]-MEAN[1,1:K]');
**&H+1 MEAN OF NORMAL DISTRIBUTION CONDITIONING ON THE PREVIOUS TIMEPOINT;
A=(MUP%EVAL(&H+1)- MU%EVAL(&H+1)[I,1])/SD%EVAL(&H+1);
PDF=1/(SQRT(2*PI)*SD%EVAL(&H+1))\*EXP(-A**2/2);
IMPS1[1,\%EVAL(&H+2)]=MU%EVAL(&H+1)[I,1]+SD%EVAL(&H+1)*PDF/(1-
PROBNORM(A));
**IMPUTED TRUNCATED NORMAL DISTRIBUTION;
END;

IF &MISSW&MISS>0 THEN DO;
MLW%EVAL(&H+1)=MEAN[1,%EVAL(&H+1)]+SQRT(VAR%EVAL(&H+1))\*PROBIT(&M
ISSW&MISS/N&G); ****MEAN[1,4]FROM ELSEWHERE;
**THRESHOLDS AT &H+ TIMEPT;
TITLE1 " &I TH IMPUTATION &G GROUP TRUNCATED PTS AND IMPUTED DATA";
PRINT MLW%EVAL(&H+1);
DO l=%EVAL(&MISSUP&MISS+1) TO %EVAL(&MISSUP&MISS+&MISSW&MISS);
MU%EVAL(&H+1)[I,1]=MEAN[1,%EVAL(&H+1)]+COV%EVAL(&H+1)\*INV(VAR1&H)*(I
MPS1[1,2:K+1]-MEAN[1,1:K]');
**&H+1 MEAN OF NORMAL DISTRIBUTION CONDITIONING ON THE PREVIOUS TIMEPOINT;
A=(MLW%EVAL(&H+1)- MU%EVAL(&H+1)[I,1])/SD%EVAL(&H+1);
PDF=1/(SQRT(2*PI)*SD%EVAL(&H+1))*EXP(-A**2/2);
IMPS1[1,%EVAL(&H+2)]=MU%EVAL(&H+1)(1,1)-SD%EVAL(&H+1)*PDF/PROBNORM(A);
  **IMPUTED TRUNCATED NORMAL DISTRIBUTION;
END;
END;
%END;
CREATE IMPS1&G VAR("TIME0":"TIME&_T"); **TIME0 IS ID VAR;
APPEND FROM IMPS1;
IF &I>1 THEN DO;
CREATE IMPS0&G VAR("TIME0":"TIME&_T"); **TIME0 IS ID VAR;
APPEND FROM IMPS0;
END;
QUIT;
%END;
%MEND;

C.3 Program for M-Step of EM algorithm

******************************************************************************
Model statement and variance-covariance structure can be easily modified in PROC MIXED to fit different repeated measures model.

Define macro variables:
&mixed define the input data set

Output data set
mean mean estimates for treatment and placebo group
r1: variance-covariance estimates for treatment group
r0: variance-covariance estimates for placebo group
******************************************************************************

%MACRO MSTEP(MIX);
ODS SELECT NONE;
DATA &MIX;
SET &MIX;
STIME=SQRRT(TIME);
RUN;
PROC MIXED DATA=&MIX;
  CLASS ID TIME1 GROUP;
  MODEL RESPONSE=GROUP GROUP*STIME STIME/ OUTPM=PREDICTED S;
  REPEATED TIME1/TYPE=AR(1) SUBJECT=ID GROUP=GROUP;
  ODS OUTPUT COVPARMS=COVPARMS;
RUN;
PROC FREQ DATA=PREDICTED;
   BY GROUP;
   TABLE PRED*TIME ;
RUN;

*ODS SELECT NONE;
PROC TABULATE DATA=PREDICTED;
   CLASS GROUP TIME;
   VAR PRED;
   TABLE TIME, GROUP*(PRED N);
   ODS OUTPUT TABLE=TABLE;
RUN;

DATA MEAN(DROP=_TYPE_ _PAGE_ _TABLE_ PRED_SUM N);
   SET TABLE;
   MEAN= PRED_SUM/N;
RUN;
PROC SORT DATA=MEAN;
   BY GROUP TIME;
RUN;
ODS SELECT ALL;

PROC IML;
 USE COVPARMS;
 READ ALL INTO COV;
 %DO G=0 %TO 1;
   COV&G=J(4,4,0);
   DO I=1 TO 4;
     DO J=1 TO 4;
       IF &G=1 THEN COV&G[I,J]=COV[3,]*COV[4,]**ABS(J-I);
       ELSE IF &G=0 THEN COV&G[I,J]=COV[1,]*COV[2,]**ABS(J-I);
     END;
   END;
 PRINT COV&G;
 CREATE R&G VAR("VAR1":"VAR4");
 APPEND FROM COV&G;
 %END;
 QUIT;

%MEND;
APPENDIX D

Programs for Simulation Study of MDT Method and Other Methods Compared
APPENDIX D

Programs for Simulation Study of MDT Method and Some Other Relevant Methods

D.1 Program for missing due to truncation method

*****************************************************************************
This SAS code performs the simulation study for MDT method und MAR assumption.
The cases for MDT assumption and other simulation conditions in this dissertation were
very similarly programmed.
*****************************************************************************

LIBNAME IMPUTE "F:\ONE MISSING";
FILENAME MYMACRO "F:\ONE MISSING"; /*AUTOCALL MACRO LIBRARY */
    /*OPTIONS MPRINT MLOGIC SYMBOLGEN YEARCUTOFF=1950; */
    OPTIONS MAUTOSOURCE SASAUTOS=(MYMACRO,SASAUTOS);
TITLE;
ODS RTF BODY="F:\ONE MISSING\5MISSING\100\MDT221000554100.RTF" BODYTITLE
STYLE=NEWRTF STARTPAGE=NO;
    /*FILENAME: VARIANCE COVARINCE SAMPLESIZE PROBOFMISSING BETA
#SIMULATION LINEAROREXPOENTIAL*/
%LET N=100;
%LET SM=100;
%LET SIGMA=2;
%LET _T=4;
%LET G=1;
%LET_PI=CONSTANT('PI');
%LET RHO=0.20;

%MACRO ONEGRP;
%DO S=1 %TO &SM ;
    PROC IML;
        COV=J(4,4,0);
        DO I=1 TO 4;
            DO J=1 TO 4;
                COV[I,J]=&SIGMA*&RHO**ABS(J-I);
            END;
        END;
    END;
BETA={5,-0.4};
MU={10,11,13,16}*BETA;
SEED=123214+&S*23;
CALL VNORMAL(Y&S,MU,COV,&N,SEED);

UNI=J(&N,1,0);
MISS=0;
ID=J(&N,1,0);
MM&S=J(&N,1,0);

DO I=1 TO &N;
  ID[I,1]=I;
  SEED[I]=SEED+I*730515+1793*(&RHO=0.8)+23*(&RHO=0.4)+1544*(&RHO=0.2);
  UNI[I,1]=RANUNI(SEED[I]);
  IF UNI[I,1]>.05 THEN DO;
    MM&S[I]=Y&S[I,4];
  END;
  ELSE DO;
    MISS=MISS+1;
    MM&S[I]=.;
  END;
END;

Y&S=Y&S||MM&S;

CALL SORT(Y&S,5);
Y&S=ID||Y&S;

ROW=%EVAL(&_T)*%EVAL(&N);
UNIVARPT&S=J(ROW,4,0);
T={0,1,3,6};

DO J=1 TO &N;
  DO I=1 TO &_T;
    K=(J-1)*%EVAL(&_T)+I;
    UNIVARPT&S[K,1]=Y&S[I,1];
    UNIVARPT&S[K,2]=T[I,1];
    UNIVARPT&S[K,3]=Y&S[I+1];
    UNIVARPT&S[K,4]=1;
  END;
END;

