STATISTICAL METHODS FOR ASSESSING THE EFFECT OF MORTALITY ON RATES OF CHANGE AND VARIABILITY IN A LONGITUDINAL STUDY OF THE ELDERLY

Christian Elizabeth Douglas

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Public Health in the Department of Biostatistics in the Gillings School of Global Public Health.

Chapel Hill
2014

Approved by:

Lloyd Edwards

John Preisser, Jr.

Chirayth M. Suchindran

Peggye Dilworth-Anderson

Pranab K. Sen

ABSTRACT

Christian Elizabeth Douglas: Statistical Methods for Assessing the Effect of Mortality on Rates of Change and Variability in a Longitudinal Study of the Elderly
(Under the direction of Lloyd Edwards)

Despite the benefits of longitudinal analysis for describing the aging process, it is not absent of complications. Failing to account for nonrandom attrition and other mechanisms that affect the ability to acquire follow-up measurements may result in estimates on a relatively healthy or advantaged sample in terms of health and economic means. In modeling the process of aging in older adults, handling of attrition requires careful attention, since attrition can affect the interpretation of the conclusions. Longitudinal studies of older adults are particularly sensitive to the truncation due to death, which is usually the largest category of nonresponse in studies of older adults. We examine the effect of death on rates of change and variability on a well-established data set of older adults leaving in the community. Our assessment utilizes models proposed to analyze data with outcomes truncated due to death.

Using proposed methods, we analyzed an imputed NC EPESE dataset allowing only truncation due to death. Simulations were completed to evaluate the models ability to estimate the rates of change under varying burdens of death. Additionally, the use of these methods in presence of non-participation and death was examined using the original NC EPESE. Allowing the missing mechanisms to depend on the outcomes of interest, simulations were conducted to describe the methods behavior in estimating rates of change for non-missing completely at random data. Finally, an assessment of

the variability about the parameter estimates was completed.

Sample size and missing completely at random burdens of death were not extremely impactful on the models ability to estimate the rates of change. However, this was not true for not missing at random data for estimates of rates of change or variability.

This dissertation is dedicated to the memory of my fathers, Wayne Douglas and Larry Hicks, Sr.

ACKNOWLEDGMENTS

First, I must thank my family, whose love and encouragement have made my accomplishments meaningful. I will be forever grateful for my parents advice and understanding, my sisters admiration, my nephews humor, and my life-partners patience during this process.

I acknowledge Lloyd Edwards for being my academic and dissertation advisor. His advice assisted in the successful completion of the program. I owe much gratitude to Dr. Xi Chen for allowing me to grow into a professional researcher and collaborator. His advice and friendship will be something that I look forward to have throughout my career. I give many thanks to Dr. Peggye Dilworth-Anderson for her insight and tactful advice on how to communicate successfully and how to establish a prosperous career. Further, I am grateful for the opportunity for being awarded a pre-doctoral fellowship from the UNC Institute on Aging Carolina Program in Health and Aging Research (CPHAR) supported by the National Institute on Aging (NIA) Pre-doctoral/Post-doctoral traineeship (5T32AG000272). I must also thank the other members of my committee: Dr. John Preisser, Dr. Chirayth M. Suchindran, and Dr. Prenab K. Sen. They have each been more than willing to provide guidance as necessary.

From Duke University, I acknowledge Dr. Dan Blazer and Bruce Burchett for providing me with access to and support of the data used in this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: LITERATURE REVIEW

1.1 Introduction

Any research with aims to understand and describe the processes and mechanisms of change over a span of time must not only collect longitudinal data but must make use of some sort of longitudinal analysis strategy. Unlike cross-sectional studies that collect data at a single moment in time on each individual in a sample, longitudinal designs attempt to measure the same set of variables on a single cohort following a specified data-measuring schedule. Compared to cross-sectional analyses, longitudinal analyses can be more efficient, more robust to model selection, and have increased statistical power (Edwards, 2000). Although longitudinal analysis has its advantages over cross-sectional designs, it comes with a set of disadvantages that includes time constraints, lack of statistical methods, and dropout. Because of the potential for selection bias and its effects on external and internal validity, Norris (1985) and Markides et al. (1982) suggest that attrition could be the greatest threat to the analysis of longitudinal data.

Within the realms of gerontological research, specifically the study of aging and human development, longitudinal data analysis has proven to be the most productive approach (Alwin and Campbell, 2001). The shift in the focus of aging research from age associations (study of older adults) to the process of aging along with the advancement in longitudinal analytical methods and availability of computing tools has made longitudinal analysis more feasible. To emphasize further, there is an indisputable difference, especially in the older highly heterogeneous adult population,

between comparing cognitive function of 70-year-olds and 75-year-olds (*cross-sectional analysis*) and the change in cognitive function as a person ages from 70 to 75 (*longitudinal analysis*). Ferraro and Kelley-Moore (2003) revealed that the cross-sectional methodology was utilized as the main source of data analysis published in an aging journal, even though the studies acquired their data longitudinally. Yet, when studying a process such as aging or an outcome correlated with aging, subjects' temporal issues must be carefully collected and analyzed by methods that allow modeling of correlations and temporal changes.

Despite the benefits of longitudinal analysis for describing the aging process, longitudinal analysis is not absent of complications. Longitudinal data collection faces retention challenges that may lead to a type of selection bias known as attrition bias (Diggle and Kenward, 1994; Elias and Robbins, 1991; Little, 1995; Mcardle and Hamagami, 1992). Failing to account for nonrandom attrition and other mechanisms that affect the ability to acquire follow-up measurements may result in estimates on a relatively healthy or advantaged sample in terms of health and economic means. Miller and Wright (1995) explained that attrition can lead to bias in two ways–by altering the sample from the original intended sample and by affecting the covariance. In modeling the process of aging in older adults, the handling of attrition requires careful attention because attrition can affect the interpretation of the inference (Norris, 1985). Longitudinal studies of older adults are particularly sensitive to truncation due to death, which is usually the largest category of non-response in studies of older adults (Markides et al., 1982; Schaie, 1996; Rhodes, 2005).

Longitudinal studies of geriatric health outcomes with truncation due to death will most likely be biased if survival status is not taken into account. When investigators are interested in estimating the trajectory of an outcome that is not mortality but is highly predictive of death, not considering survival status could lead to incorrect inferences on

a majority healthy and alive sample. This has been appropriately termed the "healthy survivor" effect by Murphy et al. (2011). To avoid bias and misleading inferences about the change in a longitudinal outcome, the joint distribution of the longitudinal outcome and survival should be modeled.

Missingness due to non-response is different from censoring due to death, for those that die during a study will not have future responses (Dufouil et al., 2004). In these cases, methods such as imputation are not appropriate. Unfortunately, very few statistical methods exist for death that occurs during follow-up compared to those methods to accommodate missingness in follow-up due to non-response. The most recent literature on truncation due to death is focused on principal stratification (Frangakis and Rubin, 2002; Frangakis et al., 2007). Most recently, Kurland et al. (2009) proposed methods for analyzing longitudinal outcomes truncated by death, with an emphasis on matching the research question to the method and interpretation of the results. Absent from their evaluation of these models were discussions on bias, estimation, and efficiency of the methods. However, estimation and bias were examined by Kurland and Heagerty (2005) in some detail for the partly conditional model, which is also referred to as the regression conditioned on being alive (RCA) model. Understanding how bias can be introduced and correctly estimating uncertainty (standard errors) and the efficiency limits of the proposed regression models are important issues for longitudinal data analysis. This refined perspective provides a more thorough literature on regression models used to analyze missing outcomes due to death.

In the following subsection, notation for the general linear, general linear mixed, and generalized linear regression models are provided. These are all popular models used to analyze longitudinal data. Section 1.3 provides the background and study design for the data used to assess and compare the models in the present study. Section 1.4

introduces notation for repeated measures with missingness and discusses the nature of missing data assumptions in longitudinal analysis. Section 1.5 explores models that incorporate death in the mean model, and Section 1.6 offers a review of the literature.

1.2 Repeated Measures Models

This dissertation utilizes three different regression models: general linear model, general linear mixed model, and generalized linear model for longitudinal data. For each model let $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{ip_i})'$, $i = 1, 2, \ldots, N$ denote an $p_i \times 1$ vector of the responses for the ith subject that are independent and are assumed to be from a distribution belonging to the class of the exponential family distributions. Let $\boldsymbol{X}_i$ denote a $p_i \times q$ known design matrix of for the ith subject. Finally, let $\boldsymbol{\beta}$ be a $q \times 1$ vector of unknown population parameters. The notation for the general linear model for repeated measures data is given as

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i, \tag{1.1}$$

where $\boldsymbol{\varepsilon}_i$ is an $p_i \times 1$ vector of random variables with mean $\mathbf{0}_{(p_i \times 1)}$ and variance $\boldsymbol{\Sigma}_{\varepsilon_i} = var(\boldsymbol{y}_i) = \boldsymbol{V}_i$, an $p_i \times p_i$ matrix with elements of the form $var(\varepsilon_{it}) = \sigma_{Y_i,tt}$ and $cov(\varepsilon_{is}, \varepsilon_{it}) = \sigma_{y_i,st}$ such that $s \neq t$.

Whereas the general linear model is useful when estimating the population-average estimates for continuous outcomes, the general linear mixed model, detailed by Laird and Ware (1982), can be used to estimate subject-specific means for repeated continuous measures and is viewed as a special case of the general linear model. The notation is as follows:

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{d}_{\boldsymbol{i}} + \boldsymbol{e}_i, \tag{1.2}$$

with

$\boldsymbol{y}_i$	is the $p_i \times 1$ vector of outcome responses for the ith unit,
$\boldsymbol{X}_i$	is the $p_i \times q$ known design matrix for the fixed effects for subject i,
$\boldsymbol{\beta}$	is the $q \times 1$ vector of unknown fixed effects parameters,
$\boldsymbol{Z}_i$	is the $p_i \times m$ design matrix for the $(m \times 1)$ random effects, $\boldsymbol{d}_i$,
$\boldsymbol{d}_i$	is the subject–specific unknown parameters,
$\boldsymbol{D}$	is the $m \times m$ covariance matrix of the $(m \times 1)$ random effects, $\boldsymbol{d}_i$ (mutually independent),
$\boldsymbol{\Sigma}_{e_i}$	is the $p_i \times p_i$ covariance matrix for the random errors, $\boldsymbol{e}_i$ (mutually independent).

In this model, the random effects, $\boldsymbol{d}_i$, and the random errors, $\boldsymbol{e}_i$, are assumed to be independent for all $i = 1, \ldots, N$. For the purpose of estimation we assume that $\boldsymbol{d}_i \sim N(\boldsymbol{0}, \boldsymbol{D})$ and $\boldsymbol{e}_i \sim N(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_i)$, so that the $var(\boldsymbol{y}_i) = \boldsymbol{V}_i = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + \sigma^2 \boldsymbol{I}$.

The generalized linear model for repeated measures, introduced by Nelder and Wedderburn (1972) uses estimating equations, proposed by Zeger et al. (1988) to estimate population averages for repeated, non-normal outcomes. Taking $\boldsymbol{y}_i$ and $\boldsymbol{X}_i$ as described above, the general notation for the generalized linear model for the marginal mean of $\boldsymbol{y}_i$ given $\boldsymbol{X}_i$ is given as

$$g\left\{E(\boldsymbol{y}_i)\right\} = g(\boldsymbol{\mu}_i) = \boldsymbol{X}_i \boldsymbol{\beta} \tag{1.3}$$

where g is a one-to-one continuous differentiable function called a link function. The link function relates the means of the response to the linear predictors, $\boldsymbol{X}_i \boldsymbol{\beta}$. Let matrix $\boldsymbol{V}_i$ represent the estimate of the covariance matrix of $\boldsymbol{y}_i$ and $\boldsymbol{R}_i(\boldsymbol{\alpha})$ be an $p_i \times p_i$ "working" correlation matrix that is identified by the vector of parameters, $\boldsymbol{\alpha}$. Then

the covariance matrix of $\boldsymbol{y}_i$ is modeled as,

$$\boldsymbol{V}_i = \phi \boldsymbol{A}_i^{\frac{1}{2}} \boldsymbol{R}_i \left(\boldsymbol{\alpha}\right) \boldsymbol{A}_i^{\frac{1}{2}},$$

where $\boldsymbol{A}_i$ is an $(p_i \times p_i)$ diagonal matrix with a variance function that is determined by the assumed probability distribution of the outcomes, $var(\mu_{ij})$, as the jth diagonal element and ϕ is a dispersion parameter that may be known or may be estimated from the data dependent upon the distribution assumption. The generalized linear model allows for the distributions of the errors to be non-normal. Further, these models focus on the estimation of the average response over the population rather than regression parameters.

1.3 Missing Data in Longitudinal Studies

1.3.1 Overview

By introducing a *data model* and a *non-response model*, we can analytically explain the effects of missing data in the analysis of longitudinal data (Laird, 1988). We will limit our overview to non-response in the outcome only and not within covariates. Using similar notation as before, let $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \ldots, y_{ip})'$, $i = 1, 2, \ldots, N$ denote a $p \times 1$ vector of the responses for the ith subject. We let $\boldsymbol{X}_i$ denote a $p \times q$ matrix of covariates for the ith subject, which contains both individual covariates and the design on time. This matrix is routinely denoted as the design matrix. Finally, we let $\boldsymbol{\beta}$ be a $q \times 1$ vector of unknown parameters and $\boldsymbol{\varepsilon}_i$ be a $p \times 1$ vector of random variables with mean $\mathbf{0}_{(p\times 1)}$ and variance $\boldsymbol{\Sigma}_{y_i}$ a $p \times p$ matrix with elements of the form $var(\varepsilon_{it}) = \sigma_{y_i,tt}$ and $cov(\varepsilon_{is}, \varepsilon_{it}) = \sigma_{y_i,st}$. Hence, the linear model for subject i takes the form,

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon}_i,$$

where $E(\boldsymbol{y}_i) = \boldsymbol{X}_i\boldsymbol{\beta}$ and $var(\boldsymbol{y}_i) = \boldsymbol{\Sigma}_{y_i}$ is the matrix of covariance parameters. The specification of the data model is completed by noting $f(\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{\beta})$ is the multivariate density of $\boldsymbol{y}_i$ conditional on $\boldsymbol{X}_i$ and $\boldsymbol{\beta}$, where inference interests are in the components of $\boldsymbol{\beta}$ and $var(\boldsymbol{y}_i) = \boldsymbol{\Sigma}_{y_i}$.

The *non-response model* is formed by letting $\boldsymbol{R}_i = (R_{i1}, R_{i2}, \ldots, R_{ip})'$ denote a $p \times 1$ vector of indicator variables for the ith subject, such that $R_{it} = 1$ if y_{it} is observed, and $R_{it} = 0$ otherwise. Let $\boldsymbol{\nu}$ denote the vector of parameters of the non-response model. The model is completed by denoting $f(\boldsymbol{R}_i|\boldsymbol{y}_i, \boldsymbol{X}_i, \boldsymbol{\nu})$ as the multivariate density of $\boldsymbol{R}_i$ given $\boldsymbol{y}_i, \boldsymbol{X}_i$, and $\boldsymbol{\nu}$. The non-response model does not describe the reasons or the processes that lead to the missing outcome variables; instead, the non-response model is a probabilistic selection mechanism given the outcome variables and covariates, $(\boldsymbol{y}_i, \boldsymbol{X}_i)$ that is central in understanding, developing, and applying modern missing data methods (Rathouz and Preisser, 2013).

Using the notation above and following the discussion from Little and Rubin (2002), we can define the complete data likelihood as

$$f(\boldsymbol{y}_i, \boldsymbol{R}_i|\boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{\nu}) = f(\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{\beta}) f(\boldsymbol{R}_i|\boldsymbol{y}_i, \boldsymbol{X}_i, \boldsymbol{\nu}). \tag{1.4}$$

The denotation of $\boldsymbol{R}_i$ allows us to partition the response vector into two components, $\boldsymbol{y}' = (\boldsymbol{y}_i^o,\ \boldsymbol{y}_i^m)$, $\boldsymbol{y}_i^o$ for the responses that are observed ($R_{it} = 1$) and $\boldsymbol{y}_i^m$ for the responses that are not observed ($R_{it} = 0$). Naturally, the dimensions of $\boldsymbol{y}_i^o$ and $\boldsymbol{y}_i^m$ may vary for each subject. Using the established notation, the density of the observed data is given as

$$f(\boldsymbol{y}_i^o, \boldsymbol{R}_i|\boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{\nu}) = \int f(\boldsymbol{y}_i^o, \boldsymbol{y}_i^m, \boldsymbol{R}_i|\boldsymbol{\beta}, \boldsymbol{\nu}, \boldsymbol{X}_i) d\boldsymbol{y}_i^m, \tag{1.5}$$

where integration is over the sample space of $\boldsymbol{y}_i^m$. Using the notion from equation (1.4),

the equation (1.5) can be expressed as

$$f(\boldsymbol{y}_i^o, \boldsymbol{R}_i|\boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{\nu}) = \int f(\boldsymbol{y}_i^o, \boldsymbol{y}_i^m|\boldsymbol{X}_i, \boldsymbol{\beta}) f(\boldsymbol{R}_i|\boldsymbol{y}_i^o, \boldsymbol{y}_i^m, \boldsymbol{X}_i, \boldsymbol{\nu}) d\boldsymbol{y}_i^m, \tag{1.6}$$

with integration over the sample space of $\boldsymbol{y}_i^m$.

1.3.2 Missing Data Assumptions for Outcome Variables, $\boldsymbol{y}$

Rubin (1976) introduced and Laird (1988) discussed a missing data hierarchy. This hierarchy helps illustrate more easily the effects of the non-response model in likelihood-based inference analysis. In the *missing at random* (MAR) scenario, the probability of the non-response process is not dependent on $\boldsymbol{y}_i^m$ given $\boldsymbol{y}_i^o$. That is, we assume

$$f(\boldsymbol{R}_i|\boldsymbol{y}_i^o, \boldsymbol{y}_i^m, \boldsymbol{X}_i, \boldsymbol{\nu}) = f(\boldsymbol{R}_i|\boldsymbol{y}_i^o, \boldsymbol{X}_i, \boldsymbol{\nu}). \tag{1.7}$$

By substituting (1.7) in (1.6) and integrating, the observed data density becomes

$$f(\boldsymbol{y}_i^o, \boldsymbol{R}_i|\boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{\nu}) = f(\boldsymbol{R}_i|\boldsymbol{y}_i^o, \boldsymbol{X}_i, \boldsymbol{\nu}) f(\boldsymbol{y}_i^o|\boldsymbol{X}_i, \boldsymbol{\beta}). \tag{1.8}$$

A stronger assumption than MAR is *missing completely at random* (MCAR). Data are said to missing completely at random when the non-response mechanism is independent of both the observed and the missing values of the outcome, ($\boldsymbol{y}$). That is,

$$Pr(\boldsymbol{R}_i|\boldsymbol{y}_i^o, \boldsymbol{y}_i^m, \boldsymbol{X}_i) = Pr(\boldsymbol{R}_i).$$

Essentially, the observed data can be considered a random sample of the population. Consequently, in general, any methods of analysis that are valid on the complete dataset will yield valid inference when the analysis is based on only observed data.

Because the missing-mechanism is independent of those observations that are

missing from the intended complete data, the parameters of the outcome model, $\boldsymbol{\beta}$, and non-response model, $\boldsymbol{\nu}$, are distinct, and MCAR or MAR data are referred to as ignorable. This ignorability speaks to the fact that MCAR and MAR data can ignore $Pr(\boldsymbol{R}_i|\boldsymbol{y}_i, \boldsymbol{X}_i)$ and obtain a valid likelihood-based analysis, provided the model for $f(\boldsymbol{y}_i|\boldsymbol{X}_i)$ is correctly specified.