CREATE UNIVARFULL&S VAR{ID TIME Y GROUP};
APPEND FROM UNIVARPT&S;
PRINT COV MU BETA MISS SEED ;

DO J=1 TO &N;
  DO I=1 TO &_T;
    K=(J-1)*%EVAL(&_T)+I;
    UNIVARPT&S[K,1]=Y&S[I,1];
    UNIVARPT&S[K,2]=T[I,1];
    UNIVARPT&S[K,3]=Y&S[I+1]*(I<=3)+Y&S[I,1]*(I=4);
UNIVARPT&S[K,4]=1;
END;
END;

Y&S=Y&S[1:4]|Y&S[6];

/*MULTIVARIATE CASE*/
CREATE MISS&S VAR{MISS};
APPEND FROM MISS;
CLOSE;
CREATE UNIVARPT&S VAR{ID TIME Y GROUP};
APPEND FROM UNIVARPT&S;
CLOSE;
CREATE IMPS1&S VAR{ID WEEK0 WEEK1 WEEK3 WEEK6};
APPEND FROM Y&S;
CLOSE;
QUIT;
ODS SELECT ALL;
DATA MISS&S;
SET MISS&S;
CALL SYMPUT("MISS&S", MISS);
RUN;

%END;

%DO S=1 %TO &SM ;
   TITLE "SIMULATION &S";
   %EMALGORITHM(1,100,&_T,&N,IMPUTE,UNIVARPT&S,IMPS0,IMPS1&S,Y,0,0,0,0,0,0,0,&
   &MISS&S);
   DATA MLW4&S;
   SET MLW4;
   RUN;
   DATA FINAL&S;
   SET COMPLETE;
   RUN;
   ODS SELECT NONE;
   PROC MIXED DATA=FINAL&S;
   CLASS TIME;
   MODEL RESPONSE=TIME/S OUTPM=PREDICTED&S;
   REPEATED TIME/TYPE=AR(1) SUBJECT=ID R=1 &N ;
   LSMEANS TIME;
   ODS OUTPUT COVPARMS=COVPARMS&S;
   RUN;
   TITLE " ";
   DATA IMPS&S;
   SET IMPS11;
   RUN;ODS SELECT ALL;
%END;

TITLE "SIMULATION RESULTS &SM";
PROC IML;
    BETABAR=0;
    COVBAR=0;
    IMPUT=0;
    IMPUTABS=0;
    MBAR=0;
    MEAN=0;
    MLW4=0;
    NMISS=0;
    MEANS=0;
    DIFFS=0;
    COVS=0;
    MLW4S=0;
    NMISSS=0;
    ITER=0;
    ITERS=0;

%DO S=1 %TO &SM ;
    USE ITERATION&S ;
    READ ALL VAR{ITER} INTO ITER&S ;
    CLOSE ;
    USE COVPARMS&S ;
    READ ALL VAR{ESTIMATE} INTO COV&S ;
    CLOSE ;
    USE FINAL&S ;
    READ ALL VAR{RESPONSE} INTO RESPONSE&S ;
    CLOSE ;
    USE MLW4&S ;
    READ ALL VAR{M} INTO MLW4&S ;
    CLOSE ;
    USE UNIVARFULL&S ;
    READ ALL VAR{Y} INTO Y&S ;
    CLOSE ;
    USE IMPS&S ;
    READ ALL VAR{TIME1 TIME2 TIME3 TIME4} INTO LAST ;
    CLOSE ;
    MEAN&S=LAST[+,]/&N ;
    DIFFERENCE&S=Y&S-RESPONSE&S ;
    DIFFERENCEABS&S=ABS(DIFFERENCE&S) ;
    DATA&S=Y&S||RESPONSE&S||DIFFERENCE&S ;
    IF &&MISS&S>0 THEN DATA&S=DATA&S[LOC(DATA&S[,3]<=0),];
    ELSE DATA&S=0 ;
    IF &&MISS&S>=0 THEN DO ;
        DIFF&S=SUM(DIFFERENCE&S)/&&MISS&S ;
        END ;
    ELSE DO ;
        DIFF&S=0 ;
        END ;
    DIFFABS&S=0 ;
    END ;
    ITER=ITER+ITER&S ;
    IMPUT=IMPUT+DIFF&S ;
    COVBAR=COVBAR+COV&S ;
D.2 Program for individual regression prediction method

This SAS code performs the simulation study for REG method and MAR assumption. The cases for MDT assumption and other simulation conditions in this dissertation were very similarly programmed.
LIBNAME IMPUTE "F:\SIMULATION\LINEAR2";
/*FILENAME MYMACRO "G:\SIMULATION"; /*AUTOCALL MACRO LIBRARY */
/*OPTIONS MPRINT MLOGIC SYMBOLGEN YEARCUTOFF=1950; */
/*OPTIONS MAUTOSOURCE
SASAUTOS=(MYMACRO,SASAUTOS);
OPTIONS MPRINT MLOGIC SYMBOLGEN;*/

TITLE;
ODS RTF BODY="F:\ONE MISSING\SMISSING\200\REG282000554100.RTF" BODYTITLE
STYLE=NEWRTF STARTPAGE=NO;

%LET N=200;
%LET SM=100;
%LET SIGMA=2;
%LET T=4;
%LET G=1;
%LET _PI=CONSTANT('PI');
%LET RHO=0.80;

%MACRO ONEGRP;
%DO S=1 %TO &SM ;
   PROC IML;
      COV=J(4,4,0);
      DO I=1 TO 4;
      DO J=1 TO 4;
      COV[I,J]=&SIGMA*&RHO**ABS(J-I);
      END;
      END;
      BETA={5,-0.4};
      MU={1 0 1 1 3 1 6}*BETA;
      SEED=123214+&S*23;
      CALL VNORMAL(Y&S,MU,COV,&N,SEED);
      UNI=J(&N,1,0);
      MISS=0;
      MM&S=J(&N,1,0);
      YI&S=J(&N,1,0);
      ID=J(&N,1,0);
      DO I=1 TO &N;
      ID[I,1]=I;
      SEED1=SEED+I*730515+1793*(&RHO=0.8)+23*(&RHO=0.4)+1544*(&RHO=0.2); UNI[I,1]=RANUNI(SEED1);
      IF UNI[I,1]>0.05 THEN DO;
      MM&S[I,1]=Y&S[I,4];
      END;
      ELSE DO;
      MM&S[I,1]=.;
      MISS=MISS+1;
      END;
CALL SORT(Y&S,(5));
Y&S=ID\|Y&S;
ROW=%EVAL(&_T)*%EVAL(&N);
UNIVARPT&S=J(ROW,5,0);
T={0,1,3,6};

PRINT COV MU BETA MISS SEED

DO J=1 TO &N;
   DO I=1 TO &_T;
      K=(J-1)*%EVAL(&_T)+I;
      UNIVARPT&S[K,1]=Y&S[J,1];
      UNIVARPT&S[K,2]=T[I,1];
      UNIVARPT&S[K,3]=Y&S[J,1+1];
      UNIVARPT&S[K,4]=1;
      UNIVARPT&S[K,5]=Y&S[J,I+1]*(I<4)+Y&S[J,I+2]*(I=4);
   END;
END;
/*MULTIVARIATE CASE*/
CREATE MISS&S VAR{MISS};
APPEND FROM MISS;
CLOSE;
CREATE UNIVARPT&S VAR{ID TIME YC GROUP YT};
APPEND FROM UNIVARPT&S;
CLOSE;

DATA MISS&S;
SET MISS&S;
CALL SYMPUT("MISS&S", MISS);
RUN;

DATA UNIVARPT&S;
SET UNIVARPT&S;
IF YT=. THEN M=1;
ELSE M=0;
RUN;

%DO I=1 %TO &\&MISS&S;
   ODS SELECT NONE;
   PROC REG DATA=UNIVARPT&S;
   MODEL YT=TIME;
   OUTPUT OUT=PRED&S&I P=YHAT;
   WHERE ID=&I;
   RUN;
   ODS SELECT ALL;
   /*PROC PRINT DATA=PRED&S&I;RUN;*/
%END;
PROC IML;
USE UNIVARPT&S;
READ ALL VAR{ID TIME YC GROUP YT M} INTO IMPUT&S;
CLOSE;
   %DO I=1 %TO &&MISS&S;
   USE PRED&S&I;
   READ ALL VAR{YC YHAT} INTO YHAT&I WHERE (TIME=6);
   CLOSE;
   H=4*&I;
   IMPUT&S[H,5]=YHAT&I[2];
%END;
CREATE IMPUT&S VAR{ID TIME YC GROUP YT M};
APPEND FROM IMPUT&S;
QUIT;

ODS SELECT NONE;
PROC MIXED DATA=IMPUT&S;
CLASS TIME;
MODEL YT=TIME;
REPEATED TIME/TYPE=AR(1) SUBJECT=ID ;
ODS OUTPUT COVPARMS=COVPARMS&S;
RUN;
ODS SELECT ALL;
%END;

PROC IML;
COVBAR=0;
COVBARS=0;
DIFF=0;
DIFFS=0;
MEAN=0;
MEANS=0;
NMISS=0;
NMISSS=0;
DENOM=0;

%DO S=1 %TO &SM;
USE IMPUT&S;
READ ALL VAR{ID TIME YC GROUP YT M} INTO IMPUT&S;
CLOSE;
/*PRINT IMPUT&S;*/
IF &&MISS&S=0 THEN DO;
DIFF&S=0;
DATA&S=0;
DENOM=DENOM+1;
%END;
ELSE DO;
DIFF&S=SUM(IMPUT&S[LOC(IMPUT&S[,6]>0),3]-
IMPUT&S[LOC(IMPUT&S[,6]>0),5])/&&MISS&S;
DATA&S=IMPUT&S[LOC(IMPUT&S[,6]>0),3][IMPUT&S[LOC(IMPUT&S[,6]>0),5];
END;
DIFF=DIFF+DIFF&S;
DIFFS=DIFFS+DIFF&S*DIFF&S;
IMPUT1&S=J(&N,&_T,0);
DO I=1 TO &N;
  DO T=1 TO &T;
    R=(I-1)*&_T+T;
    IMPUT1&S[I,T]=IMPUT&S[R,5];
  END;
END;
MEAN&S=IMPUT1&S[+]/&N;
MEAN=MEAN+MEAN&S;
MEANS=MEANS+MEAN&S[4,1]*MEAN&S[4,1];
USE COVPARMS&S;
READ ALL VAR{ESTIMATE} INTO COV&S;
CLOSE;
COVBAR=COVBAR+COV&S;
COVBARS=COVBARS+COV&S#COV&S;
NMISS=NMISS+&&MISS&S;
NMISSS=NMISSS+%EVAL(&&MISS&S)*%EVAL(&&MISS&S);
PRINT MEAN&S DATA&S DIFF&S COV&S ;
/*PRINT MEAN MEANS COVBAR COVBARS DIFF DIFFS NMISS NMISSS;*/
%END;
DIFF=DIFF/&SM;
COVBAR=COVBAR/&SM;
MEAN=MEAN/&SM;
NMISS=NMISS/&SM;
MEANSD=SQRT((MEANS-&SM*MEAN[4,]*MEAN[4,])%EVAL(&SM-1));
COVSD=SQRT((COVBARS-&SM*COVBAR#COVBAR)%EVAL(&SM-1));
DIFFSD=SQRT((DIFFS-&SM*DIFF*DIFF)%EVAL(&SM-1));
PMISSSD=SQRT((NMISSS-&SM*NMISS*NMISS)%EVAL(&SM-1))&N;
PMISS=NMISS/&N;
PRINT MEAN NMISS PMISS DIFF COVBAR DENOM;
PRINT MEANSD PMISSSD DIFFSD COVSD;
QUIT;
%MEND ONEGRP;
D.3 Program for repeated measures mixed model method

This SAS code performs the simulation study for MIXED method and MAR assumption. The cases for MDT assumption and other simulation conditions in this dissertation were very similarly programmed.

LIBNAME IMPUTE "F:\SIMULATION\OLD";
/FILENAME MYMACRO "G:\SIMULATION"; /*AUTOCLASS MACRO LIBRARY
OPTIONS MPRINT MLOGIC SYMBOLGEN YEARCUTOFF=1950; */
OPTIONS MAUTOSOURCE
SASAUTOS=(MYMACRO,SASAUTOS);
/OPTIONS MPRINT MLOGIC SYMBOLGEN;*/

ODS RTF BODY= "F:\ONE MISSING\5MISSING\5O\MIXED28500554100.RTF" BODYTITLE STYLE= NEWRTF STARTPAGE=NO;

%LET N=50;
%LET SM=100;
%LET SIGMA=2;
%LET _T=4;
%LET G=1;
%LET PI=CONSTANT('PI');
%LET RHO=0.80;

%MACRO ONEGRP;
%DO S=1 %TO &SM ;

PROC IML;

COV=J(4,4,0);
DO I=1 TO 4;
DO J=1 TO 4;
COV[I,J]=&SIGMA*&RHO**ABS(J-I);
END;
END;

BETA={5,-0.4};
MU={1 0 ,1 1.1 3.1 6}*BETA;
SEED=123214+&S*23;
CALL VNORMAL(Y&S,MU,COV,&N,SEED);
DO I=1 TO &N;
  ID[I,1]=I;
  SEED1=SEED+I*730515+1793*(&RHO=0.8)+23*(&RHO=0.4)+1544*(&RHO=0.2);
  UNI[I,1]=RANUNI(SEED1);
  IF UNI[I,1]>0.05 THEN DO;
    MM&S[I,1]=Y&S[I,4];
    END;
  ELSE DO;
    MISS=MISS+1;
    MM&S[I,1]=.;
    END;
  END;
  Y&S=Y&S||MM&S;
  CALL SORT(Y&S,(5));
  Y&S=ID||Y&S;
  ROW=%EVAL(&T)*%EVAL(&N);
  UNIVARPT&S=J(ROW,4,0);
  T={0,1,3,6};
  PRINT COV MU BETA MISS SEED;
  DO J=1 TO &N;
    DO I=1 TO &T;
      K=(J-1)*%EVAL(&T)+I;
      UNIVARPT&S[K,1]=Y&S[J,1];
      UNIVARPT&S[K,2]=T[I,1];
      UNIVARPT&S[K,3]=Y&S[J,I+1];
      UNIVARPT&S[K,4]=1;
      END;
    END;
  END;
  CREATE UNIVARFULL&S VAR(1D TIME YC GROUP);
  APPEND FROM UNIVARPT&S;
  DO J=1 TO &N;
    DO I=1 TO &T;
      K=(J-1)*%EVAL(&T)+I;
      UNIVARPT&S[K,1]=Y&S[J,1];
      UNIVARPT&S[K,2]=T[I,1];
      UNIVARPT&S[K,3]=Y&S[J,I+1]*(I<=3)+Y&S[J,I+2]*(I=4);
      UNIVARPT&S[K,4]=1;
      END;
    END;
  END;
/*MULTIVARIATE CASE*/
CREATE MISS&S VAR{MISS};
APPEND FROM MISS;
CLOSE;
CREATE UNIVARPT&S VAR{ID TIME Y GROUP};
APPEND FROM UNIVARPT&S;
CLOSE;
QUIT;

DATA MISS&S;
SET MISS&S;
CALL SYMPUT("MISS&S", MISS);
RUN;