Missing data where $(\boldsymbol{R}_i|\boldsymbol{y}_i^o)$ is related to or depends on some components of $\boldsymbol{y}_i^m$ is referred to as *non-missing at random* (NMAR) or *non-ignorable* missingness. To obtain valid inference, methods of analysis on data with NMAR require the specification of a model for the missing mechanism. The distribution of $\boldsymbol{y}_i^m$ is not the same for the completers or the target population. Instead, the distribution of $\boldsymbol{y}_i^m$ depends on $\boldsymbol{y}_i^o$ and $Pr(\boldsymbol{R}_i|\boldsymbol{y}_i, \boldsymbol{X}_i)$, which makes modeling and including the missing mechanism in analysis critical and necessary for valid inferences. Any assumptions made about the missingness process for NMAR data are wholly unverifiable from the observed data. Therefore, many authors stress the importance of conducting sensitivity analyses.

Some studies have variables observed for all subjects that could be used to denote the history of the change, presence, or absence of outcome variables. These variables are typically not part of the primary inference of the analysis and are predictive of the missing response values. Such variables are known as auxiliary variables. In the presence of auxiliary variables, $\boldsymbol{\Psi}_i$, the MAR assumption requires that the missing mechanism is independent of the missing responses given $(\boldsymbol{y}_i^o, \boldsymbol{X}_i, \boldsymbol{\Psi}_i)$. Similarly, the more stringent assumption, MCAR, requires the missing mechanism to be independent of $(\boldsymbol{y}_i, \boldsymbol{X}_i, \boldsymbol{\Psi}_i)$, when auxiliary information is present. Although auxiliary data can be helpful in meeting the MAR assumption, missingness due to MAR is not truly ignorable unless the missingness only occurs in the response variable, there exist no auxiliary information, and full likelihood analyses, $(\boldsymbol{y}_i^o|\boldsymbol{X}_i)$, are pursued.

1.4 Incorporating Death in Mean Models

1.4.1 Overview

Little and Rubin (2002) and Little (1995) discussed two general classes of factorizations of the joint model $(\boldsymbol{y}, \boldsymbol{R})$, selection models, $p(\boldsymbol{y}, \boldsymbol{R}|\boldsymbol{\beta}, \boldsymbol{\nu}) = p(\boldsymbol{R}|\boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\nu})p(\boldsymbol{y}|\boldsymbol{\beta})$, where $p(\boldsymbol{y}|\boldsymbol{\beta})$ represents the model of the complete data and $p(\boldsymbol{R}|\boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\nu})$ represents the missing data mechanism; and pattern-mixture models, $p(\boldsymbol{y}, \boldsymbol{R}|\eta, \pi) = p(\boldsymbol{y}|\boldsymbol{R}, \boldsymbol{\eta})p(\boldsymbol{R}|\boldsymbol{\pi})$, where $\boldsymbol{y}$ is conditioned on the missing data pattern $\boldsymbol{R}$. Allowing survival, $\boldsymbol{S}$, to represent survival time such as age at death or weeks from baseline until death, the joint distribution $f(\boldsymbol{y}_i, \boldsymbol{S}_i)$ denotes the probability that subject i's outcomes takes a vector of specific values and survives to a specific time, s. In regression models that describe the relationship of predictors and the longitudinal outcomes, survival must be either implicitly or explicitly modeled. The joint probability can be factored in two ways: $f(\boldsymbol{y})f(\boldsymbol{S}|\boldsymbol{y})$ and $f(\boldsymbol{y}|\boldsymbol{S})f(\boldsymbol{S})$. Depending on how or if the longitudinal outcome conditions on survival status, $\boldsymbol{S}$, the regression analysis of $\boldsymbol{y}$ can be categorized as being unconditional, fully conditional, partly conditional, or joint. When deciding which regression analysis should be considered for analysis of longitudinal data with follow-up truncated by death, Kurland et al. (2009) urged investigators to match analysis methods to research aims, for each method's target population of inference are different for each model. Each method and its target population is summarized in Table 1.1 and described in the sections that follow.

1.4.2 Unconditional Models: $f(\boldsymbol{y}_i)$

Unconditional models are useful if deaths do not occur or if deaths do not result in truncation. Considering these models would be appropriate if a researcher's question of interest is on the expected changeover time of a response in an immortal cohort or

if death does not affect the outcome. The estimation methods for these likelihood-based models implicitly impute values for those who die (Laird, 1988). Because of this fact, this method is typically not useful in gerontological research studies that are interested in the change of an outcome over time at the subject level. Yet, if the outcome of interest is focused on phenomena such as the rate of decline, recurrence, or other change following some action of a biological substance that can be collected at baseline and tested over time without requiring additional collections, then survival would not affect the outcome and the unconditional model would be a reasonable approach. Because unconditional models are assuming that death does not occur or that death does not cause truncation, the missing mechanism, survival, can be ignored without compromising the validity of the inference. This scenario follows a situation that is modeled by $f(\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{\beta}) = \int_{\boldsymbol{R}_i} f\boldsymbol{y}_i|\boldsymbol{X}_i, \boldsymbol{\beta})d\boldsymbol{y}_i^m = f(\boldsymbol{y}_i^o|\boldsymbol{X}_i, \boldsymbol{\beta})$.

1.4.3 Fully Conditional Models: $f(\boldsymbol{y}_i|\boldsymbol{S}_i = s)$

Fully conditional mean models for $\boldsymbol{y}$ given $\boldsymbol{S} = s$ follow the pattern-mixture factorization of the joint distribution of $(\boldsymbol{y}, \boldsymbol{S})$, $f(\boldsymbol{y}, \boldsymbol{S}) = f(\boldsymbol{y}|\boldsymbol{S})f(\boldsymbol{S})$. In these models, inference regards the changing-over time of the longitudinal outcome variable stratified by the time of subjects' deaths.

Pattern-Mixture

Typically, pattern-mixture models are not as popular as selection models because they do not directly model the marginal distribution of $\boldsymbol{y}$ (Little, 1993). However, when analyzing a longitudinal response with non-ignorable missing data due to death, pattern-mixture models are favored over selection models. In this setting of missingness due to death, pattern-mixture models can be completely identifiable by introducing a categorical variable in the main effects model that denotes the different strata for time

of death (Ribaudo et al., 2000; Pauler et al., 2003). Consequently, analysis will yield a mixture of distributions of the longitudinal response outcome. Each stratum will have its own trajectory of the response outcome. An advantage of this approach is accurate representation of individuals' responses over time.

In order to better understand the nature of the proposed pattern-mixture models, let's first examine the notation of the general pattern-mixture model as described by Little (1993). First, assume that there are $q_0, q_1, \ldots, q_L$ missing patterns in a population and let q_0 represent the pattern with complete responses. Let r_i take the value r for missing pattern q_r, and let n_r equal the number of subjects with q_r missing pattern such that $\Sigma_{r=0}^{L} n_r = N$ (total number of subjects). Now we have that r_i follows a multinomial distribution with probability $p(r_i = r) = \pi_r$, $r = 0, 1, \ldots, L$. Finally, we can represent the distribution of $\boldsymbol{y}_i$ as,

$$f(\boldsymbol{y}_i|\boldsymbol{r}_i, \boldsymbol{\vartheta}^{(r)}) = f(\boldsymbol{y}_{i,o}^{(r)}|r_i = r, \boldsymbol{\vartheta}_o^{(r)}) f(\boldsymbol{y}_{i,m}^{(r)}|r_i = r, \boldsymbol{y}_{i,o}^{(r)}, \boldsymbol{\vartheta}_{m,r\cdot o,r}^{(r)}). \tag{1.9}$$

$\boldsymbol{y}_{i,o}^{(r)}$ represent the observed responses in pattern q_r and $\boldsymbol{y}_{i,m}^{(r)}$ represent the missing response variables in pattern q_r. The parameters $\boldsymbol{\vartheta}_o^{(r)}$ and $\boldsymbol{\vartheta}_{m,r\cdot o,r}^{(r)}$ are functions of $\boldsymbol{\vartheta}^{(r)}$ and are assumed to be distinct for all values of r. Because death is a form of monotone missingness, the analysis within the patterns can ignore the non-response mechanism if separate analyses are conducted for each missing pattern.

Principal Stratification

Another fully conditional model is principal stratification. This method describes the average causal effects for selected principal strata defined by potential survival outcomes (Frangakis and Rubin, 2002; Hayden et al., 2005; Egleston et al., 2007, 2009). In principal stratification models, the response is estimated only in the strata of individuals expected to live for a predetermined time, s, regardless of exposure. Unlike

pattern-mixture models, principal stratification not only conditions on actual survival, it also conditions on counterfactual survival status. The attractiveness of this method is that the inference is on the principal strata that would live regardless of exposure or treatment, allowing for the separation of the effect of the exposure and death from the effect of the exposure and the outcome. This approach is most useful in analyzing treatment and intervention effects in randomized clinical trials designs. However, this approach requires many untestable assumptions about the counterfactual information that is not collected.

The notation for principal stratification involves a vector of covariates, $\boldsymbol{X}_i$, a manipulable exposure variable to which subjects can be randomized, $Z_i = z$, a survival indicator for a subject at each exposure level, $D_i(z)$, such that $D_i(z) = 0$ represents survival at exposure z, an indicator variable R_i for signifying if a subject reaches the end of the study, ($R_i = 1$ if not lost to follow-up; 0 otherwise), and the outcome for a subject at each exposure level, $Y_i(z)$. In this model, the interest lies in the estimate of the association of Z_i and $Y_i(z)$ in the stratum of patients that will survive regardless of the value of z, $D_i(z) = 0$ (alive) for all z, which can be assessed by estimating the unidentifiable survivor average causal effect (SACE), which is defined as

$$\mu = E\left\{Y_i(1)|D(z) = 0\right] - E\left[Y_i(0)|D(z) = 0\right\}.$$

Terminal Decline

A third fully conditional model uses a time scale that counts backwards from death instead of forward in years. This model is useful for measuring the "dying process" and thus utilizes the responses of decedents only (Siegler, 1975; Diehr et al., 2002; Wilson et al., 2003). Terminal decline is attractive when the researchers interest is in changes related to the imminence of death versus changes due to aging.

This fully conditional model is similar to the pattern-mixture model in that the missing pattern category determines the length of outcome vector. Unlike the pattern-mixture model, the terminal decline model allows the trajectory of the outcome over the new time scale to be estimated for the entire sample.

1.4.4 Partly Conditional Models: $f(\boldsymbol{y}_i|\boldsymbol{S}_i > s)$

Partly conditional models estimate the mean of the response conditioned on each subject being alive beyond time s. These models are different from the unconditional case where the analysis methods model the correlation structure of the repeated data implicitly and impute missing data without any differentiation between dropout due to death and dropout due to other reasons. In order to avoid this forced imputation, partly conditional models are estimated by treating longitudinal data as independent. Kurland and Heagerty (2005) call the partly conditional method "regression conditioning on being alive" (RCA). This method describes the dynamic cohort of survivors and models the change in the prevalence of the outcome among survivors at each measurement occasion.

As mentioned above, likelihood based approaches cannot directly estimate or parameterize partly conditional means; instead, the models are fit using generalized estimating equations (GEE) (Liang and Zeger, 1986) with independence working correlation. This analysis should yield consistent estimation as long as the model is correctly specified (Crowder, 1986).

1.4.5 Joint Models: $f(\boldsymbol{y}_i, \boldsymbol{S}_i)$

Although general pattern-mixture and selection models begin as joint models, their inference interests are in either the marginal or the conditional means of the longitudinal response. Joint models encompass the repeated response as well as survival data. Diehr

et al. (1995) introduced a joint model by defining the probability of being alive and healthy (PAH) and a related method to predict the PAH for a prescribed amount of time. Johnson (2002) models the PAH generally as

$$PAH(s) = P(Q(s) > q, S > s) = P(Q(s) > q|S > s)P(S > s), \tag{1.10}$$

where $S > S$ represents being alive at time s and $Q(s) > q$ represents being healthy at time s. Equation (1.10) has a very similar structure to the general pattern-mixture model and can be seen as a special case of the pattern-mixture model. However, unlike a pattern-mixture model that locks participants in specific strata, the PAH model allows subjects to move from being alive and healthy to being alive and unhealthy and vice versa. Subjects are not allowed to transition out of the dead strata once they have entered it.

1.5 Data: NC EPESE

In 1980, the Epidemiology, Demography, and Biometry Program (EDBP) initiated the "Established Populations for Epidemiologic Studies of the Elderly" (EPESE) project in order to conform to the mandate to authorize the planning, initiation, direction, coordination, and analysis of longitudinal epidemiologic studies of specific diseases and conditions affecting the elderly. Some of the most prominent purposes of the EPESE project are to study risk factors for chronic diseases in the elderly and to identify predictors of mortality, hospitalization, and placement in long-term care facilities. Specifically, the project was designed to produce estimates of the prevalence and incidence of chronic conditions, impairments, and disabilities with their associated risk factors, and to quantify the changes in these characteristics and the general functioning of individuals. EPESE results were expected to affect policies on

illness prevention practices and to lengthen the time older adults can live independently in their own homes.

Funded by the National Institute on Aging (NIA), North Carolina Established Populations of Epidemiological Studies of the Elderly (NC EPESE), officially known as the Piedmont Health Survey of the Elderly, was the fourth site added to the larger multi-center prospective population-based epidemiologic study of health status and the physical, social, and cognitive functioning of persons 65 years of age and older living in communities. An additional major goal for the data collected at the North Carolina EPESE centers was to study racial difference in mortality and health of older persons.

Established in 1986, the North Carolina cohort was a sample of 4,162 persons 65 years or older residing in households in Durham, Warren, Franklin, Granville, and Vance counties (one urban county, four rural) in the Central Piedmont area of North Carolina. The site was over 50% black and the geographic area selected was diverse, allowing both racial and urban/rural comparisons to be made regarding the distribution of certain risk factors and disease. Of the 4,162 subjects selected on the basis of a four-stage, race-stratified sampling design, 48% (including similar proportions of blacks and whites) lived in an urban setting. Participants were surveyed in person on four occasions: Wave 1 (1986-1987); Wave 2 (1989-1990); Wave 3 (1992-1993); and Wave 4 (1996-1997). At each of these waves, depression symptoms, blood pressure, and physical functioning level were among the outcomes that were measured.

The measure of depression used was the CES-D, a self-report index of depressive symptoms developed by the Center for Epidemiological Studies of the National Institute of Mental Health (Radloff, 1977). This index consists of 20 statements, each describing a symptom or absence of a symptom. Whereas the CES-D in its original form permitted graded responses for each item, the modification used in this survey allowed only two responses, ("Yes" or "No"), scored 1 or 0, respectively (Blazer et al., 1991). Blazer

et al. (1991) justified a CES-D score of 9 or more to be sufficient for categorizing those subjects who are pre-screened for being clinically depressed versus the 16 score cut-off established by Radloff (1977).

Blood pressure of all participants was measured by trained interviewers using the Hypertension Detection and Follow-Up Program protocol (1978). Participants were seated and a standard mercury column sphygmomanometer was employed. Two blood pressure measurements were taken. The outcome of interest for blood pressure for this dissertation was the average of these two measurements. We note here that nearly all the subjects were on a medication regiment to normalize blood pressure.

One of the measures of physical functioning measured in the NC EPESE was seven activities of daily living (ADL): bathing, dressing, walking, grooming, transferring from bed to chair, eating, and toileting (Katz et al., 1970; Branch et al., 1984). For each activity, it was denoted whether or not assistance was needed. The number of activities requiring any level of assistance became the physical function score for each participant.

To date, very little literature exists on the patterns and progression of depression scores, especially among older adults. Most of the literature is on the progression of the diagnosis of those older adults who are depressed (Kuchibhatla et al., 2012). Similarly, among the published articles using EPESE data, there are not many articles that provide a longitudinal account of systolic and diastolic blood pressures. The lack of these types of analyses has led to many conflicting results. The variable nature of blood pressure measurements and the many factors influencing their values make interpretation of cross-sectional results questionable and limited. Measuring physical functional independence with ADLs, li (2005), revealed evidence that supported significantly greater change in functional dependency for those who died during the study than those who remained in the study or dropped out. Participants who remained in the study or left the study both had a steady mean of mild physical functional

dependency. Combined, these three health responses represent different categories of the overall health, vitality, and independence of older adults.

1.6 Summary

Collecting longitudinal data on older populations leads to greater risks of truncation due to death. Analysis of changes in responses truncated by death is likely to be biased and thus survival should necessarily be considered in the analysis. The NC EPESE studies that were initiated by the National Institutes on Aging (NIA) to estimate the incidences and prevalences of health conditions and to uncover predictors and correlates of death and diseases should be analyzed with the most accurate techniques. Further, this study that followed the health of certain cohorts for 10 years could benefit from new techniques to incorporate death information when the interest is in the mean change over time of a morbidity outcome that is truncated by death.

In the literature there exist discussions and suggestions on incorporating missingness due to death in mean regression models. These models are assumed to be unbiased for the estimands and highlight the correct interpretation for each proposed model. Even though bias (on $\boldsymbol{\beta}$), estimation, and efficiency, $V(\hat{\boldsymbol{\beta}})$, are important components of data analysis, they have not been clearly reviewed and discussed for these models. In an attempt to provide guidelines for those analyzing gerontological longitudinal data in the presence of death, these components deserve attention and exploration.

This dissertation assesses the performance of the proposed models (unconditional, pattern-mixture, principal stratification, terminal decline, partly conditional, and joint model) for truncated longitudinal outcomes by analyzing a cohort from a well-known longitudinal study of older adults, comparing the changes in the results for different types of missing data, and assessing the effects of varying percentages of missingness due to death on bias and efficiency of the models via simulations. Chapter 2 presents

estimates of the rates of change of the four outcomes from the NC EPESE as estimated by the proposed models, when non-response is due to death only. Chapter 3 discusses the effects of the models when death and non-participation are both present. Chapter 4 provides simulations to evaluate bias and efficiency of the models under varying missing assumptions. Finally, a discussion and suggestion of needed avenues of future work are offered in Chapter 5.

Table 1.1: Summary of statistical regression models for longitudinal response and survival and its population of inference

Regression Methods	**Population of Inference**	**Research Aim**
A. Unconditional Models	An immortal cohort where subjects are expected to die or where death does not induce missingness	What is the longitudinal effect of an outcome on an immortal cohort
B. Fully Conditional Model: Pattern Mixture	Cohorts created by their survival status	What is the longitudinal change in an outcome for different survival cohorts
C. Fully Conditional Model: Principal Stratification	Cohort of those that will survive for s years regardless of treatment/exposure	What is the expected difference in an outcome for different levels of treatment/ exposure among subjects surviving a given time
D. Fully Conditional Model: Terminal Decline	Decedents	What is the behavior of a longitudinal outcome as subjects nears death
E. Partly Conditional Model	Dynamic cohort of sample survivors	What is the longitudinal trend of an outcome from a dynamic cohort
F. Joint Model	Complete mortal sample	What percentage of subjects are alive and healthy over time

CHAPTER 2: ANALYSIS OF NC EPESE TO INCORPORATE SURVIVAL IN THE ANALYSIS OF OUTCOMES TRUNCATED DUE TO DEATH

2.1 Introduction

The North Carolina Established Populations of Epidemiological Studies of the Elderly (NC EPESE) was a prominent observational prospective study that has been utilized to provide the narrative of incidences and prevalences of chronic illness, cognitive and physical impairments, and other disabilities, along with their risk factors and the changes of these characteristics of older adults as they age in the community. Analyses of this study and its sister studies have informed the needs of health care services for the prevention of illnesses plaguing adults in later life and strategies for maintaining function of older adults aging outside of health care facilities. Since the end of 2012, there have been 341 publications in the form of manuscripts, letters, and books that referenced any of the EPESE studies. From 1996 to the end of 2012, there have been 90 publications on the NC EPESE, and only 6 of the publications used a form of longitudinal methodology for its primary statistical analysis. None of the 6 publications made any distinctions between missing not due to death and missing due to death.