DATA WORK;
SET UNIVARPT&S;
IF Y=. THEN M=1;
ELSE M=0;
TIME1=TIME;
RUN;

ODS SELECT NONE;
PROC MIXED DATA=WORK ;
   CLASS TIME;
   MODEL Y=TIME/OUTP=PREDICTED&S ;
   REPEATED TIME TYPE=AR(1) SUBJECT=ID ;
   ODS OUTPUT COVPARMS=COVPARMS&S ;
RUN;

DATA PREDICTED&S(KEEP=ID TIME M PRED);
   SET PREDICTED&S;
RUN;

   PROC SORT DATA=PREDICTED&S;
   BY ID TIME;
   PROC SORT DATA=UNIVARFULL&S;
   BY ID TIME;
RUN;

DATA ALL&S;
MERGE PREDICTED&S UNIVARFULL&S;
   BY ID TIME;
RUN;

PROC SORT DATA=UNIVARPT&S;
   BY ID TIME;
RUN;
DATA MEAN&S;
MERGE PREDICTED&S UNIVARPT&S;
   BY ID TIME;
RUN;

DATA MEAN&S (KEEP=Y);
SET MEAN&S;
IF Y=. THEN Y=PRED;
RUN;

DATA ALL&S;
SET ALL&S;
DIFF=YC-PRED;
WHERE M=1;
RUN;

ODS SELECT ALL;
%END;

TITLE "SIMULATION RESULTS &SM";
PROC IML;
COVBAR=O;
COVBARS=O;
DIFF=O;
DIFFS=O;
MEAN=0;
MEANS=0;
NMISS=0;
NMISSS=0;
DENOM=0;

%DO S=1 %TO &SM ;
IF &&MISS&SA=O THEN DO;
USE ALL&S;
READ ALL VAR{YC,PRED,DIFF} INTO AVE&S;
CLOSE;
DIFF&S=SUM(AVE&S[,3])/NROW(AVE&S);
END;
ELSE DO;
DIFF&S=0; AVE&S=0; DENOM=DENOM+1;
END;

DIFF=DIFF+DIFF&S;
DIFFS=DIFFS+DIFF&S*DIFF&S;

USE MEAN&S;
READ ALL VAR{Y} INTO LAST;
CLOSE;
MEAN1&S=J(&N,&_T,0);
DO I=1 TO &N;
DO T=1 TO _T;
R=(I-1)*_T+T;
MEAN1&S[I,T]=LAST[R,1];
END;
END;
MEAN&S=MEAN1&S[+,1]/&N;
MEAN=MEAN+MEAN&S;
MEANS=MEANS+MEAN&S[4,1]*MEAN&S[4,1];
USE COVPARMS&S ;
READ ALL VAR{ESTIMATE} INTO COV&S;
CLOSE;
COVBAR=COVBAR+COV&S;
COVBARS=COVBARS+COV&S#COV&S;

NMISS=NMISS+ %EVAL(&&MISS&S);
NMISSS=NMISSS+ %EVAL(&&MISS&S)*%EVAL(&&MISS&S);

PRINT MEAN&S DIFF&S COV&S AVE&S NMISS;
%END;

DIFF=DIFF/&SM;
COVBAR=COVBAR/&SM;
MEAN=MEAN/&SM;
NMISS=NMISS/&SM;
MEANSD=SQRT((MEANS-&SM*MEAN[4,]*MEAN[4,])/%EVAL(&SM-1));
COVSD=SQRT((COVBARS-&SM*COVBAR#COVBAR)/%EVAL(&SM-1));
DIFFSD=SQRT((DIFFS-&SM*DIFF*DIFF)/%EVAL(&SM-1));
PMISSSD=SQRT((NMISSS-&SM*NMISS*NMISS)/%EVAL(&SM-1))/&N;
PMISS=NMISS/&N;
PRINT MEAN PMISS DIFF COVBAR DENOM;
PRINT MEANSD PMISSSD DIFFSD COVSD;
QUIT;

%MEND ONEGRP;

%ONEGRP;
RUN;

ODS RTF CLOSE;

D.4 Program for last observation carried forward method

*****************************************************************************
This SAS code performs the simulation study for LOCF method und MAR assumption. The cases for MDT assumption and other simulation conditions in this dissertation were very similarly programmed.
*****************************************************************************

LIBNAME IMPUTE "F:\SIMULATION\LINEAR2";
/*FILENAME MYMACRO "G:\SIMULATION"; /*AUTOCALL MACRO LIBRARY */
/*OPTIONS MPRINT MLOGIC SYMBOLGEN YEARCUTOFF=1950; */
OPTIONS MAUTOSOURCE SASAUTOS=(MYMACRO,SASAUTOS);
/*OPTIONS MPRINT MLOGIC SYMBOLGEN;*/
ODS RTF BODY="F:\ONE MISSING\MISSING\200\LOCF222000554100.RTF" BODYTITLE STYLE=NEWRTF STARTPAGE=NO;
%LET N=200;
%LET SM=100;
%LET SIGMA=2;
%LET _T=4;
%LET G=1;
%LET PI=CONSTANT('PI');
%LET RHO=0.20;
%MACRO ONEGRO;
%DO S=1 %TO &SM;
  PROC IML;
  COV=J(4,4,0);
  DO I=1 TO 4;
  DO J=1 TO 4;
  COV[I,J]=&SIGMA&&RHO**ABS(J-I);
  END;
  END;
  UNI=J(&N,1,0);
  BETA={5,-0.4};
  MU={1,0,1,1,1,3,1,6}*BETA;
  SEED=123214+&S*23;
  CALL VNORMAL(Y&S,MU,COV,&N,SEED);
  MISS=0;
  MM&S=J(&N,1,0);
  YI&S=J(&N,1,0);
  ID=J(&N,1,0);
  DO I=1 TO &N;
    ID[I,1]=I;
    SEED1=SEED*I**730515+1793*(&RHO=0.8)+23*(&RHO=0.4)+1544*(&RHO=0.2);
    UNI[I,1]=RANUNI(SEED1);
    IF UNI[I,1]>0.05 THEN DO;
      MM&S[I,1]=Y&S[I,4];
      YI&S[I,1]=Y&S[I,4];
      END;
    ELSE DO;
      MM&S[I,1]=.;
      MISS=MISS+1;
      YI&S[I,1]=Y&S[I,3];
      END;
    END;
  Y&S=Y&S||MM&S||YI&S;
  CALL SORT(Y&S,{5});
  Y&S=ID||Y&S;
  ROW=%EVAL(_T)**%EVAL(&N);
  UNIVARPT&S=J(ROW,6,0);
  T={0,1,3,6};
DO J=1 TO &N;
DO I=1 TO &_T;
K=(J-1)*%EVAL(&_T)+I;
UNIVARPT&S[K,1]=Y&S[J,1];
UNIVARPT&S[K,2]=T[I,1];
UNIVARPT&S[K,3]=Y&S[J,I+1];
UNIVARPT&S[K,4]=1;
UNIVARPT&S[K,5]=Y&S[J,I+1]*(I<4)+Y&S[J,I+2]*(I=4);
UNIVARPT&S[K,6]=Y&S[J,I+1]*(I<4)+Y&S[J,I+3]*(I=4);
END;
END;

/*MULTIVARIATE CASE*/
CREATE MISS&S VAR{MISS};
APPEND FROM MISS;
CLOSE;
CREATE UNIVARPT&S VAR{ID TIME YC GROUP YT YI};
/*YC: COMPLETE DATA, YT TRUNCATED DATA, YI LOCF DATA*/
APPEND FROM UNIVARPT&S;
CLOSE;

DATA MISS&S;
SET MISS&S;
CALL SYMPUT("MISS&S", MISS);
RUN;

DATA UNIVARPT&S;
SET UNIVARPT&S;
IF YT=. THEN M=1;
ELSE M=0;
RUN;