The six methods for incorporating death, examined by Kurland et al. (2009), have never been used to analyze the EPESE data. In addition, there have not been discussions or comments pertaining to the effect of death on bias, estimation, and efficiency in the proposed regression models. Without a full description of the

strengths and weakness of each proposed method, an understanding of the role of death in conditions that affect physical function, quality of life, and self-sufficiency are incomplete. Incomplete knowledge of the proposed methods could lead to erroneous conclusions that could misinform important policy affecting the elderly.

2.2 Effect of Death in NC EPESE

For the purpose of this dissertation, 26 of the subjects were excluded from the original 4,162 participants because they identified themselves as other than black or white. Another four subjects were not included because their ages were less than 65 at the baseline survey. Because of a focus on mortality, dates of death were collected throughout the data collection for NC EPESE and continued 10 years after the study ended. During the 10-year study period, 2,046 of the 4,132 (49.5%) eligible participants had death dates before the end of the study. These participants were 40.5% male and 55.5% black. At the time of death, the average age of the participants was 79.22 (SD=7.47) years old, and the average number of years to death from baseline was 5.09 (SD=4.52). This dissertation is interested in examining the role of death on the modeling of the mean and covariance models when longitudinal data are truncated due to death in the NC EPESE. To begin this analysis, we first assume a situation where data are only missing due to death. For this hypothetical situation, missing data from subjects because of non-response will require imputation. Hence, the NC EPESE dataset was altered by imputing these values for the outcomes of interest (depression scores, blood pressures, and ADL score) for those subjects who had missing values due to non-response. Single imputation was completed using proc mi in SAS 9.2 (Yang, 2002). Figure 2.1 describes the "completers" (no imputation required) as those subjects who survived beyond the study and provided complete data for each outcome of interest for each measurement occasion and those subjects who died before the end of the study

but provided complete information up to their deaths.

The effect of death in the modified NC EPESE was graphically assessed by dividing the decedents into two subgroups by baseline age–those greater than and equal to 85-years-old and those younger than 85. The population-expected means for the depression scores, the systolic and diastolic blood pressures, and the physical function scores up to 10 years prior to death were plotted. To serve as a reference, the end-of-study survivors were divided similarly into two subgroups and with their outcome trajectories graphed with their decedent counterparts.

In Figure 2.2, the decedents and survivors have similar mean CES-D scores 10 years from death (end of study). As the decedents approached death, their trajectories increased more sharply than the survivors' trajectories in both cohorts. Although both trends increased over time, the trajectories were statistically different ($p < 0.001$) for those less than 85-years-old and those greater than or equal to 85.

The effect of death in Figure 2.3 for systolic blood pressure is not as apparent as in the depression data. Nonetheless, the systolic blood pressure trend for the survivors younger than 85 were nearly constant whereas the decedents experienced a decline. Both groups declined in the 85 and older graph and were not statistically different.

The population-expected mean diastolic blood pressure (Figure 2.4) for the survivors had a steeper decline over the 10-year period than the decedents in both age groups. The trends for the survivors and decedents were similar in each graph, albeit the trend for decedents 85-years-old and older was shifted down approximately two units.

Physical functional is known to be highly predictive of death. The results in Figure 2.5 supports this relationship. As decedents get closer to death, the number of ADLs accomplished alone decreased. If the decedents are older than 85, this trend is more severe.

All of these non-mortality outcomes have some association with death and make it

difficult to correctly analyze and interpret findings without considering survival. The six regression models introduced by Kurland et al. (2009) are options for analyzing the outcomes with a treatment of survival.

2.3 Analysis of Modified NC EPESE

The unconditional, pattern-mixture models and terminal decline models (Kurland et al., 2009) for each outcome were fitted using a linear mixed model with a random intercept and slope as described by Laird and Ware (1982). The standard linear mixed effect model is written as

$$\boldsymbol{y}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{d}_i + \boldsymbol{e}_i \tag{2.1}$$

where $\boldsymbol{y}_i$ is a $p_i \times 1$ of observations on person i for $i = 1, \ldots, N$; $\boldsymbol{X}_i$ is a $p_i \times q$ known, constant design matrix for person i; $\boldsymbol{\beta}$ is a $q \times 1$ vector of unknown, constant population parameters; $\boldsymbol{Z}_i$ is a $p_i \times 2$ known and constant design matrix for person i; $\boldsymbol{d}_i = \begin{pmatrix} d_{i0} \\ d_{i1} \end{pmatrix}$ is the corresponding 2×1 vector of unknown random effects (random intercept and slope); and $\boldsymbol{e}_i$ is a $p_i \times 1$ vector of unknown random errors. Vectors $\boldsymbol{d}_i$ and $\boldsymbol{e}_i$ are assumed to be from a Gaussian distribution and independent with mean $E(\boldsymbol{d}_i) = \boldsymbol{0}$ and $E(\boldsymbol{e}_i) = \boldsymbol{0}$ and $var(\boldsymbol{d}_i) = \boldsymbol{D}$, where $\boldsymbol{D} = \begin{pmatrix} \sigma^2_{d_0} & \sigma_{d_0 d_1} \\ \sigma_{d_0 d_1} & \sigma^2_{d_1} \end{pmatrix}$ and $var(\boldsymbol{e}_i) = \boldsymbol{\Sigma}_{e_i} = \sigma^2\boldsymbol{I}_i$. Each model assumed homogenous variance for subjects measurements across time with no expected correlation between the measurements for all subjects, $\sigma^2\boldsymbol{I}_i$, (conditional on the random effects) and allowed the random effects to be independent with unique variances. Hence $var(\boldsymbol{y}_i) = \boldsymbol{V}_i = \boldsymbol{Z}_i\boldsymbol{D}\boldsymbol{Z}_i^T + \sigma^2\boldsymbol{I}_i$.

To estimate these models when data are not complete, an E-M algorithm proposed by Dempster et al. (1977) and used by Laird and Ware (1982) and Dempster et al.

(1981b) may be used to obtain maximum likelihood estimates of $\boldsymbol{\beta}$, $\boldsymbol{d}_i$, $\boldsymbol{D}$, and σ^2. To see this, we take $\boldsymbol{\theta}$ to the vector of the covariance components and then note that the estimates for the fixed and random effects when variance is unknown is given as

$$\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}) = \left(\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{W}}_i \boldsymbol{X}_i \right)^{-1} \sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{W}}_i \boldsymbol{y}_i \tag{2.2}$$

and

$$\hat{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}) = \hat{\boldsymbol{D}} \boldsymbol{Z}_i^T \hat{\boldsymbol{W}}_i (\boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})). \tag{2.3}$$

These are the weighted least square equations with estimates for $\hat{\boldsymbol{W}}_i = \hat{\boldsymbol{V}}_i^{-1}$ with $\hat{\boldsymbol{V}}_i = \hat{\boldsymbol{\Sigma}}_{e_i} + \boldsymbol{Z}_i \hat{\boldsymbol{D}} \boldsymbol{Z}_i^T$, where $\hat{\boldsymbol{\Sigma}}_{e_i}$ is the estimate of the variance-covariance matrix of $\boldsymbol{e}_i$. The variances of these estimates are defined as

$$\hat{var}\left[\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}) \right] = \left(\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{W}}_i \boldsymbol{X}_i \right)^{-1} \tag{2.4}$$

and

$$\hat{var}\left[\hat{\boldsymbol{d}}(\hat{\boldsymbol{\theta}}) \right] = \hat{\boldsymbol{D}} \boldsymbol{Z}_i^T \left\{ \hat{\boldsymbol{W}}_i - \hat{\boldsymbol{W}}_i \boldsymbol{X}_i \left(\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{W}}_i \boldsymbol{X}_i \right)^{-1} \boldsymbol{X}_i^T \hat{\boldsymbol{W}}_i \right\} \boldsymbol{Z}_i \hat{\boldsymbol{D}} \tag{2.5}$$

If $\boldsymbol{d}_i$, $\boldsymbol{e}_i$, and $\boldsymbol{y}_i$ were to be observed, then closed-forms of the maximum likelihood estimates of Σ_{e_i} and $\boldsymbol{D}$ based on quadratic forms in $\boldsymbol{d}_i$ and $\boldsymbol{e}_i$ for $i = 1, \ldots, N$ can be obtained. For the variance structure assumed for the linear mixed models in this analysis ($var(\boldsymbol{d}_i) = \boldsymbol{D}$ a 2×2 nonnegative definite matrix and the $var(\boldsymbol{e}_i) = \boldsymbol{\Sigma}_{\boldsymbol{e_i}} = \sigma^2 \boldsymbol{I}_i$) these estimates would take the form

$$\hat{\sigma}^2 = \sum_{i=1}^{N} \boldsymbol{e}_i^T \boldsymbol{e}_i / \sum_{i=1}^{N} p_i = t_1 / \sum_{i=1}^{N} p_i \tag{2.6}$$

and

$$\hat{\boldsymbol{D}} = N^{-1} \sum_{i=1}^{N} \boldsymbol{d}_i \boldsymbol{d}_i^T = \boldsymbol{t}_2 / N. \tag{2.7}$$

The equations above show that the sufficient statistics of the covariance components are t_1 and the non-redundant components of the vector $\boldsymbol{t}_2$. An estimate of $\boldsymbol{\theta}$ could be used to approximate the estimates of the missing sufficient statistics by setting the sufficient statistics to their expectations given the observed outcome vector, $\boldsymbol{y}_i$. Before these equations can be denoted, we must define $\hat{\boldsymbol{\theta}}$, $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})$ and $\hat{\boldsymbol{d}}_i(\hat{\boldsymbol{\theta}})$ to be estimates of $\boldsymbol{\theta}$, $\boldsymbol{\beta}$, and $\boldsymbol{d}_i$, respectively. Estimates of the sufficient statistics, t_1 and $\boldsymbol{t}_2$ are computed as

$$\begin{aligned}
\hat{t_1} &= E\left\{ \sum_{i=1}^{N} \boldsymbol{e}_i^T \boldsymbol{e}_i | \boldsymbol{y}_i, \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}} \right\} \\
&= \sum_{i=1}^{N} \left\{ E\left[\boldsymbol{e}_i^T \boldsymbol{e}_i | \boldsymbol{y}_i, \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}} \right] \right\} \\
&= \sum_{i=1}^{N} \left[\hat{\boldsymbol{e}}_i(\hat{\boldsymbol{\theta}})^T \hat{\boldsymbol{e}}_i(\hat{\boldsymbol{\theta}}) + \mathrm{tr}(var\left\{ \boldsymbol{e}_i | \boldsymbol{y}_i, \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}} \right\}) \right]
\end{aligned} \tag{2.8}$$

and

$$\begin{aligned}
\hat{\boldsymbol{t}_2} &= E\left\{ \sum_{i=1}^{N} \boldsymbol{d}_i \boldsymbol{d}_i^T | \boldsymbol{y}_i, \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}} \right\} \\
&= \sum_{i=1}^{N} \left\{ E\left[\boldsymbol{d}_i \boldsymbol{d}_i^T | \boldsymbol{y}_i, \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}} \right] \right\} \\
&= \sum_{i=1}^{N} \left\{ \hat{\boldsymbol{d}}_i(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{d}}_i(\hat{\boldsymbol{\theta}})^T + var\left(\boldsymbol{d}_i | \boldsymbol{y}_i, \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}} \right) \right\}
\end{aligned} \tag{2.9}$$

where $\hat{\boldsymbol{e}}_i(\hat{\boldsymbol{\theta}}) = E(\boldsymbol{e}_i | \boldsymbol{y}_i, \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}), \hat{\boldsymbol{\theta}}) = \boldsymbol{y}_i - \boldsymbol{X}_i \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}) - \boldsymbol{Z}_i \hat{\boldsymbol{d}}_i(\hat{\boldsymbol{\theta}})$. The maximum likelihood estimates of the parameters are found by starting with a suitable initial value for $\hat{\boldsymbol{\theta}}$ and then iterating between 2.8 and 2.9 (evaluation-steps) and (2.6) and (2.7) (maximizing-steps) until arriving at convergence.

An alternative method for computing the ML estimates is the Newton-Raphson (N-R) algorithm for linear mixed-effects models, which are based on the first- and second-order partial derivatives of the log-likelihood functions. The log-likelihood of the stacked responses, $\boldsymbol{y}_i$, used to derive estimates is denoted as

$$l(\boldsymbol{y};\boldsymbol{\beta},\boldsymbol{\theta}) = -\ln(2\pi)\sum_{i=1}^{N}\frac{p_i}{2} - \frac{1}{2}\sum_{i=1}^{N}\ln|\boldsymbol{V}_i| - \frac{1}{2}\sum_{i=1}^{N}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta})^T\boldsymbol{V}_i^{-1}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta}). \quad (2.10)$$

As detailed by Jennrich and Schluchter (1986), the N-R algorithm iteratively computes new parameter values from current parameter values, using

$$\begin{bmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}^{\circ} \\ \boldsymbol{\theta}^{\circ} \end{bmatrix} - \begin{bmatrix} \boldsymbol{H}_{\beta\beta} & \boldsymbol{H}_{\beta\theta} \\ \boldsymbol{H}_{\theta\beta} & \boldsymbol{H}_{\theta\theta} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{s}_{\beta} \\ \boldsymbol{s}_{\theta} \end{bmatrix} \quad (2.11)$$

with

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{\beta\beta} & \boldsymbol{H}_{\beta\theta} \\ \boldsymbol{H}_{\theta\beta} & \boldsymbol{H}_{\theta\theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 l}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}} & \frac{\partial^2 l}{\partial\boldsymbol{\beta}\partial\boldsymbol{\theta}} \\ \frac{\partial^2 l}{\partial\boldsymbol{\theta}\partial\boldsymbol{\beta}} & \frac{\partial^2 l}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}} \end{bmatrix} \quad (2.12)$$

and

$$\boldsymbol{s} = \begin{bmatrix} \boldsymbol{s}_{\beta} \\ \boldsymbol{s}_{\theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial l}{\partial\boldsymbol{\beta}} \\ \frac{\partial l}{\partial\boldsymbol{\theta}} \end{bmatrix}. \quad (2.13)$$

$\boldsymbol{H}$ is referred to as the Hessian matrix, and $\boldsymbol{s}$ is often described as the gradient or score vector. During the computation algorithm, these values are evaluated using the current values of the parameters.

Both the E-M and N-R estimation methods implicitly impute values for those beyond their time in the study. Therefore, when death is not included in the estimation, as in unconditional models, the population of interest is the original target population, which has been described as an immortal sample (Dufouil et al., 2004). When longitudinal outcomes are due to death, this estimation method may not be the most

appropriate option due to selection bias (Little and Rubin, 2002).

The pattern-mixture model, as denoted previously in equation 1.9, would average over all of the missing patterns, which would imply an implicit extrapolation within each pattern. This type of analysis is not useful when the missing patterns are due only to death. Therefore, Pauler et al. (2003) recommends considering death as a joint outcome instead of a nuisance parameter. Under this advice, the pattern-mixture model in the analysis of the modified NC EPESE is given as follows

$$\boldsymbol{y}_i^r = \boldsymbol{X}_i\boldsymbol{\beta}^r + \boldsymbol{Z}_i\boldsymbol{d}_i^r + \boldsymbol{e}_i^r, \tag{2.14}$$

where r represents the cohorts who died between the first follow-up and the second follow-up ($r = 1$), between the second follow-up and the third follow-up ($r = 2$), and the completers ($r = 3$). The covariance structures for pattern-mixture structures are the same used in the unconditional specifications for each cohort, r. This regression method is attractive because it should give accurate depictions of the trajectories of an outcome for each survival cohort, but requires conditioning on death, which is not known at baseline.

The last method that was fitted using a linear mixed model was the terminal decline model. Terminal decline in this dissertation is confined to the prognostic trend among those subjects that die before the end of the study's observational period. In order to model the terminal decline, the temporal change of interest shifts from the years post-baseline to the years from death. By letting 0 represent the time of death, the years from death are demarcated by their negative magnitudes. Thus, the model specification remains the same as those represented in the unconditional model except that the years post-baseline variable is replaced by the years from death variables. To denote the new design of time, the design matrix in the equation (2.15) is represented by the matrix $\boldsymbol{A}_i$. The matrix $\boldsymbol{U}_i$ is the design matrix for the random intercept and slope of

time from death. The structures of the covariance matrices remain unchanged from the unconditional and pattern-mixture models. Hence, the terminal decline model is modeled as

$$\boldsymbol{y}_i = \boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{U}_i\boldsymbol{d}_i + \boldsymbol{e}_i. \tag{2.15}$$

The primary interest of the principal-stratification method is estimating the unidentifiable survivor average causal effect (SACE)(Frangakis and Rubin, 2002; Hayden et al., 2005; Holland, 1986; Robins and Greenland, 2000; Rubin, 1974, 2000), which is defined as the difference in the outcome for those with the exposure or treatment and the outcome of those without the exposure given participants would survive despite their assigned exposure group. Using the notation introduced in Chapter 1, the SACE takes the following form for a continuous outcome

$$\mu^* = E\left[Y_i(1)|D(z) = 0\right] - E\left[Y_i(0)|D(z) = 0\right]. \tag{2.16}$$

Although others have proposed estimation methods for the SACE for randomized studies (Gilbert et al., 2003; Hayden et al., 2005; Zhang and Rubin, 2003), the SACE in this dissertation uses the estimation method proposed by Egleston et al. (2007). This estimation method was developed to be used in observational studies and estimates the SACE by using a set of unidentifiable assumptions. This estimation was designed to correct the bias that may be a result of both non-response and baseline differences for those with or without the exposure. To describe the estimation method, the notation present in Chapter 1 must be revisited.

First, let $\boldsymbol{X}$ be a vector of covariates, which included the baseline age centered at baseline. The exposure categories were the race-gender groups (white-female, white-male, black-female, and black-male). The exposure indicator is denoted as Z (1 if the race-gender of interest; 0 otherwise). The survival indicator for the given exposure is

denoted by $D(Z)$, such that $D(Z) = 0$ represents survival for exposure status Z. An indicator variable R signifies if a subject reaches the end of the study, ($R = 1$ if not lost to follow; 0 otherwise), and the outcome for exposure is given as $Y(Z)$. For each exposure value, n independent and identically distributed observed data were gathered,

$$\boldsymbol{O} = \{O_i = \{\boldsymbol{X}_i, Z_i, D_i, R_i(\text{ if } D_i = 0), Y_i(\text{ if } D_i = 0 \text{ and } R_i = 1)\}, i = 1, \ldots, n\}.$$

The four principal strata are defined as

1. Individuals who would survive regardless of exposure, $D(0) = D(1) = 0$. (S1)

2. Those who would die if they have the exposure but survive if they do not, $D(0) = 0, D(1) = 1$. (S2)

3. Those who would die regardless of exposure, $D(0) = D(1) = 1$. (S3)

4. Those who would die if they do not have the exposure of interest and would survive if they do, $D(0) = 1, D(1) = 0$. (S4)

The assumptions evoked to identify SACE are given below.

1. Stable Unit Treatment Value Assumption (Rubin, 1980), which states that individual's potential outcomes are not dependent on the exposure status of either the other participants potential outcomes or the mechanism in which the exposure was acquired.

2. Monotonicity is an assumption described by Gilbert et al. (2003) and Zhang and Rubin (2003). This assumption states that acquiring the exposure is not protective to death and implies that principal stratum S4 does not exist. For example, the principal stratum in which an individual is expected to survive if she were a black female but is expected to die if she were not a black female

is not allowed in this estimation. This assumption may be violated for some of the exposure levels for evidence exists of gender-race death associations (Yao and Robert, 2011).

3. Strong ignorability of "treatment" assignment (Rosenbaum and Rubin, 1983) implies that developing an exposure is independent of the potential outcomes given the covariates. In an observational study, this assumption means that the exposure statuses (exposed vs. not exposed) are similar within each covariate level. Thus, the probability of an individual surviving given membership or non-membership in the exposure group given the covariates is denoted as $g_z(\boldsymbol{X}) = P\left[D(Z) = 0|\boldsymbol{X}\right]$. Further $E\left[Y(Z)|D(Z) = 0, \boldsymbol{X}\right] = E\left[Y|D = 0, Z = z, \boldsymbol{X}\right]$.

4. For those who survive, the non-response of the non-mortality outcome is independent of the value of the outcomes within levels of exposure status and covariates. This provides a situation that is similar to the missing at random (MAR) assumption. Coupled with the previous assumption, this assumption makes it possible to identify the expected mean of the outcome for those who would survive but have missing outcomes within the exposure status and covariates. That is, $h_z(\boldsymbol{X}) = E\left[Y(Z)|D(Z) = 0, \boldsymbol{X}\right]$.

The quantity displayed in the results section is the estimate to $E\left[Y(1)|D(Z) = 0, R = 1, \boldsymbol{X}\right]$ at each survey follow-up period. From the above assumptions, we have the following:

$$\begin{aligned} E\left[Y(1)|S1\right] &= E\left[Y(1)|D(1) = 0\right] \\ &= E\left\{E\left[Y(1)|D(1), \boldsymbol{X}\right]\right\} \\ &= \sum E\left[Y(1)|D(1) = 0, \boldsymbol{X}\right] P\left[D(1) = 0|\boldsymbol{X}\right] \\ &= \sum h_1(\boldsymbol{X}) g_1(\boldsymbol{X}). \end{aligned}$$

The mean for each race-gender "exposure" was estimated for each measurement occasion using ordinary least squares regression ($h_1(\boldsymbol{X})$) and the $g_1(\boldsymbol{X})$ was estimated from a logistic regression model. Only one covariate, baseline age centered about the mean, was considered for both models.

The trends produced by the partly conditional model estimate the expected population mean trend on the subject being alive. This method is useful when the interest is the regression of a repeated measure on the participant being alive. That is, $E(Y_{ij}|X_{ij}, S_i > t_j)$. Likelihood methods like the linear mixed model discussed above do not directly parameterize partly conditional models for the estimation method imposes responses for decedents. Rather, Kurland and Heagerty (2005) demonstrate that the generalized estimating equations (Liang and Zeger, 1986) with an independent working correlation directly parameterize the regression model for the target population of those who are alive at the time of collection.

For the partly conditional mean $\mu_{ij}^U = E(Y_{ij}|X_{ij})$, an unbiased, linear quasi-score equation for the regression parameter vector β^U is given by

$$U(\beta^U) = \sum_{i=1}^{N} \sum_{j=1}^{p_i} \frac{\partial \mu_{ij}^U}{\partial \beta} (Y_{ij} - \mu_{ij}^U). \tag{2.17}$$

By letting A_{ij} be an indicator variable that equals 1 if $S_i > t_j$ and 0 otherwise, the quasi-score contributions can be restricted to ensure inference on the target population. The new quasi-score equation becomes

$$U(\beta^A) = \sum_{i=1}^{N} \sum_{j=1}^{p_i} A_{ij} \frac{\partial \mu_{ij}^A}{\partial \beta^A} (Y_{ij} - \mu_{ij}^A). \tag{2.18}$$

This model will yield consistent estimators and valid inference for μ_{ij}^A using a directly

parameterized partly conditional regression model with a link function g.

$$g(\boldsymbol{\mu}_i^{\boldsymbol{A}}) = \boldsymbol{X}_i\boldsymbol{\beta}^A, \tag{2.19}$$

where $\mu_{ij}^A = E(y_{ij}|X_{ij}, S_i > t_j) = E(y_{ij}|X_{ij}, A_{ij} = 1)$ given that the regression model is correctly specified (Crowder, 1986). The outcomes for this model were fitted as a generalized linear regression model with an identity link and an independent working correlation structure.

The probability of being alive and healthy (PAH) (Johnson, 2002) for each participant i was calculated according to the following equation:

$$PAH(s)_i = P(\boldsymbol{Q}_i(s) > q, \boldsymbol{S}_i > s|\boldsymbol{X}_i) = P(\boldsymbol{Q}_i(s) > q|\boldsymbol{S}_i > s)P(\boldsymbol{S}_i > s), \tag{2.20}$$

where $\boldsymbol{Q}(s)_i$ represents the dichotomous health variable and $\boldsymbol{S}_i$ represent the survival time variable. Utilizing repeated logistic regression with an unstructured correlation matrix, the probability of being not depressed ($<$ 9 depressive symptoms), the probability of being non-hypertensive (systolic blood pressure $<$ 140 or diastolic blood pressure $<$ 90), and the probability of not having any activities of daily living limitations (ADL$<$ 1) were calculated for each follow-up period and each gender-race combination. The probability of survival was estimated from a Cox proportional hazards model given by

$$\lambda(s|\boldsymbol{x}) = \lambda_0(s)\exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p) = \exp(\boldsymbol{\beta}'\boldsymbol{x}), \tag{2.21}$$

where $\boldsymbol{x}$ is a p dimensional vector of covariates (centered age at baseline, gender, and race) and $\boldsymbol{\beta}$ is a p-dimensional vector of regression. The covariates are the same for both the logistic and hazard models to make interpretations simpler. The bias in this model is introduced in the estimates of the probabilities of the repeated measure. If missing in the outcome can be assumed to be missing at random (MAR), then the

likelihood models would produce unbiased estimates. In contrast, missingness that is not missing at random (NMAR) could introduce bias in the estimation.

2.4 Results

Figure 2.6(a)-(f), Figure 2.7(a)-(f), Figure 2.8(a)-(e) and Figure 2.9(a)-(f) depict the results of the analysis of the Duke EPESE data using the different proposed methods for assessing the rates of change based on research inquiries for depression scores (CES-D), systolic blood pressure, diastolic blood pressure, and physical function (ADLs), respectively. In all of the models, the baseline variables–age (centered about the mean), race (1 if subjects identify as black and 0 if subjects identify as white), and gender (1 if subjects are identified as male and 0 if subjects are identified as female) – are included as covariates with the exception of the principal stratification model. The principal stratification model uses one's race and gender identification as exposures and age as the only covariate.

In Figure 2.6(a), black females have the highest initial depression value and white males have the lowest baseline depression score. However, each group experiences similar annual rates of change ranging from 0.86-0.90 with women displaying a slightly lower rate. Figure 2.6(b) offers evidence that those that did not die during the study had fewer depressive symptoms than the decedents. Those participants who survived to the first follow-up had the most dissimilar rates of change in CES-D scores. Most notable was the decline in CES-D for black males. Examining the trajectories of the number of depressive symptoms of those individuals that are expected to survive until the end of the study despite their race-gender classification, we found that three years post-baseline scores were lower than the scores at baseline. Further, all groups, except for white males, had estimated CES-D scores at six and ten years post-baseline that follow a similar rate of increase in CES-D (Figure 2.6(c)). Nearing death (Figure

2.6(d)), women had about four depressive symptoms, which is approximately one more than their male counterparts. The terminal decline trends were similar by race. The dynamic cohort's depression rates of change, as modeled by the partly conditional model, were two units lower than those reported in the model assuming an immortal cohort (unconditional model), but the trends were similar. Although men began the study with a higher probability of being alive and having less than nine depressive symptoms (0.92 vs. 0.89; Figure 2.6(f)), over time women became more likely to be healthy and alive.

The unconditional regression model of systolic blood pressure resulted in similar baseline and annual rates of change by gender (Figure 2.7(a)). Women had baseline systolic blood pressures of 144 millimeters of mercury (mmHg), which was higher than their male counterparts. Moreover, the women's rates of decline were only half of the decline rates for the males (0.11 mmHg per year). Women completers in the pattern-mixture graph (Figure 2.7(b)) had the lowest initial systolic blood pressure compared to the cohorts of women that did not complete the study. Yet, their annual rates of change were positive and nearly constant, whereas the rates of change of the other female survival cohorts represented annual rates of decline ranging from 0.55-0.94. The initial systolic blood pressures for the men were the same across the survival cohorts, but their annual rates of decline were smallest among the completers. In Figure 2.7(c), the systolic blood pressure oscillated between a higher mean systolic blood pressure and a slight lower mean, except for white males, who experienced increases in their mean systolic blood pressure after the first follow-up. Mean systolic blood pressure slightly decreased as participants approached their deaths. The terminal rates of decline were similar by race (Figure 2.7(d)). The blood pressure regressed on those being alive (partly conditional) mirrored the rates from the unconditional model. Depicted in Figure 2.7(f), women were less than 50% likely to be alive and non-hypertensive while

men were approximately 50% alive and non-hypertensive. As the groups aged, the men's probability of being alive and non-hypertensive declined more rapidly than the women's rates.

Rates of change in the mean diastolic blood pressure for the immortal cohort were similar across race-gender groups. Additionally, black women and men had similar baseline values for diastolic blood pressure measurements, which remained true for white men and women (Figure 2.8(a)). In Figure 2.8(b), the baseline values of diastolic blood pressure are very similar across the survival cohorts for each race-gender group. Nonetheless, white men and women who died after the first follow-up only experienced small declines in their diastolic blood pressure measurements annually. The trend of diastolic blood pressure for those who would survive regardless of race or gender had similar baseline values to the unconditional and pattern-mixture completers' regression models, except for white males. Although the baseline values are similar, the rates of change for the principal stratification were much higher than the other two methods. Men's mean diastolic blood pressure declined more rapidly than the womens as they neared death (Figure 2.8(d)). Figure 2.8(e) displays the estimated means of diastolic blood pressure for those alive at the given follow-up occasion. These values are similar to the unconditional model results.

Unconditional regressed ADLs graph in Figure 2.9(a) display similar baseline values by race and similar rates of increase by gender. Figure 2.9(b) offers evidence that those that did not die during the study had fewer physical functioning limitations than the study decedents. The rates of change were similar by race-gender groups for $r = 1$ and $r = 2$ survival cohorts. For those who would survive regardless of race-gender assignment, their rates of change were similar by gender. The men seem to have a leveling of physical functioning dependency, while women continued to experience increases in their mean physical function limitations (Figure 2.9(c)). As

women approach death, their physical function dependency increased more rapidly than men (0.30 versus 0.25). The physical functioning estimated means of the mutable population had comparable means to the mean estimates of the unconditional model for each race-gender group (Figure 2.9(e)). Black women were the most physically limited at baseline with a PAH of 0.63. The others had probabilities greater than 0.70. Even though the annual rate of decline was slightly higher for males than females (0.07 annual rate of decline), the graphical trends were alike.

2.5 Simulation of Varying Death Burden

To evaluate the ability of the unconditional, pattern-mixture, and partly conditional models to estimate rates of change in the Center for Epidemiologic Studies Depression (CES-D) scores, systolic and diastolic blood pressures, and activities of daily living (ADL) of a complete dataset without bias for various MCAR burdens of death, we simulated a sample from a theoretical population. Each of the four outcomes were treated as continuous outcomes and were generated from the mixed model with a random intercept and slope as described in equation (2.1). Four waves of longitudinal outcomes were simulated from a normal distribution with mean $\boldsymbol{X_i\beta}$ and covariance $\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i\boldsymbol{D}_i\boldsymbol{Z}_i^T + \sigma^2\boldsymbol{I}_i$, where $\boldsymbol{D}$ was allowed to be unstructured. For each outcome, the design matrix $\boldsymbol{X}_i$ consisted of a column vector of ones for the intercept, a column vector for time of measurements post-baseline (0, 3, 6, 10 years), a column vector of baseline age centered about the mean, a column vector of indicators for identifying as black, a column vector of indicators for identifying as male, a column vector indicating time by race, and a column vector indicating time by sex. The design matrix of the random effects, ($\boldsymbol{Z}_i$), was constructed as a column vector of ones for the intercept and a column vector for the time of measurements post-baseline (0, 3, 6, 10 years). The race-gender combination values were treated as multinomial random variables and

were generated accordingly: white males ($\pi = 0.16$), black males ($\pi = 0.19$), and white females ($\pi = 0.29$). Age was simulated assuming it was from a normal distribution dictated by the mean and standard deviation in each race-gender group from the NC EPESE. Similarly, time of measurements mirrored the NC EPESE. Thus, we assume measurements were only possible at baseline, and 3, 6, and 10 years post-baseline. The values of the parameters $\boldsymbol{\beta}, \sigma^2$, and $\boldsymbol{D}$ used in the simulation are given for each outcome below.

CES-D score: $\boldsymbol{\beta}^T = \begin{pmatrix} 3.243 & 0.086 & 0.065 & 0.356 & -0.686 & 0.001 & 0.003 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 5.32 & -0.14 \\ -0.14 & 0.39 \end{pmatrix} \boldsymbol{Z}_i^T + 7.16\boldsymbol{I}_i$$

Systolic BP: $\boldsymbol{\beta}^T = \begin{pmatrix} 143.29 & 0.088 & 0.002 & 0.756 & -2.326 & -0.017 & -0.136 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 139.69 & -3.38 \\ -3.38 & 0.91 \end{pmatrix} \boldsymbol{Z}_i^T + 314.05\boldsymbol{I}_i$$

Diastolic BP: $\boldsymbol{\beta}^T = \begin{pmatrix} 77.258 & 0.799 & -0.289 & 2.775 & 1.529 & -0.0006 & -0.010 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 47.66 & -2.26 \\ -2.26 & 0.39 \end{pmatrix} \boldsymbol{Z}_i^T + 92.94 \boldsymbol{I}_i$$

ADL score: $\boldsymbol{\beta}^T = \begin{pmatrix} 0.742 & 0.211 & 0.114 & 0.243 & 0.133 & 0.001 & -0.063 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z} \begin{pmatrix} 1.52 & -0.002 \\ -0.002 & 0.04 \end{pmatrix} \boldsymbol{Z}^T + 1.78 \boldsymbol{I}_i$$

A 1,000 samples of complete data were generated for each sample size– $N = 100$, $N = 500$, and $N = 1000$. After the complete datasets were generated, death indicators were created from a Bernoulli random generating function to simulate participants leaving the study because of death. Subjects in the simulated datasets became at risk of death following baseline responses. One death indicator simulated a 10% death rate per survey wave following baseline for an overall death rate of approximately 27% of baseline participants. Another missing scheme allowed the death rate to increase as the survey years increased by simulating a death indicator with a death rate of 10% after baseline, 20% after the first follow-up, and 30% after the second follow-up, resulting in an expected 50% of baseline participants dying before the end of the study. The final indicator simulated a death rate of 30% per post-baseline survey wave resulting in an expected overall death rate of 66%.

2.6 Simulations Results

The choice is made to fit the unconditional, pattern-mixture, and partly conditional models as described earlier for each of the 1,000 samples and for the three death percentage scenarios. All analyses were performed using SAS v9.2. For the unconditional and pattern-mixture models, maximum likelihood estimation was provoked. The partly conditional models were estimated by generalized estimating equations using an identity working correlation matrix and empirical standard errors to account for the repeated continuous measures per subject. The relative bias of the estimation for each mean per race-gender combination was computed. Tables 2.1 - 2.4 give the mean bias for each method by the different death percentage scenario. Pattern-mixture models could only be performed for the 500 and 1,000 sample sizes. The sample size of $N = 100$ did not provide enough participants for the survival cohorts for some of the missing schemes.

Tables 2.1-2.4 give the mean relative bias of the estimates for each race-gender combination five years post-baseline for the unconditional, pattern-mixture completers, and partly conditional models. The unconditional model was able to estimate the mean systolic and mean diastolic blood pressures with minimal bias relative to the true means for each sample size and death percentage burdens. When estimating the mean CES-D sores in the sample size of 500 and 1,000, the unconditional model seemed to perform better when the overall percentage of death is small. The model that uses only the completers had larger relative bias than the other models for all outcomes. Mean depression scores were able to be estimated using the partly conditional model with a relative biases that were smaller than the pattern-mixture model. The bias increased as the overall percentage of death increased. None of the models were stable at estimating the mean number of activities of daily living (ADL). Black males experienced alarmingly high relative biases than the other race-gender groups for ADL. One explanation is

because this groups estimation depends on the estimation of all parameters.

2.7 Discussion

By considering survival in the estimation of depression, systolic and diastolic blood pressures, and functional dependency over time, the analyses presented in this paper contribute to the previous understanding of the nature of these outcome through a new level of reliability of the estimates as well as depictions of the outcomes trajectories. Because the NC EPESE was part of an inaugural study on older adults living in America and the only dataset that allowed for adequate race comparison, the dataset has been studied intensely, in particular, for the outcome of depression. Blazer et al. (1991) examined the association of age and depression using a cross-sectional regression analysis. Their analysis concluded that age and depression had an indirect relationship when adjusted for gender, income, physical disability, cognitive impairment, and social support. Although our analysis did not account for some of the key covariates associated with depression, each of the six models supported a direct association of aging and depression, except for the trend for those black males who died before the second follow-up. Furthermore, the graphs presented in 2.6(a)-(f) offer one of the few longitudinal trends of depression on the North Carolina Established Populations of Epidemiological Studies of the Elderly (NC EPESE). Thus, these results contributes to what we know about depression over time for older adults.

Previous studies of blood pressure of the NC EPESE have asserted many conclusions of hypertension in relation to race (Howard et al., 2009; Blazer et al., 2001; Gold et al., 1996; Svetkey et al., 1993). For example, blacks in the NC EPESE have been previously shown to have higher prevalence rates of hypertension than whites. Moreover, unlike systolic blood pressure, diastolic blood pressure for older adults has been reported to only be associated with mortality for whites (Blazer et al., 2001) Figure 2.7(f) supports

that when survival status and hypertension are modeled together, a gender association is more prominent than a race association. The joint model supports similar results by gender in terms of both baseline values and rates of change in the probability of being alive and healthy over time. However, race associations with blood pressure are visible in the analysis of diastolic blood pressure.

Currently, mobile disability is defined as difficulty or dependency in carrying out activities essential to independent living and desired activities important to ones quality of life; it is typically screened through Activities of Daily Living (ADL) and Instrumental Activities of Daily Living (IADL) citeptopinkova. Although this study utilized only the ADL score to define limitations of physical functioning, the previously reported racial gap in disability remained supported citeptaylor.