ODS SELECT NONE;
PROC MIXED DATA=UNIVARPT&S;
   CLASS TIME;
   MODEL YI=TIME;
   REPEATED TIME/TYPES=AR(1) SUBJECT=ID ;
   ODS OUTPUT COVPARMS=COVPARMS&S;
RUN;
ODS SELECT ALL;

PROC IML;
COVBAM;
COVBAR=0;
COVBARS=0;
DIFF=0;
DIFFS=0;
MEAN=0;
MEANS=0;
NMISS=0;
NMISSS=0;
DENOM=0;

%DO S=1 %TO &SM;

USE UNIVARP&S;
READ ALL VAR{ID TIME YC GROUP YT YI M } INTO IMPUT&S;
CLOSE;
IF &&MISS&S=0 THEN DO;
DIFF&S=0;
DATA&S=0;
DENOM=DENOM+1;
END;
ELSE DO;
DIFF&S=SUM(IMPUT&S[LOC(IMPUT&S[,7]>0),3]-IMPUT&S[LOC(IMPUT&S[,7]>0),6])/&MISS&S;
DATA&S=IMPUT&S[LOC(IMPUT&S[,7]>0),3]]IMPUT&S[LOC(IMPUT&S[,7]>0),6];
END;
DIFF=DIFF+DIFF&S;
DIFFS=DIFFS+DIFF&S*DIFF&S;
IMPUT1&S=J(&N,&-T,O);
DO I=1 TO &N;
DO T=1 TO &-T;
R=(I-I)*&-T+T;
IMPUT1&S[I,T]=IMPUT&S[R,6];
END;
END;

MEAN&S=IMPUT1&S[+,]/&N;
MEAN=MEAN+MEAN&S;
MEANS=MEANS+MEAN&S[4,1]*MEAN&S[4,1];

USE COVPARMS&S;
READ ALL VAR{ESTIMATE} INTO COV&S;
CLOSE;
COVBAR=COVBAR+COV&S;
COVBARS=COVBARS+COV&S#COV&S;

NMISS=NMISS+&&MISS&S;
NMISSS=NMISSS+&&MISS&S)*%EVAL(&&MISS&S);%
PRINT MEAN&S DATA&S DIFF&S COV&S ;
/*PRINT MEAN MEANS COVBAR COVBARS DIFF DIFFS NMISS NMISSS;*/

%END;

DIFF=DIFF/&SM;
COVBAR=COVBAR/&SM;
MEAN=MEAN/&SM;
NMISS=NMISS/&SM;
MEANSD=SQR((MEANS-&&SM*MEAN[4,]*MEAN[4,])/%EVAL(&SM-1));
COVSD=SQR((COVBARS-&&SM*COVBAR#COVBAR)/%EVAL(&SM-1));
DIFFSD=SQRT((DIFFS-&SM*DIFF*DIFF)/EVAL(&SM-1));
PMISSSD=SQRT((NMISSS-&SM*NMISS*NMISS)/EVAL(&SM-1)/&N);
PMISS=NMISS/&N;
PRINT MEAN COVBAR PMISS DIFF NMISS DENOM ;
PRINT MEANSD COVSD PMISSSD DIFFSD;
QUIT;

%MEND ONEGRP;

%ONEGRP;
RUN;
ODS RTF CLOSE;
APPENDIX E

Programs for MDT Method in conjunction with Multiple Imputation Method including Bootstrap Sampling Algorithm
APPENDIX E

Programs for MDT Method in conjunction with Multiple Imputation Method including Bootstrap Algorithm

E.1 Program for bootstrap sampling of IMPS data and application of MDT method to bootstrap samples

*******************************************************************************

This program creates bootstrap samples and obtains the model parameter estimates by applying MDT method to bootstrap samples.

Input data set
Impute.imps multivariate format of IMPS data
  Variables: ID, week0-week6, group and flag

Output data sets:
impute.parameter1 contains the parameter estimates for treatment group from each bootstrap samples by MDT method
  variables: sample number, beta, variance and correlation estimates
impute.parameter0 contains the parameter estimates for placebo group from each bootstrap samples by MDT method
  variables: sample number, beta, variance and correlation estimates
impute.mean0: average of parameter estimates from bootstrap samples from
  impute.parameter0 for placebo group
impute.mean1: average of parameter estimates from bootstrap samples from
  impute.parameter1 for treatment group
impute.covariance0: variance-covariance estimates of model parameters for placebo group
impute.covariance1: variance-covariance estimates of model parameters for treatment group

*******************************************************************************

LIBNAME IMPUTE "C:\MI IMPUTATION";
FILENAME MYMACRO "C:\MI IMPUTATION\IMPUTE FINAL"; /*AUTOCALL MACRO LIBRARY */
OPTIONS MPRINT MLOGIC SYMBOLGEN YEARCUTOFF=1950;
OPTIONS MAUTOSOURCE
SASAUTOS=(MYMACRO,SASAUTOS);
TITLE;
%LET N_BOOT=500;
%LET _T=4;
%LET _PI=CONSTANT('PI');
%LET N=95;
%LET _PI=CONSTANT('PI');
%GLOBAL _PI;

PROC PRINTTO LOG="G:\MI IMUTATION\IMPUTE FINAL\BOOTSTRAPMVN\NBOOT.TXT" NEW;RUN;
/*PROC PRINTTO LOG=LOG;RUN;*/

ODS RTF BODY="C:\MI IMUTATION\IMPUTE FINAL\BOOTSTRAP\N_BOOT.RTF" BODYTITLE
STYLE=NEWRTF STARTPAGE=NO;

DATA WORKRL (RENAME=(WEEK0=TIME1 WEEK1=TIME2 WEEK3=TIME3 WEEK6=TIME4
GROUP=GRP));
SET IMPUTE.IMPS;
RUN;

PROC PRINT DATA=WORKRL;RUN;

%MACRO BOOTSAmp;

DATA IMPS01(WHERE=(GRP=0))
   IMPS11(WHERE=(GRP=1));
      /*CREATE ONE DATASET FOR EACH TREATMENT*/
   SET WORKRL;
RUN;

%DO B=2 %TO &N_BOOT;
     /*CREATE &N_BOOT BOOTSTRAP REPLICATION*/
     /*CREATE INDEPENDENT SETS OF REPLICATIONS IN TERMS OF TREATMENT*/

   DATA IMPS0&B;
   DO I=1 TO NOBS;
   PT=CEIL(RANUNI(0)*NOBS);
   SET IMPS01 NOBS=NOBS POINT=PT; /*USE TREATMENT-SPECIFIC DATA*/
   OUTPUT;
   END;
   STOP;
RUN;

   DATA IMPS1&B;
   DO I=1 TO NOBS;
   PT=CEIL(RANUNI(0)*NOBS);
   SET IMPS11 NOBS=NOBS POINT=PT; /*USE TREATMENT-SPECIFIC DATA*/
   OUTPUT;
   END;
   STOP;
RUN;
/* PROC PRINT DATA=IMPS0&B;RUN;
PROC PRINT DATA=IMPS1&B;RUN;*/
%END;

%DO B=1 %TO &N_BOOT;
  TITLE "BOOTSTRAP &B";
  DATA WORK&B;
  SET IMPS0&B IMPS1&B ;
  DROP I;
  RUN;
  PROC SORT DATA=WORK&B;
  BY GRP FLAG DESCENDING ID ;
  RUN;
  ODS SELECT NONE;
  PROC FREQ DATA=WORK&B;
  BY GRP;
  TABLE FLAG ID;
  ODS OUTPUT FREQ.BYGROUP2.TABLE2.ONEWAYFREQS=STATT;
  ODS OUTPUT FREQ.BYGROUP1.TABLE1.ONEWAYFREQS=STATP;
  RUN;
  %LET MISS0UP=0;
  %LET MISS13UP=0;
  %LET MISS1LW=0;
  %LET MISS14UP=0; /*PUT &MISS0UP &MISS1LW &MISS13UP &MISS14UP;*/