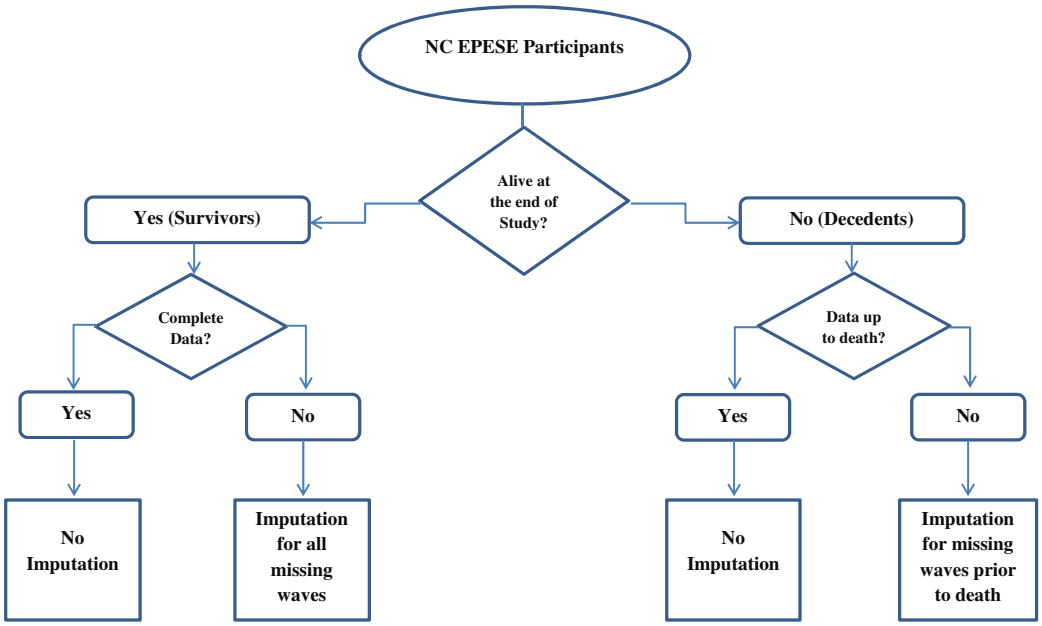
The major contribution of this study was the use of advanced statistical methods that included survival status in the estimation of the longitudinal means of outcomes from a subpopulation (NC EPESE) of the popular longitudinal study of older adults. Additionally, the appropriate populations and aims of these models were presented. These current findings, along with previous results of these outcomes, strengthen the understanding of accurate changes of the outcome measures as a cohort of older adults become older.

From the simulations with death indications that did not depend on the covariates or the outcome (MCAR), we noticed that the estimated means for the immortal cohorts, completers, and dynamic cohort of survivors five years post-baseline were reliable at estimating the systolic and diastolic blood pressures, despite the burden of death. However, there is evidence that the pattern-mixture carries a slight increase of bias in its estimation of the means of the race-gender groups. The mean depression scores were estimated by the different models with very little bias, generally, but the models performed better when the sample size increased and the percentage of those who died

was lower. The relative biases of the estimates of the mean of activities of daily living scores were notably larger than the other outcomes, especially for black males. The mean estimates for black males were dependent on all of the parameter estimates. The reason for this occurrence is not quite clear, but we suspect that there may exist a vulnerability in our simulation used to generate the data. With the exception of the results from ADL, the models seem to perform quite well when the missing assumption is MCAR.

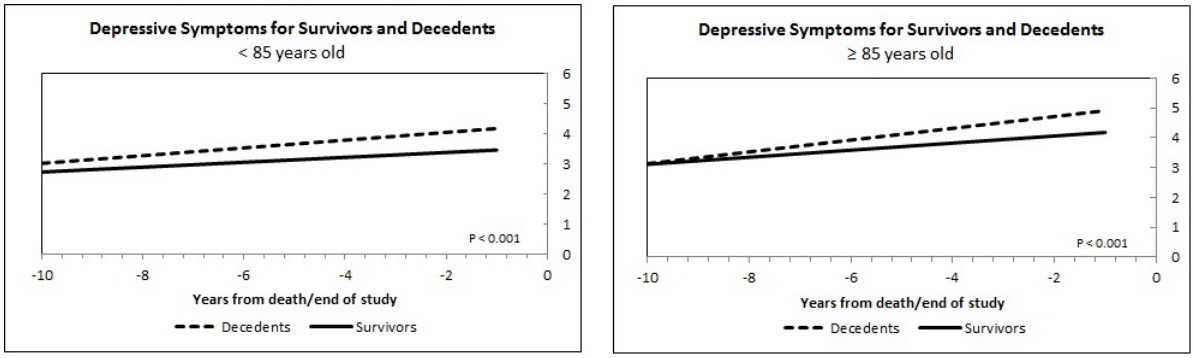
In conclusion, the target of this study was to revisit the analysis of outcomes from the NC EPESE and re-analyze the outcomes with models that integrate survival status. The results presented in this study extend our knowledge of these outcomes longitudinally and provide descriptions of survival-incorporating methods. Further, the simulations offer insight that the different means produced by the survival incorporating methods are similar when compared to the initial population of interest.

Figure 2.1: Data imputation decision chart.



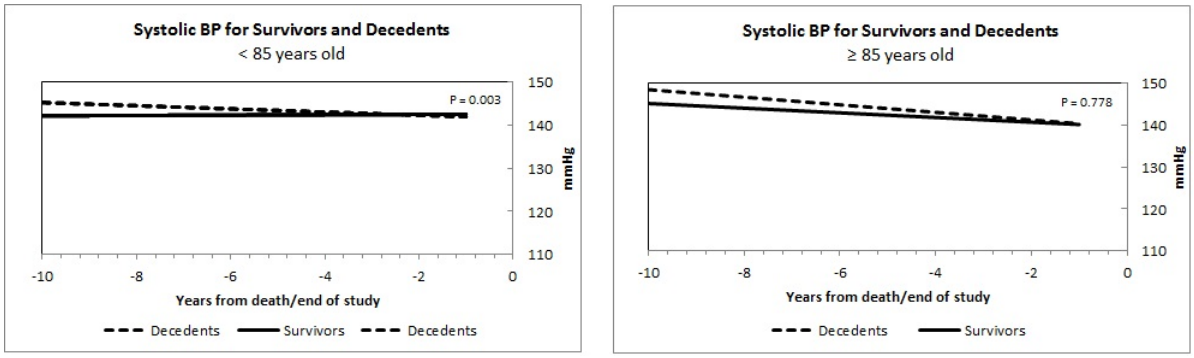
Describes the subjects whose outcome variables are subject to imputation in the North Carolina EPESE.

Figure 2.2: Mean Depression Score for those > 85 and those ≥ 85 years old.



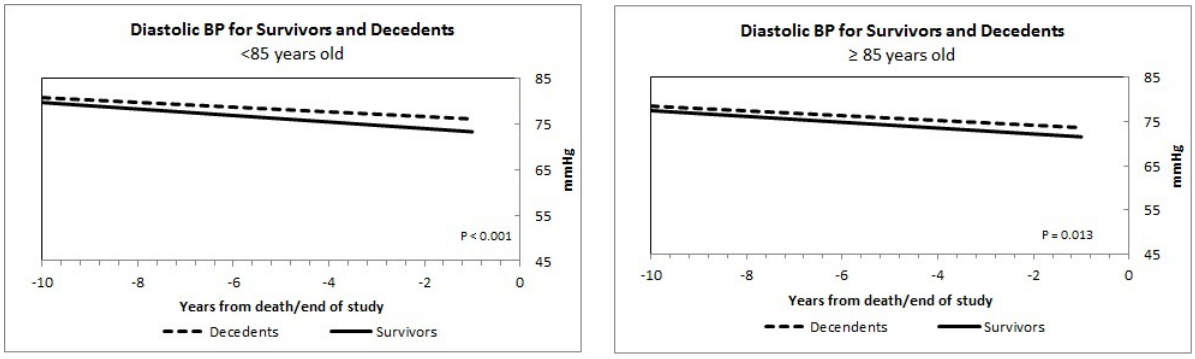
The p value in each panel corresponds to the test of the difference in the rate of change due to death.

Figure 2.3: Mean Systolic Blood Pressure for those > 85 and those ≥ 85 years old.



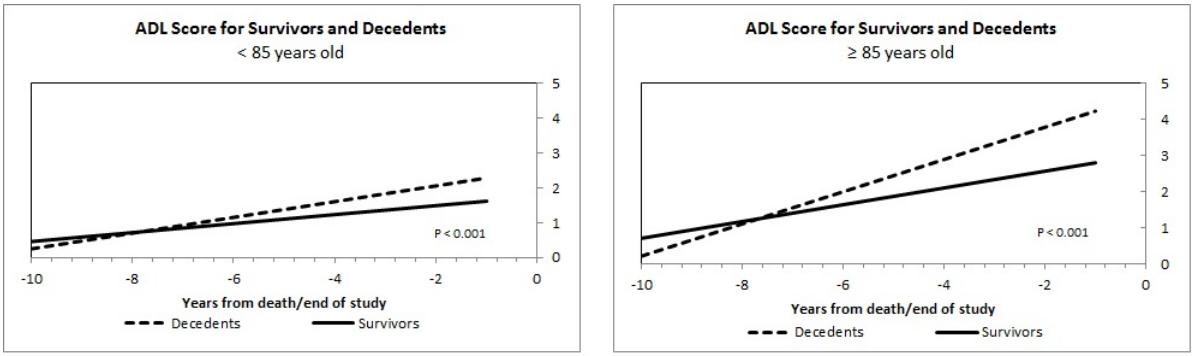
The p value in each panel corresponds to the test of the difference in the rate of change due to death.

Figure 2.4: Mean Diastolic Blood Pressure for those > 85 and those ≥ 85 years old.



The p value in each panel corresponds to the test of the difference in the rate of change due to death.

Figure 2.5: Mean Functional Score for those > 85 and those ≥ 85 years old.



The p value in each panel corresponds to the test of the difference in the rate of change due to death.

Figure 2.6: Fitted trajectories of CES-D scores for EPESE participants

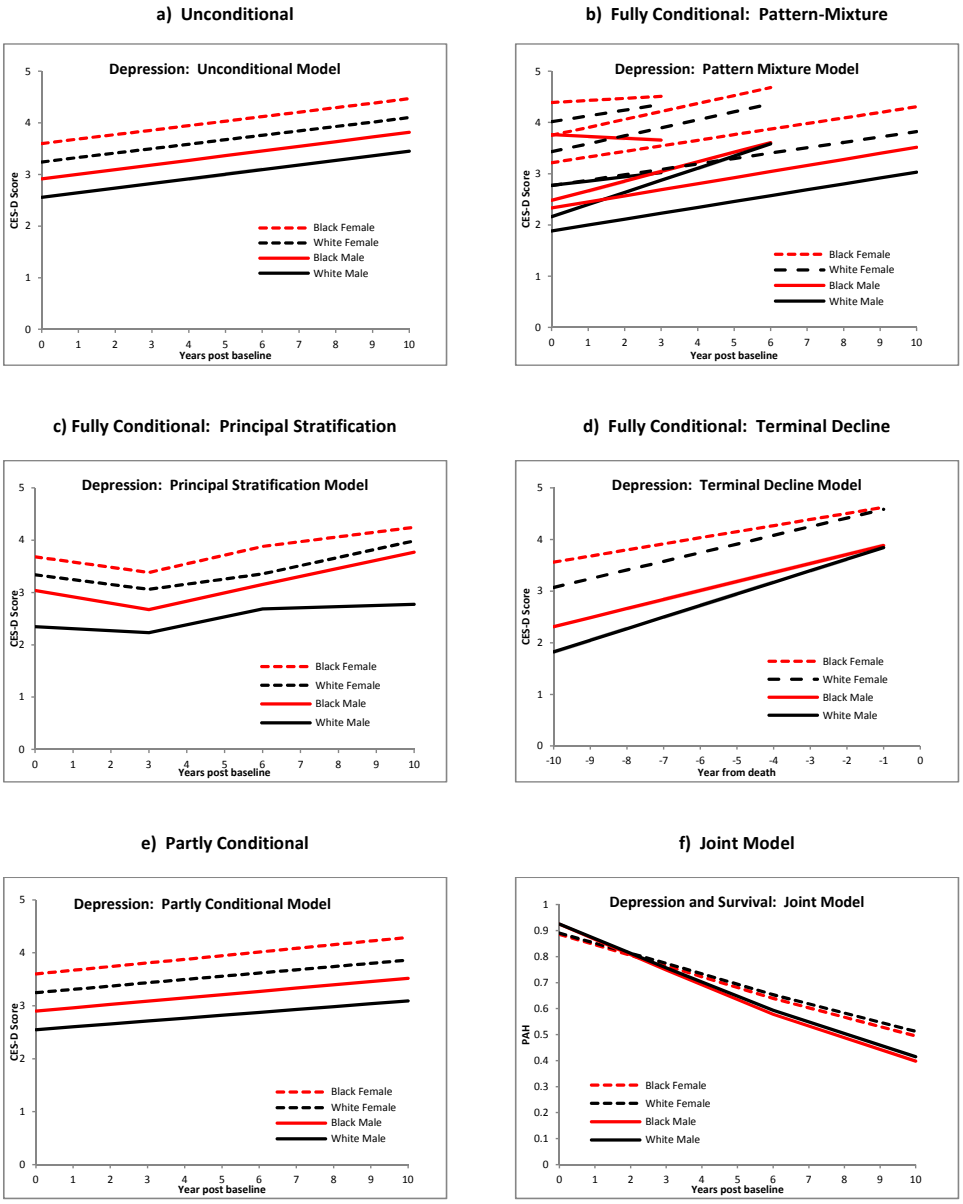


Figure 2.7: Fitted trajectories of systolic blood pressure for EPESE participants

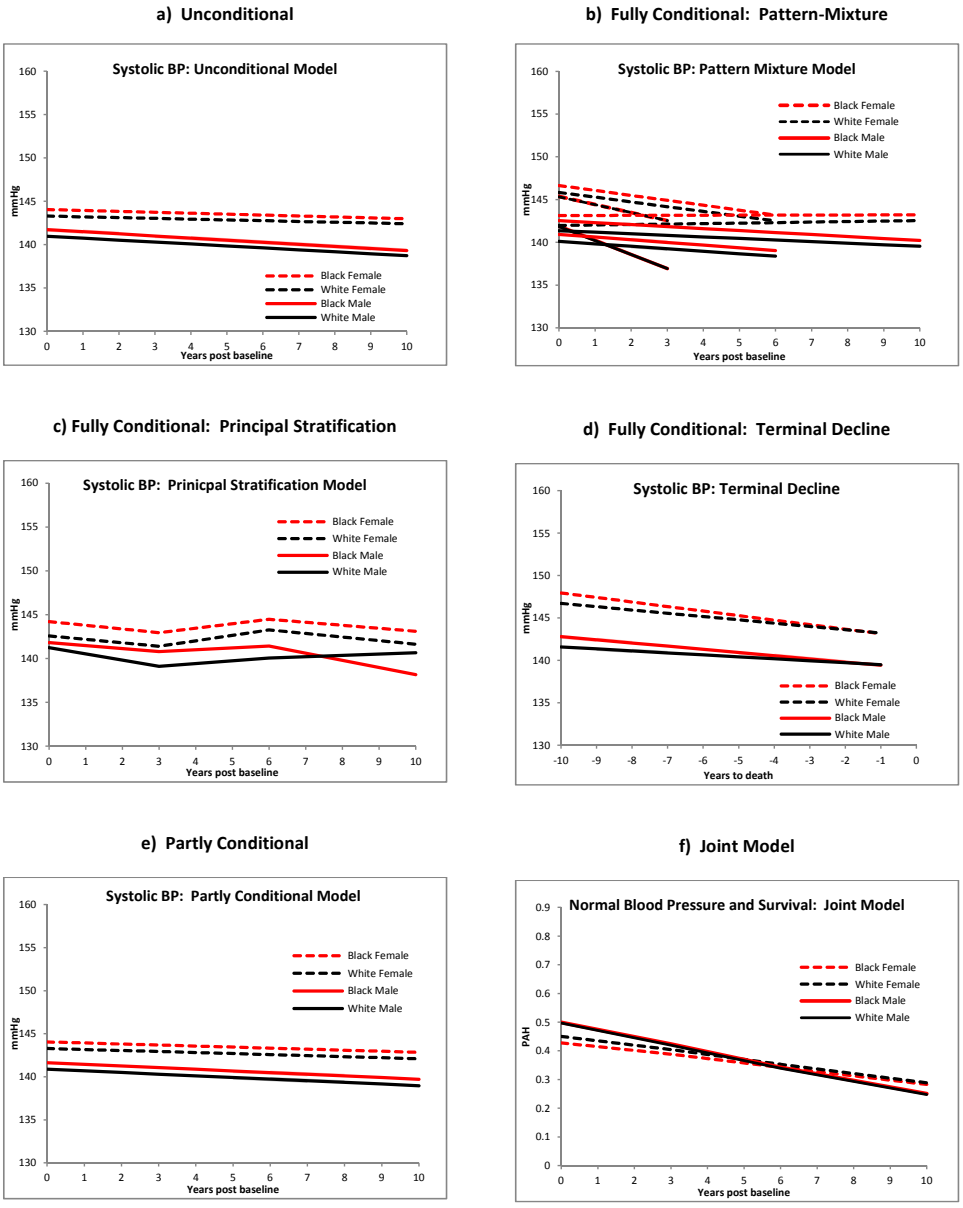


Figure 2.8: Fitted trajectories of diastolic blood pressure for EPESE participants