DATA STAT0;
/*CREATE MACRO VARIABLE FOR MDT IN PLACEBO GROUP*/
  SET STATP;
  IF F_FLAG="FAILURE" THEN CALL SYMPUT('MISS0UP,FREQUENCY);
  RUN;

DATA STAT1;
/*CREATE MACRO VARIABLE FOR MDT(UPPER AND LOWER) IN TREATMENT GROUP*/
  SET STATT;
  IF ID=4506 THEN CALL SYMPUT('MISS1LW,FREQUENCY);
  IF ID=2121 THEN CALL SYMPUT('MISS14UP,FREQUENCY);
  IF ID=2372 THEN DO;
    CALL SYMPUT('MISS13UP,FREQUENCY);
    MISS14UP=SYMGET('MISS14UP)+FREQUENCY;
    CALL SYMPUT('MISS14UP,MISS14UP);
  END;
  RUN;

  %PUT &MISS0UP &MISS1LW &MISS13UP &MISS14UP;

DATA IMPUTE.IMPS1&B(WHERE=(ID>24));
/*CREATE ONE DATASET FOR EACH PLACEBO BOOTSTRAP SAMPLE*/
SET WORK&B;
   ID= _N_; 
   DROP GRP ;
RUN;

DATA IMPUTE.IMPS0&B(where=(ID<=24));
/*OUTPUT EACH PLACEBO BOOTSTRAP SAMPLE*/
SET WORK&B;
   ID= _N_; 
   DROP GRP ;
RUN;

/*PROC PRINT DATA=IMPUTE.IMPS0&B;
PROC PRINT DATA=IMPUTE.IMPS1&B;
RUN;*/

DATA IMPUTE.WORK&B;
SET WORK&B;
   ID= _N_; 
   TIME=0;RESPONSE=TIME1; GROUP=GRP; OUTPUT;
   TIME=1;RESPONSE=TIME2; GROUP=GRP; OUTPUT;
   TIME=3;RESPONSE=TIME3; GROUP=GRP; OUTPUT;
   TIME=6;RESPONSE=TIME4; GROUP=GRP; OUTPUT;
   DROP TIME1-TIME4 GRP FLAG ;
RUN;

/*PROC PRINT DATA=IMPUTE.WORK&B;RUN;*/

TITLE "BOOTSTRAP &B";
%EMALGORITHM(1,100,&_T,&N,IMPUTE,WORK&B,IMPS0&B,IMPS1&B,Y,&MISS0UP,&MISS13UP,0,0,&MISS0UP,&MISS14UP,0,&MISS1LW);

DATA MEAN&B;
SET MEAN;
RUN;

%DO G=0 %TO 1;
DATA COVPARMSS&G&B;
SET COVPARMS&B;
WHERE GROUP="GROUP &G";
SAMPLE=SYMGET('B');
DROP SUBJECT GROUP;
RUN;

DATA MEANS&G&B;
SET MEAN&B;
WHERE GROUP=&G;
SAMPLE=SYMGET('B');
RUN;

DATA MUP3&G&B;
SET MUP3&G&B;
SAMPLE=SYMGET('B');
RUN;

DATA MUP4&G&B;
SET MUP4&G&B;
SAMPLE=SYMGET('B');
RUN;
%END;

DATA MLW41&B;
SET MLW41&B;
SAMPLE=SYMGET('B');
RUN;
%END;

%DO J=2 %TO &N-BOOT;
PROC APPEND BASE=MUP301 DATA=MUP30&J;
PROC APPEND BASE=MUP401 DATA=MUP40&J;
PROC APPEND BASE=MUP311 DATA=MUP31&J;
PROC APPEND BASE=MUP411 DATA=MUP41&J;
PROC APPEND BASE=MLW411 DATA=MLW41&J;
PROC APPEND BASE=COVPARMSS01 DATA=COVPARMSS0&J;
PROC APPEND BASE=COVPARMSS11 DATA=COVPARMSS1&J;
PROC APPEND BASE=MEANS01 DATA=MEANS0&J;
PROC APPEND BASE=MEANS11 DATA=MEANS1&J;
RUN;
%END;
%MEND;

%BOOTSMAP;

TITLE "BOOTSTRAP RESULTS (BOOTSAMPLE=&N_BOOT)";
***TRANSFORM COVARIANCE PARAMETER OF PLACEBO GROUP TO FLAT FORMAT******;
DATA COVPARMSS0 ;
SET COVPARMSS01;
IF COVPARM="VARIANCE" THEN VARIANCE=ESTIMATE;
ELSE DO;
RETAIN VARIANCE;
CORRELATION=ESTIMATE;
END;
RUN;

DATA COVPARMSS0;
SET COVPARMSS0;
WHERE CORRELATION NE .;
DROP COVPARM ESTIMATE;
RUN;
***TRANSFORM COVARIANCE PARAMETER OF TREATMENT GROUP TO FLAT FORMAT********;

DATA COVPARMSS1;
SET COVPARMSS1;
IF COVPARM="VARIANCE" THEN VARIANCE=ESTIMATE;
ELSE DO;
RETAIN VARIANCE;
CORRELATION=ESTIMATE;
END;
RUN;

DATA COVPARMSS1;
SET COVPARMSS1;
WHERE CORRELATION NE.;
DROP COVPARM ESTIMATE;
RUN;

***TRANSFORM MEAN ESTIMATES OF PLACEBO GROUP TO FLAT FORMAT********;

DATA MEANS0;
SET MEANS01;
IF TIME=0 THEN XBAR0=MEAN;
IF TIME=1 THEN DO;
RETAIN XBAR0;
XBAR1=MEAN;
END;
IF TIME=3 THEN DO;
RETAIN XBAR0 XBAR1;
XBAR3=MEAN;
END;
IF TIME=6 THEN DO;
RETAIN XBAR0 XBAR1 XBAR3;
XBAR6=MEAN;
END;
RUN;
DATA MEANS0;
SET MEANS0;
WHERE XBAR6 NE.;
DROP MEAN GROUP TIME;
RUN;

DATA MEANS0BETA;
SET MEANS0;
BETA0=XBAR0;
BETA1=XBAR1-XBAR0;
DROP XBAR0-XBAR6;
RUN;

***TRANSFORM MEAN ESTIMATES OF TREATMENT GROUP TO FLAT FORMAT******;
DATA MEANS1;
SET MEANS1;
IF TIME=0 THEN XBAR0=MEAN;
IF TIME=1 THEN DO;
RETAIN XBAR0;
XBAR1=MEAN;
END;
IF TIME=3 THEN DO;
RETAIN XBAR0 XBAR1;
XBAR3=MEAN;
END;
IF TIME=6 THEN DO;
RETAIN XBAR0 XBAR1 XBAR3;
XBAR6=MEAN;
END;
RUN;
DATA MEANS1;
SET MEANS1;
WHERE XBAR6 NE .;
DROP MEAN GROUP TIME;
RUN;
DATA MEANS1BETA;
SET MEANS1;
BETA0=XBAR0;
BETA1=XBAR1-XBAR0;
DROP XBAR0-XBAR6;
RUN;
DATA MUP301(RENAME=(MUP=MUP3)); SET MUP301; RUN;
/*PROC PRINT DATA=MUP301;RUN;*/
DATA MUP401(RENAME=(MUP=MUP4)); SET MUP401; RUN;
/*PROC PRINT DATA=MUP401;RUN;*/
DATA MUP311(RENAME=(MUP=MUP3)); SET MUP311; RUN;
/*PROC PRINT DATA=MUP311;RUN;*/
DATA MUP411(RENAME=(MUP=MUP4)); SET MUP411; RUN;
/*PROC PRINT DATA=MUP411;RUN;*/
DATA MLW411(RENAME=(MLW=MLW4)); SET MLW411;
*PROC PRINT DATA=MLW411;
RUN;
PROC SORT DATA= MUP301;
BY SAMPLE;
RUN;
PROC SORT DATA=MUP401;
BY SAMPLE;
RUN;
PROC SORT DATA=MUP311;
BY SAMPLE;
RUN;
PROC SORT DATA=MUP411;
BY SAMPLE;

PROC SORT DATA=MLW411;
BY SAMPLE;
RUN;
PROC SORT DATA=COVPARMSS0;
BY SAMPLE;

PROC SORT DATA=COVPARMSS1;
BY SAMPLE;

PROC SORT DATA=MEANS0BETA;
BY SAMPLE;

PROC SORT DATA=MEANS1BETA;
BY SAMPLE;
RUN;

/*PROC PRINT DATA=MEANS0BETA;
PROC PRINT DATA=MEANS1BETA;RUN;
PROC PRINT DATA=PARAMETER0;
PROC PRINT DATA=PARAMETER1;RUN;*/

**************MERGE ALL THE PARAMETER ESTIMATES OF PLACEBO TO PARAMETER0 DATA SET**************;
DATA PARAMETER0;
MERGE MEANS0BETA COVPARMSS0 MUP301 MUP401;
BY SAMPLE;
RUN;

PROC CORR DATA=PARAMETER0 COV;
ODS OUTPUT COV=COVARIANCE0 SIMPLESTATS=MEAN0;
RUN;

DATA COVARIANCE0;
SET COVARIANCE0;
KEEP BETA0 BETA1 VARIANCE CORRELATION MUP3 MUP4;
RUN;

DATA IMPUTE.COVARIANCE0;
SET COVARIANCE0;
RUN;

DATA MEAN0;
SET MEAN0;
KEEP VARIABLE MEAN;
RUN;
DATA IMPUTE.MEAN0;
SET MEAN0;
RUN;

DATA IMPUTE.PARAMETER0;
SET PARAMETER0;
RUN; /*OUTPUT MEAN AND VARIANCE-COVARIANCE MATRIX ESTIMATES OF PARAMETERS */

***************MERGE ALL THE PARAMETER ESTIMATES OF TREATMENT TO PARAMETER1 DATA SET**************;

DATA PARAMETER1;
MERGE MEANS1BETA COVPARMSS1 MUP311 MUP411 MLW411;
BY SAMPLE;
RUN;
PROC CORR DATA=PARAMETER1 COV;
ODS OUTPUT COV=CovARIANCE1 SIMPLESTATS=MEAN1;
RUN;
DATA COVARIANCE1;
SET COVARIANCE1;
KEEP BETA0 BETA1 VARIANCE CORRELATION MUP3 MUP4 MLW4;
RUN;
DATA IMPUTE.COVARIANCE1;
SET COVARIANCE1;
RUN;
DATA MEAN1;
SET MEAN1;
KEEP VARIABLE MEAN;
RUN;
DATA IMPUTE.MEAN1;
SET MEAN1;
RUN;
DATA IMPUTE.PARAMETER1;
SET PARAMETER1;
RUN; /*OUTPUT MEAN AND VARIANCE-COVARIANCE MATRIX ESTIMATES OF PARAMETERS */
ODS RTF CLOSE;

E.2 Program for combining the MDT method with multiple imputation for IMPS data set

This code combines the MDT method with multiple imputation for IMPS data. The model parameter estimates were obtained from original IMPS data using MDT method, the variance-covariance of the model parameter estimates were obtained from bootstrap sampling.
Input data sets

Impute.imps: multivariate format of IMPS data
Variables: ID week0-week6 group and flag
impute.covariance0: variance-covariance estimates of model parameter for placebo group
impute.covariance1: variance-covariance estimates of model parameter for treatment group

LIBNAME MI "C:\MI IMPUTATION\";
OPTIONS MPRINT MLOGIC SYMBOLGEN YEARCUTOFF=1950;
/*BOOTSTRAP WITH PARAMETERS FROM THE MDT METHODS WITH AR(1) AND SQRT(TIME) BOOTSTRAP NUMBER IS 1000*/

PROC PRINTTO LOG=LOG; RUN;
%LET _T=4;
%LET MISSUP03=5;
%LET MISSUP04=5;
%LET MISSLW03=0;
%LET MISSLW04=0;
%LET MISSLW13=0;
%LET MISSLW14=1;
%LET MISSLW13=1;
%LET MISSLW14=2;
%LET MISSLW13=0;
%LET MISSLW14=1;
%LET MISSLW13=0;
%LET MISSLW14=1;
%LET MISSLW13=1;
%LET MISSLW14=2;
%LET MISSLW13=0;
%LET MISSLW14=1;
%LET MISSLW13=6;
%LET MISSLW14=7;
%LET N_IMPUTE=10;

PROC PRINT DATA=MI.IMPS;RUN;
DATA MI.IMPS11;
SET MI.IMPS;
WHERE GROUP=1;
DATA MI.IMPS01;
SET MI.IMPS;
WHERE GROUP=0;
RUN;

PROC SORT DATA=MI.IMPS11;
BY FLAG WEEK3;
RUN;
PROC SORT DATA=MI.IMPS01;
BY FLAG WEEK3;
RUN;
PROC PRINT DATA=MI.IMPS01;
PROC PRINT DATA=MI.IMPS11;
RUN;

DATA IMPSM10;
SET MI.IMPS01; 
DROP GROUP; 
DATA IMPSMI1; 
SET MI.IMPS11; 
DROP GROUP; 
RUN; 

PROC PRINT DATA=IMPSMI0; 
PROC PRINT DATA=IMPSMI1; 
RUN; 

%MACRO MI(MI0,MI1); 

PROC IML; 
USE &MIO; 
READ ALL INTO MIO; 
USE &MI1; 
READ ALL INTO MI1; 
CLOSE; 
PRINT MIO MI1; 

%DO I=1 %TO &N_IMPUTE; 
SEED0=07161973; 
SEED1=05151973; 
SEED2=03221978; 
SEED6=03221234; 

MU0={5.2396,4.7971,4.4732,4.1558}; 
MU1={5.2635,4.5359,4.0032,3.4812}; 

%DO G=0 %TO 1; 

USE MI.COVARIANCE&G; 
READ ALL INTO SIGMA&G; 

CALL VNORMAL(PAR&G, MU&G, SIGMA&G,1); 
PRINT PAR&G; 
MEAN&G=J(4,1,0); 
MEAN&G[1]=PAR&G[1]; 
MEAN&G[3]=PAR&G[1]+PAR&G[2]*SQRT(3); 

COV&G=J(4,4,0); 
DO I=1 TO 4; 
DO J=1 TO 4; 
COV&G[I,J]=PAR&G[3]*PAR&G[4]**ABS(J-I); 
END; 
END; 

PRINT COV&G MEAN&G; 
%END;
MUP03=PAR0[,5];
MUP04=PAR0[,6];
MUP13=PAR1[,5];
MUP14=PAR1[,6];
MLW14=PAR1[,7];
PRINT MUP03 MUP04 MUP13 MUP14 MLW14;