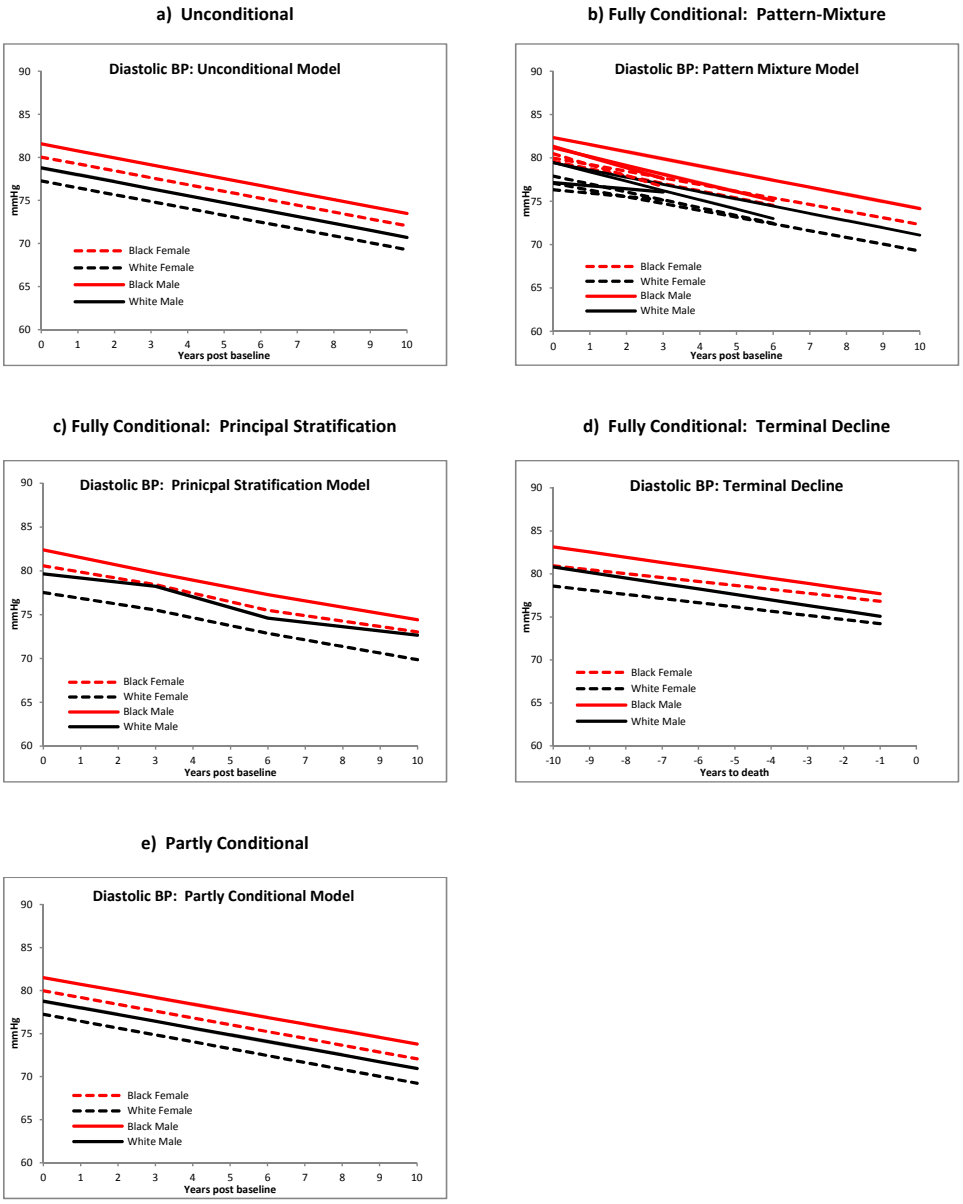


Figure 2.9: Fitted trajectories of ADL Scores for EPESE participants

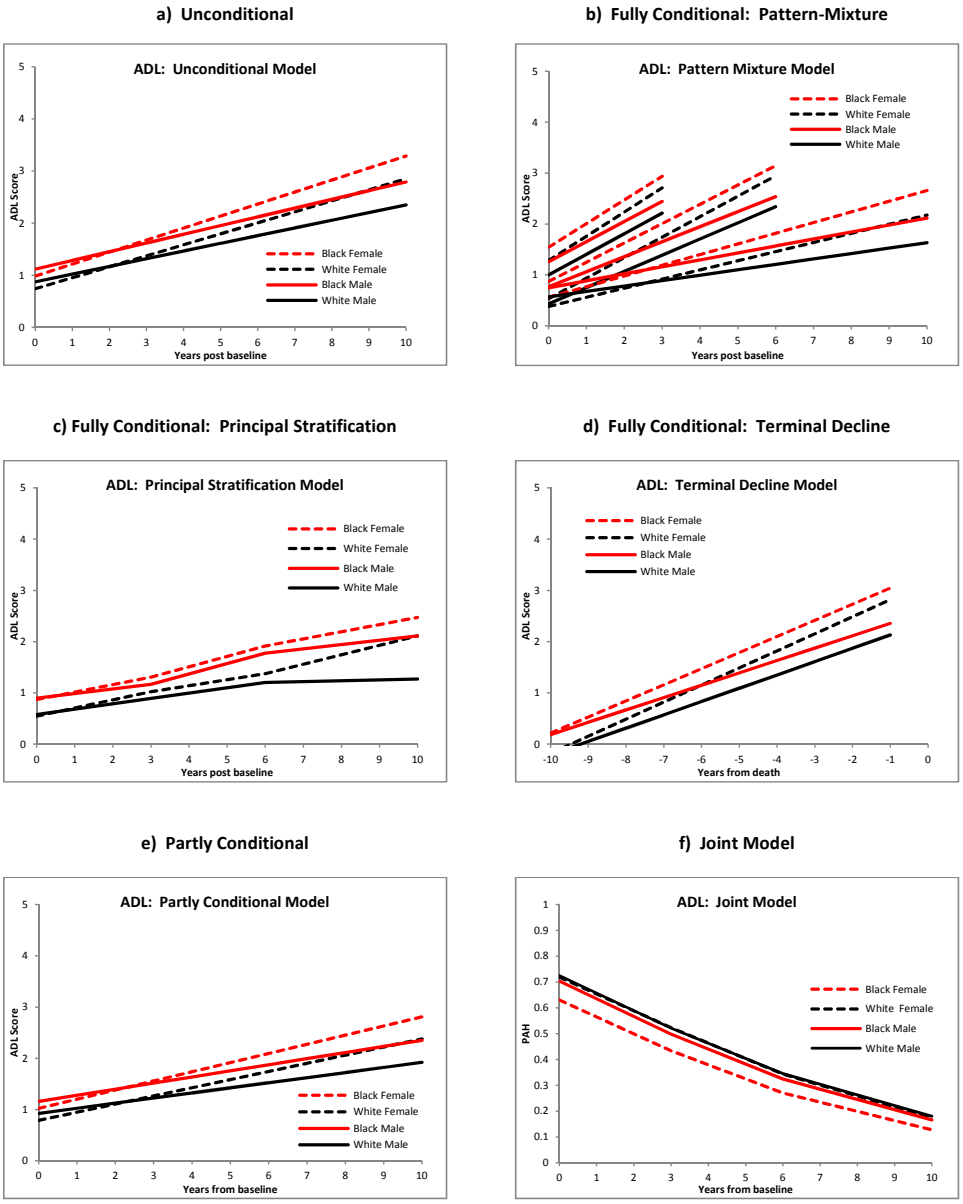


Table 2.1: CES-D - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1000 simulated samples with three follow-up times at varying percentages of death per wave

		Sample size= 100			**Sample size**= 500			**Sample size**= 1000		
Model	**Mean**	10%	10-20-30%	30%	10%	10-20-30%	30%	10%	10-20-30%	30%
Unconditional	White Females	-0.60	-1.21	0.16	0.20	0.10	0.39	-0.10	0.00	-0.33
		(0.0063)	(0.0065)	(0.0074)	(0.0026)	(0.0028)	(0.0031)	(0.0020)	(0.0020)	(0.0024)
	White Males	-1.19	-1.49	-0.51	0.25	0.28	0.30	0.13	0.39	0.46
		(0.0089)	(0.0093)	(0.0107)	(0.0040)	(0.0042)	(0.0049)	(0.0027)	(0.0029)	(0.0034)
	Black Females	0.45	-0.17	0.51	0.23	0.42	0.51	-0.28	-0.18	-0.33
		(0.0052)	(0.0053)	(0.0062)	(0.0022)	(0.0023)	(0.0027)	(0.0156)	(0.0017)	(0.0019)
	Black Males	0.23	-0.12	0.06	0.37	0.23	0.54	-0.02	0.22	0.45
		(0.0077)	(0.0080)	(0.0090)	(0.0035)	(0.0035)	(0.0041)	(0.0024)	(0.0025)	(0.0029)
Pattern	White Females				0.26	0.48	0.27	-0.12	-0.42	-0.54
Mixture					(0.0028)	(0.0036)	(0.0043)	(0.0021)	(0.0025)	(0.0032)
Completers	White Males				0.67	0.58	1.08	0.06	0.79	0.50
					(0.0044)	(0.0054)	(0.0064)	(0.0029)	(0.0037)	(0.0045)
	Black Females				0.01	0.19	0.44	-0.28	-0.54	-0.26
					(0.0024)	(0.0031)	(0.0036)	(0.0017)	(0.0021)	(0.0024)
	Black Males				0.42	0.32	1.28	-0.07	0.61	0.82
					(0.0038)	(0.0046)	(0.0054)	(0.0025)	(0.0032)	(0.0040)
Partly	White Females	-0.70	-1.49	0.08	0.23	0.26	0.38	-0.13	-0.22	-0.44
Conditional		(0.0064)	(0.0069)	(0.0080)	(0.0027)	(0.0030)	(0.0034)	(0.0020)	(0.0021)	(0.0026)
	White Males	-1.02	-1.14	-0.63	0.45	0.48	0.58	0.12	0.54	0.55
		(0.0093)	(0.0099)	(0.0118)	(0.0042)	(0.0045)	(0.0052)	(0.0027)	(0.0030)	(0.0036)
	Black Females	0.26	-0.38	0.52	0.11	0.22	0.46	-0.30	-0.35	-0.35
		(0.0054)	(0.0057)	(0.0068)	(0.0022)	(0.0025)	(0.0029)	(0.0016)	(0.0019)	(0.0020)
	Black Males	0.26	0.24	0.06	0.38	0.49	0.75	-0.03	0.39	0.64
		(0.0079)	(0.0084)	(0.0099)	(0.0036)	(0.0038)	(0.0044)	(0.0024)	(0.0027)	(0.0032)

Table 2.2: Systolic BP - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1,000 simulated samples with three follow-up times at varying percentages of death per wave

		Sample size= 100			**Sample size**= 500			**Sample size**= 1000		
Model	**Mean**	10%	10-20-30%	30%	10%	10-20-30%	30%	10%	10-20-30%	30%
Unconditional	White Females	-0.07	-0.14	-0.01	0.02	0.01	0.04	-0.02	-0.01	-0.04
		(0.0006)	(0.0006)	(0.0007)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	White Males	-0.07	-0.11	-0.04	0.01	0.01	0.01	0.01	0.03	0.03
		(0.0007)	(0.0007)	(0.0008)	(0.0003)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0003)
	Black Females	0.05	-0.01	0.06	0.02	0.02	0.04	-0.03	-0.02	-0.03
		(0.0005)	(0.0005)	(0.0006)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	Black Males	0.05	0.02	0.02	0.01	0.02	0.01	0	0.02	0.04
		(0.0007)	(0.0007)	(0.0008)	(0.0003)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0003)
Pattern	White Females				-0.03	0.04	0.01	-0.02	-0.05	-0.05
Mixture					(0.0003)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0003)
Completers	White Males				0.05	0.04	0.06	0.01	0.06	0.05
					(0.0003)	(0.0004)	(0.0005)	(0.0002)	(0.0003)	(0.0003)
	Black Females				-0.00	0.01	0.05	-0.03	-0.05	-0.03
					(0.0002)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0002)
	Black Males				0.02	0.00	0.11	-0.00	0.05	0.07
					(0.0003)	(0.0004)	0.0005)	(0.0002)	(0.0003)	(0.0003)
Partly	White Females	-0.07	-0.15	-0.00	0.02	0.02	0.03	-0.02	-0.02	-0.04
Conditional		(0.0006)	(0.0006)	(0.0007)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	White Males	-0.07	-0.09	-0.04	0.02	0.02	0.02	0.01	0.03	0.04
		(0.0007)	(0.0008)	(0.0009)	(0.0003)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0003)
	Black Females	0.05	-0.22	0.07	0.01	0.02	0.05	-0.03	-0.02	-0.03
		(0.0005)	(0.0006)	(0.0006)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	Black Males	0.05	0.04	0.03	0.01	0.02	0.04	-0.01	0.03	0.47
		(0.0007)	(0.0007)	(0.0008)	(0.0003)	(0.0002)	(0.0004)	(0.0002)	(0.0002)	(0.0003)

Table 2.3: Diastolic BP - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1,000 simulated samples with three follow-up times at varying percentages of death per wave

		Sample size= 100			**Sample size**= 500			**Sample size**= 1000		
Model	**Parameters**	10%	10-20-30%	30%	10%	10-20-30%	30%	10%	10-20-30%	30%
Unconditional	White Females	-0.07	-0.13	-0.01	0.02	0.01	0.03	-0.02	-0.01	-0.04
		(0.0005)	(0.0006)	(0.0007)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	White Males	-0.07	-0.10	-0.05	0.00	0.00	0.00	0.00	0.02	0.03
		(0.0006)	(0.0007)	(0.0008)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	Black Females	0.04	-0.01	0.06	0.02	0.02	0.04	-0.03	-0.02	-0.03
		(0.0005)	(0.0005)	(0.0006)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0002)
	Black Males	0.04	0.02	0.02	0	0.02	0.01	0.00	0.02	0.03
		(0.0006)	(0.0006)	(0.0007)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
Pattern	White Females				0.02	0.04	0.01	-0.02	-0.04	-0.04
Mixture					(0.0002)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0003)
Completers	White Males				0.03	0.03	0.05	0.00	0.05	0.04
					(0.0003)	(0.0004)	(0.0004)	(0.0002)	(0.0003)	(0.0003)
	Black Females				-0.00	0.01	0.05	-0.02	-0.04	-0.02
					(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	Black Males				0.01	-0.00	0.09	-0.01	0.04	-0.06
					(0.0003)	(0.0003)	(0.0004)	(0.0002)	(0.0002)	(0.0003)
Partly	White Females	-0.07	-0.14	-0.04	0.02	0.01	0.02	-0.02	-0.02	-0.04
Conditional		(0.0005)	(0.0006)	(0.0007)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	White Males	-0.07	-0.09	-0.04	0.01	0.01	0.01	0.00	0.02	0.03
		(0.0006)	(0.0007)	(0.0008)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	Black Females	0.04	-0.02	0.06	0.01	0.02	0.04	-0.03	-0.02	-0.03
		(0.0005)	(0.0005)	(0.0006)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0002)
	Black Males	0.04	0.03	0.02	0.00	0.01	0.03	-0.00	0.02	0.04
		(0.0006)	(0.0006)	(0.0007)	(0.0003)	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)

Table 2.4: ADL - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1,000 simulated samples with three follow-up times at varying percentages of death per wave

		Sample size= 100			**Sample size**= 500			**Sample size**= 1000		
Model	**Parameters**	10%	10-20-30%	30%	10%	10-20-30%	30%	10%	10-20-30%	30%
Unconditional	White Females	-1.79	-2.34	-1.18	-0.90	-1.01	-0.77	-1.27	-1.20	-1.46
		(0.0055)	(0.0057)	(0.0064)	(0.0023)	(0.0024)	(0.0027)	(0.0017)	(0.0018)	(0.0020)
	White Males	0.94	0.69	1.36	2.00	2.02	2.01	1.98	2.18	2.23
		(0.0072)	(0.0074)	(0.0084)	(0.0032)	(0.0033)	(0.0038)	(0.0022)	(0.0023)	(0.0027)
	Black Females	1.02	0.53	1.1	0.81	0.79	1.02	0.39	0.47	0.35
		(0.0044)	(0.0448)	(0.0051)	(0.0019)	(0.0020)	(0.0022)	(0.0013)	(0.0014)	(0.0015)
	Black Males	3.66	3.44	3.52	3.49	3.59	3.61	3.37	3.56	3.73
		(0.0059)	(0.0061)	(0.0068)	(0.0027)	(0.0028)	(0.0031)	(0.0018)	(0.0019)	(0.0022)
Pattern	White Females				-0.86	-0.65	-1.01	-1.25	-1.52	-1.52
Mixture					(0.0025)	(0.0033)	(0.0038	(0.0019)	(0.0022)	(0.0028)
Completers	White Males				2.35	2.29	2.61	1.93	2.49	2.32
					(0.0036)	(0.0044)	(0.0052)	(0.0024)	(0.0030)	(0.0036)
	Black Females				0.64	0.73	1.05	0.41	0.22	0.44
					(0.0021)	(0.0027)	(0.0031)	(0.0015)	(0.0018)	(0.0021)
	Black Males				3.57	3.41	4.39	3.34	3.86	3.97
					(0.0029)	(0.0036)	(0.0043)	(0.0020)	(0.0025)	(0.0031)
Partly	White Females	-1.88	-2.52	-1.20	-0.89	-0.87	-0.85	-1.26	-1.35	-1.45
Conditional		(0.0056)	(0.0060)	(0.0068)	(0.0023)	(0.0026)	(0.0029)	(0.0018)	(0.0018)	(0.0022)
	White Males	1.02	0.96	1.30	2.15	2.18	2.20	1.97	2.26	2.33
		(0.0074)	(0.0078)	(0.0091)	(0.0033)	(0.0035)	(0.0041)	(0.0022)	(0.0024)	(0.0028)
	Black Females	0.93	0.38	1.57	0.73	0.78	1.06	0.39	0.37	0.36
		(0.0045)	(0.0047)	(0.0055)	(0.0019)	(0.0021)	(0.0024)	(0.0014)	(0.0015)	(0.0016)
	Black Males	3.73	3.69	3.56	3.53	3.58	3.89	3.36	3.67	3.82
		(0.0060)	(0.0064)	(0.0073)	(0.0027)	(0.0029)	(0.0033)	(0.0019)	(0.0020)	(0.0024)

CHAPTER 3: ASSESSMENT OF MODELS FOR ANALYZING LONGITUDINAL OUTCOMES TRUNCATED DUE TO DEATH AND NON-PARTICIPATION

3.1 Introduction

Very few if any observational studies that collect repeated measures on a sample over several waves are able to achieve complete data collection. This is especially true for large observational studies of older adults. Rhodes (2005), who performed a meta-analysis on the characteristics of attrition as reported in 57 studies that analyzed participants 50 years and older in 13 prestigious gerontological journals over a span of 30 years, reported an average overall attrition rate of 34%. Several reasons contributed to this drop-out, but death, illness, and lack of interest were typically the main culprits. Over the past several decades, the characteristics of those participants in the different categories have been established. For instance, those who are lost to follow-up in earlier waves are typically lost due to lack of interest; however, the individuals who are lost at later follow-up occasions are usually lost because of illness or death (Norris, 1985; Schaie, 1996). Furthermore, those who are not retained in the study due to illness or death have been shown to perform lower biologically, cognitively, and functionally than those that leave for other reasons (Rhodes, 2005; Rabbit et al., 1994).

Generally, older adults, who are lost to attrition are outperformed by those who remain. In this situation, internal and external validity are compromised from the overrepresentation of healthier participants and underrepresentation of the effect on the response as individuals get older. To protect from bias, it is imperative to account

for attrition. Moreover, treating different reasons for loss-to-follow-up as the same may not be valid or the best practice.