%DO G=0 %TO 1;
%DO H=2 %TO 3;
COV=COV&G;
K=%EVAL(&H);
VAR1&H=COV[1,K,1:K];
VAR%EVAL(&H+1)=COV[K+1,K+1];
COV%EVAL(&H+1)1=COV[K+1,1:K];
SIGMA%EVAL(&H+1)=VAR%EVAL(&H+1)-
COV%EVAL(&H+1)1*INV(VAR1&H)*COV%EVAL(&H+1)1';
SD%EVAL(&H+1)=SQRT(SIGMA%EVAL(&H+1));
**VARIANCES OF CONDITIONAL NORMAL DISTRIBUTION;
%LET MISS=&G%EVAL(&H+1);
IF &G=O THEN
MU%EVAL(&H+1)&G=J(%EVAL(&&MISSUP&MISS+&&MISS+&&MISSLW&MISS+&&MISSMAR&MISS),1,0);
IF &G=1 THEN MU%EVAL(&H+1)&G=J((&MISSUP14+&MISSLW14+&MISSMAR&MISS),1,0);

IF %EVAL(&&MISSUP&MISS)>0 THEN DO;
DO S=1 TO &MISSUP&MISS;  *PRINT "&G GROUP &H+1 TIME UPPER";  *PRINT U;
MU%EVAL(&H+1)&G[S,1]=MEAN&G[%EVAL(&H+1),1]+COV%EVAL(&H+1)1*INV(VAR1&H)*(MI&G[S,2:K+1]-MEAN&G[1,K,1]);
*&H+1 MEAN OF NORMAL DISTRIBUTION CONDITIONING ON THE PREVIOUS TIMEPOINT;
ZLOWER=(MUP&G%EVAL(&H+1)-MU%EVAL(&H+1)&G[S,1])/SD%EVAL(&H+1);
B=CDF('NORMAL',ZLOWER);A=1-CDF('NORMAL',ZLOWER);
U=A*RANUNI(SEED2*S*%EVAL(&H)*%EVAL(&G+1))+B;
MI&G[S,%EVAL(&H+2)]=MU%EVAL(&H+1)&G[S,1]+QUANTILE('NORMAL',U)*SD%EVAL(&H+1);
**IMPUTED TRUNCATED NORMAL DISTRIBUTION;
IF MI&G[S,%EVAL(&H+2)]>7 THEN MI&G[S,%EVAL(&H+2)]=7;
END;
END;

IF &MISSLW&MISS>0 THEN DO;
DO S=%EVAL(&MISSUP&MISS+1) TO %EVAL(&MISSUP&MISS+&MISSLW&MISS);
MU%EVAL(&H+1)&G[S,1]=MEAN&G[%EVAL(&H+1),1]+COV%EVAL(&H+1)1*INV(VAR1&H)*(MI&G[S,2:K+1]-MEAN&G[1,K,1]);
*&H+1 MEAN OF NORMAL DISTRIBUTION CONDITIONING ON THE PREVIOUS TIMEPOINT;
ZUPPER=(MLW&G%EVAL(&H+1)-MU%EVAL(&H+1)&G[S,1])/SD%EVAL(&H+1);
A=CDF('NORMAL',ZUPPER);
U=A*RANUNI(SEED6*S%EVAL(&H)*%EVAL(&G+1));
IF MI&G[S,%EVAL(&H+2)]<1 THEN MI&G[S,%EVAL(&H+2)]=1;
END;
END;

IF &&MISSMAR&MISS>O THEN DO;
IF &G=0 THEN DO; 
S1=%EVAL(&&MISSUP&MISS+&&MISSLW&MISS+1);
S2=(&MISSUP&MISS+&MISSLW&MISS+&MISSMAR&MISS); 
END;
IF &G=1 THEN DO;
S1=%EVAL(&MISSUP14+&MISSLW14+1);
S2=(&MISSUP14+&MISSLW14+&MISSMAR&MISS); 
END;
DO S=S1 TO S2;
MU%EVAL(&H+1)&G[S,1]=MEAN&G[%EVAL(&H+1),1]+COV%EVAL(&H+1)1*INV(VAR1&H)*(MI&G[S,2:K+l]'-MEAN&G[l:K,l]);

*H+1 MEAN OF NORMAL DISTRIBUTION CONDITIONING ON THE PREVIOUS TIMEPOINT;
A=SEED1*S*%EVAL(&H)*%EVAL(&G+1);
AA=MU%EVAL(&H+1)&G[S,1]+RANNOR(A)*SD%EVAL(&H+1);
MI&G[S,%EVAL(&H+2)]=AA;
IF MI&G[S,%EVAL(&H+2)]<1 THEN MI&G[S,%EVAL(&H+2)]=1;
IF MI&G[S,%EVAL(&H+2)]>7 THEN MI&G[S,%EVAL(&H+2)]=7;
END;
END;

PRINT MI0 MI1;

MIIMPS&I=MI0// MI1; PRINT MIIMPS&I;
CREATE MIIMPS&I VAR("TIME0":"TIME&_T");
**TIME0 IS ID VAR;
APPEND FROM MIIMPS&I;

%END;
QUIT;

%DO J=1 %TO &N IMPUTE;

DATA MIIMPSS&J;
  SET MIIMPS&J;
  IMPUT=SYMGET('J');
  _IMPUTATION_=INPUT(IMPUT,BEST4.);
  IF _N_<=25 THEN GROUP=0;
  IF _N_>24 THEN GROUP=1;
RUN;

%END;

%DO J=2 %TO &N IMPUTE;
PROC APPEND BASE=MIIMPSS1 DATA=MIIMPSS&J;
RUN;

%END;
DATA MIIMPSS1 (RENAME=(TIME0=ID));
SET MIIMPSS1;
RUN;

DATA MIIMPSS1;
SET MIIMPSS1;
IMPS=TIME1; TIME=0; OUTPUT;
IMPS=TIME2; TIME=1; OUTPUT;
IMPS=TIME3; TIME=3; OUTPUT;
IMPS=TIME4; TIME=4; OUTPUT;
DROP TIME1-TIME4;
RUN;

%MEND;

%MI(IMPSM0,IMPSM1);

PROC PRINT DATA=MIIMPSS1;RUN;

ODS RTF BODY="E:\MI IMPUTATION\IMPUTE FINAL\MI ANALYZE\MIMDT1000MVN.RTF"
BODYTITLE STYLE=NEWRTF STARTPAGE=NO;

TITLE "MULTIPLE IMPUTATION WITH MDT METHOD";
TITLE2 "PARAMETERS ARE RANDOMLY DRAWING FROM MULTIVARIATE NORMAL DISTRIBUTION";
DATA MIIMPSS1;
SET MIIMPSS1;
STIME=SQR(TIME);
RUN;

ODS SELECT NONE;

PROC MIXED DATA=MIIMPSS1;
CLASS TIME;
MODEL IMPS= GROUP STIME GROUP*STIME/SOLUTION COVB;
REPEATED TIME/TYPE=AR(1) SUBJECT=ID GROUP=GROUP;
ODS OUTPUT SOLUTIONF=MIXPARMS COVB=MIXCOVB;
BY _IMPUTATION_;
RUN;
ODS SELECT ALL;

PROC PRINT DATA=MIXPARMS; RUN;
PROC PRINT DATA=MIXCOVB; RUN;

PROC MIANALYZE PARMs=MIXPARMS EDF=283
    COVB(EFFECTVAR=ROWCOL)=MIXCOVB;
    MODELEFFECTS INTERCEPT GROUP STIME GROUP*STIME;
RUN;

ODS RTF CLOSE;
Rong Liu was born on May 15, 1973 in Taiyuan, Shanxi Province, People’s Republic of China. She graduated from Taiyuan No. 5 High School in 1992. In 1997, she graduated with a Bachelor of Medicine from Shandong Medical University, Shandong, China. In 2000, she received a M.P.H. from Beijing Epidemiology and Microbiology Institute, Beijing, China. She also earned a M.S. in Statistics from University of Virginia in 2004. She began her Ph.D. studies in Biostatistics at Virginia Commonwealth University in the fall of 2003. Rong Liu has been an intern at Wyeth Consumer Healthcare, Richmond, Va. In addition, she also worked as research assistant for Beijing Epidemiology and Microbiology Institute from 2000 to 2001.