Several methods have been proposed to account for non-response and death. As discussed previously, parametric likelihood models provide valid inference about the parameters, say $\boldsymbol{\beta}$, given the missing data is missing completely at random (MCAR)(Rubin, 1976). Yet, the estimation methods for incomplete data may induce selection bias when they implicitly impute the missing data by conditioning on the observed data. To accommodate left and right censoring in the linear mixed model, Hughes (1999) modified the E-M algorithm originally posed by Dempster et al. (1981a) and utilized for the random effect models by Laird and Ware (1982). His procedure was based on an example of the Monte Carlo E-M Algorithm (MCEM) introduced by Wei and Tanner (1990). He observed (Q_{ij}, C_{ij}) for subject i at time j, where Q_{ij} is the response that could be censored and C_{ij} is the censoring indicator. When the response is not censored, $C_{ij} = 0$, then $Q_{ij} = Y_{ij}$. Reaching the floor or the ceiling is represented by $C_{ij} = -1$ (implies $Y_{ij} < Q_{ij}$) and $C_{ij} = 1$ (implies $Y_{ij} > Q_{ij}$), respectively. Just as in Chapter 2, maximum likelihood estimates can be obtained from this estimation model by employing the E-M algorithm. Including the new observed variables, $\boldsymbol{Q}$ and $\boldsymbol{C}$. The M-step of the algorithm is written as

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= \left(\boldsymbol{X}^T\hat{\boldsymbol{W}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\hat{\boldsymbol{W}}E(\boldsymbol{y}|\boldsymbol{C},\boldsymbol{Q},\hat{\boldsymbol{\theta}}) \\
\hat{\boldsymbol{D}} &= \sum_{i=1}^{N} E\left(\boldsymbol{d}_i\boldsymbol{d}_i^T|\boldsymbol{C}_i,\boldsymbol{Q}_i,\hat{\boldsymbol{\theta}}\right)/N \\
\sigma^2 &= \sum_{i=1}^{N} E\left(\boldsymbol{e}_i^T\boldsymbol{e}_i|\boldsymbol{C}_i,\boldsymbol{Q}_i,\hat{\boldsymbol{\theta}}\right)/\sum_{i=1}^{N} p_i
\end{aligned} \tag{3.1}$$

where $\hat{\boldsymbol{\theta}}$ is a vector of values for the model parameters. The E-step involves solving the expectations in equation (3.1). By letting $\int_{y_i(C,Q)}$ denote the integral over all of the $\boldsymbol{y}_i$ that are consisted with the observed $\boldsymbol{C}_i$ and $\boldsymbol{Q}_i$, and letting f denote a generic density

function, the conditional densities of the covariance components can be written as

$$\begin{aligned} f(\boldsymbol{d}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) &= \int_{y_i(C,Q)} f(\boldsymbol{d}_i|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}) f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) dy_{i(C,Q)} \\ f(\boldsymbol{e}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) &= \int_{y_i(C,Q)} f(\boldsymbol{e}_i|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}) f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) dy_{i(C,Q)}. \end{aligned} \tag{3.2}$$

Then we have,

$$\begin{aligned} E\left(\boldsymbol{d}_i\boldsymbol{d}_i^T|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}\right) &= \int_{d_i} \boldsymbol{d}_i\boldsymbol{d}_i^T f(\boldsymbol{d}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) dd_i \\ &= \int_{d_i} \boldsymbol{d}_i\boldsymbol{d}_i^T \int_{y_i(C,Q)} f(\boldsymbol{d}_i|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}) f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) dy_{i(C,Q)} dd_i \\ &= \int_{y_i(C,Q)} \left(\int_{d_i} \boldsymbol{d}_i\boldsymbol{d}_i^T f(\boldsymbol{d}_i|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}) dd_i\right) f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) dy_{i(C,Q)} \\ &= \int_{y_i(C,Q)} E\left(\boldsymbol{d}_i\boldsymbol{d}_i^T|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}\right) f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) dy_{i(C,Q)}. \end{aligned} \tag{3.3}$$

Similarly, we have

$$E\left(\boldsymbol{e}_i^T\boldsymbol{e}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}\right) = \int_{y_i(C,Q)} E\left(\boldsymbol{e}_i^T\boldsymbol{e}_i|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}\right) f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}) dy_{i(C,Q)}. \tag{3.4}$$

The quantities $E\left(\boldsymbol{d}_i\boldsymbol{d}_i^T|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}\right)$ and $E\left(\boldsymbol{e}_i^T\boldsymbol{e}_i|\boldsymbol{y}_i, \hat{\boldsymbol{\theta}}\right)$ are the expectations that were given in the E-step in Chapter 2. Thus, for censored data, the E-step requires averaging the previous E-step (discussed in Chapter 2) over $\boldsymbol{y}$, consistent with the observed censoring pattern. To provide a general solution to the equations (3.3) and (3.4), Hughes (1999) prescribed using Monte Carlo methods paired with the Gibbs sampler approach (Gelfand and Smith, 1990). The method requires sampling from $\boldsymbol{y}_i$ from

$f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}})$ and then using the sample to compute

$$\begin{aligned} E\left(\boldsymbol{d}_i\boldsymbol{d}_i^T|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}\right) &\approx \sum_{l=1}^{L} E\left(\boldsymbol{d}_i\boldsymbol{d}_i^T|\boldsymbol{y}_i^l, \hat{\boldsymbol{\theta}}\right)/L \\ E\left(\boldsymbol{e}_i^T\boldsymbol{e}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}\right) &\approx \sum_{l=1}^{L} E\left(\boldsymbol{e}_i^T\boldsymbol{e}_i|\boldsymbol{y}_i^l, \hat{\boldsymbol{\theta}}\right)/L \\ E\left(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}}\right) &\approx \sum_{l=1}^{L} \boldsymbol{y}_i^l/L \end{aligned} \tag{3.5}$$

where $\boldsymbol{y}_i^l \sim f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}})$. The L samples are generated through the Gibbs sampler approach which requires an initial value of $\boldsymbol{y}_i$ that is chosen from a distribution that is close to $f(\boldsymbol{y}_i|\boldsymbol{C}_i, \boldsymbol{Q}_i, \hat{\boldsymbol{\theta}})$. Once an initial value has been selected, new values of $\boldsymbol{y}_i$ can be generated by iteratively sampling from the univariate conditional distributions given as $f(y_{ij}|y_{ik:k\neq j}, \hat{\boldsymbol{\theta}})$, for all $j = 1, \ldots, p_i$ where $C_{ij} \neq 0$.

The more popular methods suggested for modeling both non-participation and death are semi-parametric regression methods through weighting. Appropriately, these models are known as weighting generalized estimation equations (WGEE). The models proposed are usually modifications of the class of weighted estimating equations introduced by Robins et al. (1995). These models have been shown to provide consistent and asymptotically normal estimators of regression parameters given the probability of non-response at a given time t, which depends only on the past values of covariates and responses up to time t, that is, $t-1$. Further, the probability of a non-response model can be specified given the past observed data. These models are preferred because they are typically computationally simple and do not require the joint modeling of the response and missing mechanism (Rosenbaum, 1987). Rajan and Leurgans (2010) presented a weighted generalized estimation equations approach that accounts for death and monotone non-participation by treating the two categories of attrition as different events and modeling the two events separately.

Additionally, Kurland and Heagerty (2005) described handling monotone non-response in the regression conditioned on being alive (RCA) models. Shardell and Miller (2008) extends the literature on RCA models describing a weighted generalized estimating equation that estimates outcomes on those who are alive and considers death and non-monotone missing of time varying covariates and outcomes. The Rajan and Leurgans (2010) and Kurland and Heagerty (2005) approaches are presented below. Both models begin with the marginal mean regression model of $\boldsymbol{y}_i$ given $\boldsymbol{X}_i$,

$$E(\boldsymbol{y}_i|\boldsymbol{X}_i) = g^{-1}(\boldsymbol{X}_i\boldsymbol{\beta}).$$

This is the generalized linear model described in equation (1.3). Under missing completely at random (MCAR), these models have been shown to have consistent and asymptotically normal estimators of the regression parameters by solving the following generalized estimating equation (GEE):

$$U(\hat{\boldsymbol{\beta}}) = N^{-1/2}\sum_{i=1}^{N} \boldsymbol{D}_i^T \boldsymbol{V}_i^{-1}(\boldsymbol{y}_i - \boldsymbol{\mu}_i) = 0, \tag{3.6}$$

where $\boldsymbol{y}_i$ is a vector of responses for subject i, $\boldsymbol{\mu}_i = E(\boldsymbol{y}_i)$, $\boldsymbol{D}_i = \partial\boldsymbol{\mu}/\partial\boldsymbol{\beta}^T$, and $\boldsymbol{V}_i$ is a $p_i \times p_i$ invertible working covariance matrix of $\boldsymbol{y}_i$, where p_i is the length of $\boldsymbol{y}_i$ and $\boldsymbol{V}_i = \phi\boldsymbol{A}_i^{\frac{1}{2}}\boldsymbol{R}_i(\boldsymbol{\alpha})\boldsymbol{A}_i^{\frac{1}{2}}$. $\boldsymbol{A}_i$ is an $(p_i \times p_i)$ diagonal matrix with a variance function that is determined by the assumed probability distribution of the outcomes along the diagonal, and ϕ is a dispersion parameter that may be known or may be estimated from the data dependent upon the distribution assumption.

Rajan and Leurgans (2010) proposed a class of weighted estimating equations with two indicators of missing patterns–due to non-response and death. This was accomplished by defining two indicator variables R_{ij} and S_{ij} to indicate non-response and death, respectively. Three states were defined:

1. $R_{ij} = S_{ij} = 0$ signifies that the subject i was observed at occasion j.

2. $R_{ij} = 0$ and $S_{ij} = 1$ denotes that subject i was decease by time j. Death implies a monotone missing pattern, that is, if $S_{ij} = 1$ then $S_{i(j+1)} = 1$.

3. $R_{ij} = 1$ and $S_{ij} = 0$ indicates that subject i is alive but was not observed at time j.

The non-participation missing pattern was assumed to be monotonic. Moreover, covariates were assumed to have complete data. Further, the random variable pair (R_{ij}, S_{ij}) was assumed to satisfy the following probabilistic model for a subject being responsive or observed:

$$\begin{aligned} &P(R_{ij} = S_{ij} = 0 | R_{i(j-1)} = S_{i(j-1)} = 0, X_{ij}, y_{ij}) \\ &\quad = P(R_{ij} = S_{ij} = 0 | R_{i(j-1)} = S_{i(j-1)} = 0, X_{i(j-1)}, y_{i(j-1)}) \\ &\quad = P(R_{ij} = 0 | R_{i(j-1)} = S_{i(j-1)} = 0, X_{i(j-1)}, y_{i(j-1)}) \\ &\quad \times P(S_{ij} = 0 | R_{i(j-1)} = S_{i(j-1)} = 0, X_{i(j-1)}, y_{i(j-1)}) \end{aligned} \tag{3.7}$$

The above equation shows that the joint probability of the two indicator random variables given the past covariates and responses can be factored into two conditional probability distributions. These conditional probabilities of the non-participation and death indicators must be bounded and not equal to 0 to ensure consistent and asymptotic normal estimates (Robins and Greenland, 2000). Under this assumption, the probability of being responsive does not depend on a subject's current or future responses, which are akin to the missing at random assumption (MAR)(Rubin, 1976).

By assuming (3.7), we are able to identify $E(y_{ij}|X_{ij})$ in the presence of missing outcomes due to death and non-response in terms of the observed random variables,

R_{ij} and S_{ij}, as

$$E(y_{ij}|X_{ij}) = \int\int\ldots\int E(y_{ij}|R_{ij}, S_{ij}, X_{ij}, y_{i(j-1)}) \times \prod_{t=1}^{j} dF(y_{it}|R_{it}.S_{it}, X_{it}, y_{i(t-1)})dy_{ij} \tag{3.8}$$

This estimate as described is a weighted average of $E(y_{ij}|R_{ij}, S_{ij}, X_{ij}, y_{i(j-1)})$ with the specific weights $\prod_{t=1}^{j} f(y_{ij}|R_{ij}.S_{ij}, X_{ij}, y_{i(j-1)})$. If we denote $\psi_{ij} = P(R_{ij} = S_{ij} = 0|R_{i(j-1)} = S_{i(j-1)} = 0, X_{i(j-1)}, y_{i(j-1)})$, then ψ_{ij} can be defined using a vector of $q \times 1$ unknown parameters $\boldsymbol{\alpha}$. That is, $\psi_{ij} = \psi_{ij}(\boldsymbol{\alpha})$ which is usually chosen to be a multinomial function parameterized by $\boldsymbol{\alpha}$. Next, we define $\pi_{ij} = \psi_{i1}(\boldsymbol{\alpha}) \times \cdots \times \psi_{ij}(\boldsymbol{\alpha})$ to be the probability that subject i responds at time j. When assumption (3.7) holds, $\pi_{ij}(\boldsymbol{\alpha})$ is the conditional probability of observing participant i at time j given past data. This leads to a diagonal matrix of weight observations for subject i of the form $\Phi_i(\boldsymbol{\alpha}) = diag((1 - r_{i1})(1 - s_{i1})/\pi_{i1}(\boldsymbol{\alpha}), \ldots, (1 - r_{ij})(1 - s_{ij})/\pi_{ij}(\boldsymbol{\alpha}))$. In order to improve efficiency, the generalized estimating equations described in (3.6) was modified to include data from individuals with incomplete data. This was accomplished by adding a term with zero expectations to the estimating equations, and defining the new generalized estimating equation as

$$U(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}) = N^{-1/2} \sum_{i=1}^{N} \left\{ \boldsymbol{D}_i^T \boldsymbol{V}_i^{-1} \Phi_i(\hat{\boldsymbol{\alpha}})(\boldsymbol{y}_i - \boldsymbol{\mu}_i) - (\Phi_i(\hat{\boldsymbol{\alpha}}) - 1)\phi(\boldsymbol{y}_i; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}}) \right\} = 0 \tag{3.9}$$

where $\Phi_i(\hat{\boldsymbol{\alpha}})$ is the diagonal matrix of weights for subject i and $\phi(\boldsymbol{y}_i; \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\alpha}})$ is the conditional probability of $\boldsymbol{y}_i$ given the covariates and the observed response data. Using the form of the Horvitz-Thompson estimator (Horvitz and Thompson, 1952), the estimator of the mean is given as

$$\hat{\boldsymbol{\mu}} = N^{-1} \sum_{i=1}^{N} \left\{ \frac{(1 - \boldsymbol{R}_i)(1 - \boldsymbol{S}_i)\boldsymbol{y}_i}{\pi(\boldsymbol{X}_i, \hat{\boldsymbol{\alpha}})} - \frac{(1 - \boldsymbol{R}_i)(1 - \boldsymbol{S}_i) - \pi(\boldsymbol{X}_i, \hat{\boldsymbol{\alpha}})}{\pi(\boldsymbol{X}_i, \hat{\boldsymbol{\alpha}})} E(\boldsymbol{y}_i|\boldsymbol{R}_i.\boldsymbol{S}_i, \boldsymbol{X}_i) \right\}. \tag{3.10}$$

Similar to Rajan and Leurgans (2010), Kurland and Heagerty (2005) modified their regression conditioning on being alive (RCA) models by adapting the inverse probability of censoring-weighted generalized estimating equations (IPCW-GEE) (Robins and Greenland, 2000). This approach, as seen above, involves modeling the drop-out pattern. To apply IPCW-GEE to estimate $\mu_{ij}^A = E(y_{ij}|X_{ij}, S_i > t) = E(y_{ij}|X_{ij}, A_{ij} = 1)$ weights, π_{ij}^A must be estimated so that

$$E\left(\frac{R_{ij}}{\pi_{ij}} y_{ij}|X_{ij}, S_i > t\right) = E(y_{ij}|X_{ij}, S_i > t),$$

where R_{ij} reflects the missing status and $\pi_{ij}^A = P(R_{ij} = 1|X_{i(j-1)}, Y_{i(j-1)}, S_i > t)$.

Additionally, Kurland and Heagerty (2005) offer a hierarchy of missingness of RCA models to highlight when drop-out is ignorable and to define how estimation methods can be altered to accommodate missing data (Table 3.1). In the case of missing completely at random (MCAR), $P(R_{ij} = 1|y_{i1}, \ldots, y_{1(j-1)}, yij, S_i > t_j) = P(R_{ij}|S_i > t_j)$. By taking the expected value with respect to the joint probability of the response ($\boldsymbol{y}$) and the pattern of dropout $\boldsymbol{R}$ given the subject being alive ($A_{ij} = 1$), estimates of $\boldsymbol{\beta}^A$ can be shown to be consistent.

$$\begin{aligned}
E_{y,R|A}\left[U(\boldsymbol{\beta}^A)\right] &= E_{y,R|A}\left\{\sum_{i=1}^{N}\sum_{j=1}^{p_i} A_{ij}\frac{\partial \mu^{y|A}}{\partial \beta} R_{ij}(y_{ij} - \mu_{ij}^{y|A})\right\} \\
&= \sum_{i=1}^{N}\sum_{j=1}^{p_i} A_{ij} E_{y,R|A}\left[R_{ij}\frac{\partial \mu^{y|A}}{\partial \beta}(y_{ij} - \mu_{ij}^{y|A})\right] \\
&= \sum_{i=1}^{N}\sum_{j=1}^{p_i} A_{ij} Ey|A\left[\frac{\partial \mu^{y|A}}{\partial \beta}(y_{ij} - \mu_{ij}^{y|A}) E_{R|y,A}(R_{ij} = 1|A_{ij})\right] \qquad (3.11) \\
&= \sum_{i=1}^{N}\sum_{j=1}^{p_i} A_{ij} Ey|A\left[\frac{\partial \mu^{y|A}}{\partial \beta}(y_{ij} - \mu_{ij}^{y|A})\right] P(R_{ij}|A_{ij} = 1) \\
&= 0, \text{ if } \mu_{ij}^A = E(y_{ij}|X_{ij}, A_{ij} = 1).
\end{aligned}$$

The above equation shows that under this missing assumption, estimation of $\boldsymbol{\beta}^A$ is consistent if the model is specified correctly. MCAR missing does not affect the consistency of the estimates but could impact the efficiency of regression estimators of $\boldsymbol{\beta}^A$.

Under the missing at random (MAR) assumption, we have

$$\begin{aligned} P(R_{ij} = 1 | y_{i1}, \ldots, y_{1(j-1)}, y_{ij}, S_i > t_j) \\ = P(R_{ij} = 1 | y_{i1}, \ldots, y_{1(j-1)}, S_i > t_j) \\ = \pi_{ij}^A \end{aligned} \tag{3.12}$$

For this case, it can be shown that the dropout process is not ignorable and must be modeled correctly to obtain valid inference about μ_{ij}^A. On the other hand, the model of survival is ignorable. Again, we weight the quasi-score equation by the inverse of the censoring weights, $\pi_{ij}^A = P(R_{ij} = 1 | y_{i1}, \ldots, y_{i(j-1)}, S_i > t_j)$ and take the expectation of the observed data distribution to obtain

$$\begin{aligned} E_{y,R|A}\left[U(\boldsymbol{\beta}^A)\right] &= \sum_{i=1}^{N} \sum_{j=1}^{p_i} A_{ij} E_{y|A}\left[E_{R|y,A}\left(\frac{R_{ij}}{\pi_{ij}^A}\right) \frac{\partial \mu^{y|A}}{\partial \beta}(y_{ij} - \mu_{ij}^{y|A})\right] \\ &= \sum_{i=1}^{N} \sum_{j=1}^{p_i} A_{ij} E_{y|A}\left[\frac{\pi_{ij}^A}{\pi_{ij}^A} \frac{\partial \mu^{y|A}}{\partial \beta}(y_{ij} - \mu_{ij}^{y|A})\right] \text{ if } R_{ij} \text{ is MAR} \\ &= 0 \text{ if } \mu_{ij}^A = E(y_{ij} | X_{ij}, S_i > t_j). \end{aligned} \tag{3.13}$$

Whereas MCAR and MAR model the missing pattern given being alive, $S_i > t_j$, MCAR-S and MAR-S models the missing pattern conditioned on being alive and survival time, $S_i = s, s > t_j$. These missing assumptions can accommodate dropout for RCA data in a similar manner using the IPC weights as cited in Table 3.1. Shardell and Miller (2008) extended the work of Kurland and Heagerty (2005) to show how to specify and estimate appropriate weights for RCA models for non-monotonic missing

in the outcome and covariates.

3.2 Vulnerabilities of the Survival Incorporating Models

Unconditional models have been described as being inappropriate if death is dependent on observed and unobserved responses, $f(\boldsymbol{R}|\boldsymbol{y^o}, \boldsymbol{y^m})$. In Chapter 2, evidence was presented to suggest that missing due to death is missing not at random (MNAR) for each response – depression, systolic and diastolic blood pressures, and physical functioning dependence. Further challenges of the unconditional model are to estimate unbiased estimates of the population regression parameters with missing due to non-response.

As defined and when used with missing due to death only, the pattern-mixture and terminal decline methods regress the response value over complete data (sample) for the population of interest. However, the inclusion of non-response for other reasons compromises the ability of the estimation process to produce unbiased estimators. Subsequently, the parameters of the regression mean will be under the same scrutiny as the unconditional models. Under MCAR, the estimates will be unbiased and the inference will remain valid. Similarly, MAR data will produce unbiased estimates, but dissimilarly, the target of inference will be altered. Efficiency for either case is not as strong as it would be under completely observed data.

Without the inverse probability weights and the correct specifications of the missing pattern and mean model, the generalized estimating equations would be unlikely to produce a consistent and asymptotically normal regression estimator in the presence of non-MCAR non-response. The methods of principal stratification, partly conditional, and joint, as models defined in this dissertation, evoke generalized estimating equations to contribute to the estimation of their mean regression parameters. Those components dependent on GEE for estimation will be susceptible to bias.

3.3 NC EPESE Data with non-participation and death

For the purpose of this analysis, 26 of the subjects were excluded because they identified themselves as other than black or white. Another four subjects were not included because their ages were less than 65 at the baseline survey. The analysis sample was completed by excluding one subject for not providing a value for any of the outcomes for the initial interview. The resulting baseline dataset was 34.9% male and 54.6% black with a mean age of 73.57-years-old. During the 10-year study period 2,045 of the 4,131 remaining subjects in the sample had death dates before the end of the study. A total of 712 individuals either died during the study but dropped from the study at a survey that was prior to the measurement occasion of their deaths or survived during the surveillance period but exited the study before the study concluded. When determining missing status, individuals were only categorized as missing if all of the four outcomes had missing values at a given measurement occasion. Those subjects who had a date of death that occurred before the end of the surveillance period and a response for the survey period that occurred just before death were considered as dropout due to death. Participants that had a response for at least one of the outcome measurements at the third follow-up were considered as completers. All other individuals were labeled as the cohort that had non-response for other reasons. The cohort of subjects with non-response due to death was more likely to be black, male, and older on average than the cohorts of the completers or those with non-response for other reasons (Table 3.2).

3.4 Results

Figures 3.1-3.4 present the resulting trajectories of the analysis of the NC EPESE data with non-response and death using the different proposed methods for assessing the rates of change based on research inquiries and truncation due to death for depression

scores (CES-D), systolic and diastolic blood pressures, and physical function (ADLs), respectively, for those subjects who were 73.57-years-old at baseline. In all of the models, the baseline covariates were age (centered about the baseline mean), race (1 if subjects identified as black and 0 if subjects identified as white), and gender (1 if subjects are identified as male and 0 if subjects are identified as female). Table 3.3 provides the annual rates of change for the first five years for each regression method by outcome for the data with missingness due to death and non-response and for the data with missingness due to death only (imputed values).

3.4.1 Depression

In the previous analysis of imputed data, the unconditional method yielded estimates that were similar across the race-gender groups. This conclusion was upheld for the unconditional analysis of the dataset without any imputation, except that the estimates of the rates of change were approximately 0.06 units smaller and the rate of change was no longer statistically significant. As in the previous unconditional analysis, the estimates for the intercept, gender, race, and age remained significant for the new analysis.

Those individuals who did not survive beyond the first follow-up survey demonstrated a decline in their number of depression symptoms; nonetheless, the remaining two cohorts exhibited increases in depression symptoms over time. The race-gender groups for the second follow-up and final follow-up survival cohorts had annual rates of increase that ranged from 0.04 to 0.06, with the exception of white males in the second follow-up survival cohort whose slope was 0.16. All of these slopes are less than the estimated slopes from the analysis of the imputed data (0.10-0.24). Moreover, the annual rates of change in the existing analysis for the 6- and 10-year survival cohorts were more similar than their rates in the previous analysis. The effect of the covariates

had no change from the previous results for those individuals with death dates beyond the end of the study surveillance period. Only the intercept and gender effects were statistically significant in the analysis of the data with non-response and truncation due to death for the first follow-up survival cohort. The age effect, which was significant in the analysis of the imputed data, was no longer significant for the unaltered dataset. In the second follow-up survival cohort for depression, intercept and gender remained significant and race and the effect of gender over time became statistically significant.

The principal stratification method for estimating depression concluded that the annual rates of change were small and negative across race-gender groups for the present data analysis, while the results from the previous results were small and positive across race-gender groups. Over the first five years, as demonstrated in Figure 1c, depression symptoms had a significant decline annually for the principal stratification model. The negative slopes in the principal stratification were also seen in the slopes of the race-gender cohorts for the pattern-mixture model for the first follow-up survival cohort, yet the baseline values of depressive symptoms were similar to the initial values of the third follow-up survival cohort (Figures 3.1b-c).

In the terminal decline analysis, black females nearly had no slope, which is 0.12 units lower than the slope of the black females in the terminal decline results of the imputed data. White females and black males had comparable slopes in the current analysis, which was similar to the results for the previous data, yet the data analysis of the NC EPESE without imputation was about 0.07 units lower. In the previous result, the terminal decline rates of change were akin to the pattern-mixture's second follow-up survival cohort conclusions. Nevertheless, the terminal decline findings for the unaltered dataset do not share this trend. In the previous analysis, the intercept, gender, and trend effects were significant. These variables remained significant in the terminal decline analysis of the unaltered data, along with the trend for race becoming

significant and the age covariate losing significance.

The partly conditional regression of the NC EPESE without imputation resulted in small and negative annual rates of change for the first five years, while the race-gender annual rates of change for the partly conditional analysis performed on the imputed data were small but positive. The slopes for each race-gender are like those reported from the principal stratification. This was not the case for the imputed data.

The other regression methods resulted in different trends and significant covariates in their analysis of the dataset that allowed non-response to be missing for non-death causes than the results from the data with truncation only due to death. In contrast, the joint model that offers a method to account for the survival and the non-mortality outcome values simultaneously produced similar annual rates of change for both datasets.

3.4.2 Systolic Blood Pressure

The unconditional modeling of the systolic blood pressure on the NC EPESE dataset without imputation resulted in varying annual rates of change for each race-gender groups, unlike the imputed data results that expressed comparable slopes by gender. Black men experienced the greatest annual decline (-0.30), which was also true in the unconditional analysis of the NC EPESE representing non-response only due to death (-0.24). The decline for white males trailed just behind black males as in the previous analysis of the death-only dataset, except at a lesser magnitude (-0.17 vs. -0.22). Black women declined at a rate of 0.06 annually compared to the 0.11 annual decline in systolic blood pressure for the non-response due to death-only dataset. Unexpectedly, white females experienced an increased annual rate of change (0.06) instead of a decline (-0.09) as in the prior results of the analysis of the death-only dataset. Nearly all of the slopes from the previous model rendered annual rates of change that were similar

by gender, yet this pattern was not sustained in the current analysis.

Analysis for those individuals who were alive at the 3-year survey but not beyond concluded with rates of annual change for the female groups that were higher than the rates of change from the analysis with the modified NC EPESE dataset (Figure 3.2 b). On the other hand, the slopes for males in the original data were lower than the rate for males in the imputed data. Notably, the annual rate of change for white males was drastically different (-1.28 vs. -2.97). Additionally, estimated annual rates of change for the second follow-up survival cohort were lower for the females in the dataset that allowed other reasons for non-response than the dataset with only dropout due to death and higher for the males in the dataset with all missing than the males with death non-response only with alarming differences for white females (-0.24 vs. -0.55) and black males (-0.52, -0.31). The analysis for those individuals who were alive until the end of the study but could have non-response had drastically higher rates of change for females than the females who did not have non-response. Black males also experienced an increase but it was more modest. White males in the dataset that had non-response not due to death produced a smaller annual rate of change than the rate of change for the males in the dataset with dropout due to death.

Gender and the intercept were statistically significant in the original analysis for all three survival cohorts. In the pattern-mixture analysis for the data with MCAR non-response, only the intercept was significant for the first follow-up survival cohort. Age became significant while the change-over time no longer explained the variation in the changes in systolic blood pressure for the second follow-up survival cohort. For the last survival cohort, the trend over time joined gender and intercept as significant covariates.

All of the race-gender groups for the all-inclusive non-response NC EPESE had positive estimated slopes for the principal stratification trajectories, whereas the

estimated slopes for males in the principal stratification analysis of the imputed NC EPESE experienced a decline. Furthermore, Table 3.3 shows that the slopes from two datasets are highly dissimilar in magnitude. Moreover, the estimates of the slopes for females in the non-manipulated NC EPESE were markedly higher than the data with imputation, 0.51 vs. 0.04 for black females and 0.33 vs. 0.12 for white females.

When comparing the annual rates of change estimates from the analysis completed on the current NC EPESE and the imputed NC EPESE, the terminal decline estimates have the least differences. All of the slopes report decreasing trends for each race-gender group as previous reported from the analysis on the modified dataset. Moreover, the magnitudes were alike with the exception of the slope reported for white females, which is lower in magnitude (-0.31 vs. -0.58). In the new analysis, the years from death was no longer significant, but the intercept and gender effect sustained their significance.

The estimates from the partly conditional model and the joint model were very similar for the race-gender groups in the present dataset as in the previous imputed dataset. All of the estimates of rates of change declined; white females had a smaller decline in the unaltered dataset than in the modified one for the partly conditional analysis. Almost identical conclusions were produced in both datasets.

3.4.3 Diastolic Blood Pressure

The analysis for the diastolic blood pressure did not have any difference in the covariates that were statistically significant for the two datasets for any of the methods. Consequently, the estimated annual rate of change for the first five years was nearly the same for each of the methods for the two datasets. Notably, in the principal stratification model the estimates were lower for the data with all types of non-response than the death-only drop-out dataset, except for white females.

3.4.4 Activities of Daily Living

Similar to the diastolic blood pressure regression results, the activities of daily living had almost identical conclusions for each method for both datasets. However, the inference had some minor changes. The unconditional and the partly conditional regression analyses of the NC EPESE, allowing all drop-out types, resulted in an addition of the significance of a race trend. The pattern-mixture, terminal decline, and joint models had similar annual rates of change and the same significant covariates for each dataset. The other methods produced almost identical results in the analysis of the two datasets, but the principal stratification method did not. Each of the annual rates of increase was lower in the analysis of the non-imputed data than in the imputed data.

3.5 Simulation of MAR and NMAR Death and Non-Participation

In order to assess the proposed models' abilities to accurately estimate the means of the Center for Epidemiologic Studies Depression (CES-D) scores, systolic and diastolic blood pressures, and activities of daily living (ADL) of a complete dataset without bias when a subject's missing status (death or non-response) is dependent on study covariates and the response, a sample was generated from a theoretical population. Each of the four outcomes were treated as continuous outcomes and were generated from the mixed model with a random intercept and slope as described in equation (2.1). Four waves of longitudinal outcomes were simulated from a normal distribution with mean $\boldsymbol{X_i\beta}$ and covariance $\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i\boldsymbol{D}_i\boldsymbol{Z}_i^T + \sigma^2\boldsymbol{I}_i$, where $\boldsymbol{D}$ was allowed to be unstructured. For each outcome, the design matrix $\boldsymbol{X}_i$ consisted of a column vector of ones for the intercept, a column vector for time of measurements post-baseline (0, 3, 6, 10 years), a column vector of baseline age centered about the mean, a column vector of indicators for identifying as black, a column vector of indicators for identifying as male,

a column vector indicating time by race, and a column vector indicating time by sex. The design matrix of the random effects, ($\boldsymbol{Z}_i$), was constructed as a column vector of ones for the intercept and a column vector for the time of measurements post- baseline (0, 3, 6, 10 years). The race-gender combinations values were treated as multinomial random variables and were generated accordingly, white males ($\pi = 0.16$), black males ($\pi = 0.19$), and white females ($\pi = 0.29$). Age was simulated while assuming it was from a normal distribution dictated by the mean and standard deviation in each race-gender group from the NC EPESE. Similarly, time of measurements mirrored the NC EPESE. Thus, we assume measurements were only possible at baseline, and 3, 6, and 10 years post-baseline. The values of the parameters $\boldsymbol{\beta}, \sigma^2$, and $\boldsymbol{D}$ used in the simulation are given for each outcome below.

CES-D score: $\boldsymbol{\beta}^T = \begin{pmatrix} 3.243 & 0.086 & 0.065 & 0.356 & -0.686 & 0.001 & 0.003 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 5.32 & -0.14 \\ -0.14 & 0.39 \end{pmatrix} \boldsymbol{Z}_i^T + 7.16\boldsymbol{I}_i$$

Systolic BP: $\boldsymbol{\beta}^T = \begin{pmatrix} 143.29 & 0.088 & 0.002 & 0.756 & -2.326 & -0.017 & -0.136 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 139.69 & -3.38 \\ -3.38 & 0.91 \end{pmatrix} \boldsymbol{Z}_i^T + 314.05\boldsymbol{I}_i$$

Diastolic BP: $\boldsymbol{\beta}^T = \begin{pmatrix} 77.258 & 0.799 & -0.289 & 2.775 & 1.529 & -0.0006 & -0.010 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 47.66 & -2.26 \\ -2.26 & 0.39 \end{pmatrix} \boldsymbol{Z}_i^T + 92.94\boldsymbol{I}_i$$

ADL score: $\boldsymbol{\beta}^T = \begin{pmatrix} 0.742 & 0.211 & 0.114 & 0.243 & 0.133 & 0.001 & -0.063 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z} \begin{pmatrix} 1.52 & -0.002 \\ -0.002 & 0.04 \end{pmatrix} \boldsymbol{Z}^T + 1.78\boldsymbol{I}_i$$

A 1,000 samples of complete data were generated for each sample size – $N = 100$, $N = 500$, and $N = 1000$. After the complete datasets were generated, missing indicators were created to represent participants leaving the study because of death or non-participation. Subjects in the simulated datasets became at risk of death or dropout following baseline responses.

For convenience, an ordinal ranking for the missing categories, such as death ($k = 1$), non-response ($k = 2$), and completers was assumed. Following this assumption, a cumulative regression model was used to simulate missing categories with known probabilities for each wave for each outcome with dependence on the baseline covariates and the previously observed response values. Missing at random due to death and non-response was modeled using the cumulative regression model described below:

$$logit(\gamma_{ik}) = logit[P(y_i \leq k)] = \eta_k + \boldsymbol{x}^T\boldsymbol{\beta} \tag{3.14}$$

with $\boldsymbol{x}_i$ representing a vector of the covariates – sex (male=1, 0 otherwise), race

(black=1, 0 otherwise), baseline age centered about the baseline mean, lag measurement occasion, $t_{i(j-1)}$ (lag wave), the additional effects on the probability over the lag measurement occasion given one identifies as male (male-lag wave interaction) and given one identifies as black (black-lag wave interaction), and the previous (lag) response outcome value $(y_{i(j-1)})$– and $\boldsymbol{\beta}$ representing their corresponding regression parameters: $\boldsymbol{\beta}^T = \begin{pmatrix} 0.1 & 0.05 & 0.001 & 0.001 & -0.02 & -0.02 & .01 \end{pmatrix}$. η_k represents the intercept for the k cumulative logit. The η_k by outcome is given as

$$\text{CES-D Score: } \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -1.6 \\ -1.2 \end{pmatrix} \quad \text{Systolic BP: } \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -3.0 \\ -2.6 \end{pmatrix}$$

$$\text{Diastolic BP: } \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -2.4 \\ -2.0 \end{pmatrix} \quad \text{ADL: } \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -1.6 \\ -1.2 \end{pmatrix}$$

Using the estimates of the probabilities as a parameter in a random Bernoulli generator, these models led to 31% missing due to death and 13% missing due to non-response for each outcome.

Non-missing at random (NMAR) indicators were simulated in a similar manner. For the NMAR assumption, we allow the probabilities to depend on the current response outcomes (y_{ij}). The design matrix, $\boldsymbol{x}_i$, described above replaces $t_{i(j-1)}$ with t_{ij} and $y_{i(j-1)}$ with y_{ij}. The parameter vector, $\boldsymbol{\beta}^T$ and the outcome-specific intercepts, η_k's, remain the same as defined for the previous MAR models. Using the probabilities estimated from the new cumulative models, we obtained 30% of the individuals missing due to death and 13% missing due to non-participation for each outcome.

3.6 Simulations Results

We chose to fit the unconditional, pattern-mixture, and partly conditional models as described earlier for each of the 1,000 samples of the $N = 500$ and $N = 1000$

simulated datasets with the MAR and NMAR missing profiles. All analyses were performed using SAS v9.2. For the unconditional and pattern-mixture models, maximum likelihood estimation was provoked. The partly conditional models were estimated by generalized estimating equations using an identity working correlation matrix and empirical standard errors to account for the repeated continuous measures per subject. The bias of the estimation for each mean per race-gender combination was computed. Tables 3.4 - 3.7 give the mean relative bias for each method.

When the missingness was assumed to be not missing at random (NMAR – dependent on outcomes that may not be observed), the mean estimates of CES-D scores and systolic and diastolic blood pressures for the immortal cohort, the study completers, and the dynamic survival group were underestimated (negative relative bias). Moreover, the unconditional model, which is modeled using the linear mixed model, was able to estimate the mean depression and blood pressure values with minimal bias. However, the linear mixed model was not as robust against bias when only information from the completers was utilized. As one would expect, the partly conditional model, which is estimated by generalized estimating equations using an identity working correlation, did not estimate the means of the outcomes without substantial bias for either missing at random (MAR) or not missing at random (NMAR). The outcome measuring physical functional limitations did not support any trends regarding the bias present in the estimation by missing assumption, samples size, or model type.

3.7 Discussion

Through applying the advanced longitudinal regression methods that incorporate survival in a dataset with outcomes truncated due to death and non-response and comparing the results to a previous analysis from a dataset with outcomes truncated due to death only, we have revealed that some of the methods were able to provide

similar results despite the presence of non-response for other reasons. The annual rates of change in Table 3 for the joint models are nearly identical for both sets of data. Following these results, evidence suggests that the estimates of probability of being alive and healthy (PAH) were the most steady. The consistency of the PAH values are estimated in a manner that is similar to the weighted estimating equations.

Despite the percentage of values that were imputed in the previous analysis, annual rates of change were primarily alike by model for activities of daily living (ADLs) and diastolic blood pressure. Although Table 3.2 indicates that systolic and diastolic blood pressures had comparable missingness due to death and non-response, systolic blood pressure annual rates of change and inference were not similar to previous results for any of the statistical modeling methods. Correspondingly, depression had dissimilar outcomes by model but its percentage of non-response not due to death was more than 60% greater than the percentage of non-response for ADLs. One possible explanation for the models different performance for the outcomes of systolic and diastolic blood pressures is that the systolic blood pressures contained a higher level of variance than diastolic blood pressures. This variance in systolic blood pressures was likely reduced due to the single imputation of those values that were not truncated due to death.

By conducting simulations, we were able to gain distinctive evidence of the models' performance under varying conditions. These simulation results allowed us to compare the models' ability to estimate the true means of the original population when missing is due to MAR or NMAR death and non-response. The biases for the NMAR missing scenario for these outcomes were much greater in magnitude and were typically negative. The underestimation of the population means provided evidence of a possible "healthy survivor" effect influencing the estimates. Further, the simulation results provided evidence that suggest that investigators should be cautious when choosing an estimation model and should examine which missing assumptions can be considered for

their data.

3.8 Conclusion

Very few studies have been published on the performance of the proposed models for non-mortality outcomes truncated due to death. This study presented a comparison analysis of the models analyzing data with missing due to death with imputed non-response and data with MAR or NMAR missing due to death and non-response to better understand how the assumptions affected the estimates and inference. Some of the proposed methods for analyzing data to account for survival have been shown to be more sensitive to imputation than other models. The simulations for depression and blood pressure measures were in support of using unconditional models when death and non-response is missing at random (MAR). Without proper weights, the simulation results supported the fact that the linear mixed model is able to produce minimal biased estimates under the assumption of MAR and assuming the sample is immortal, unlike the partly conditional model.

Figure 3.1: Fitted trajectories of CES-D scores for EPESE participants

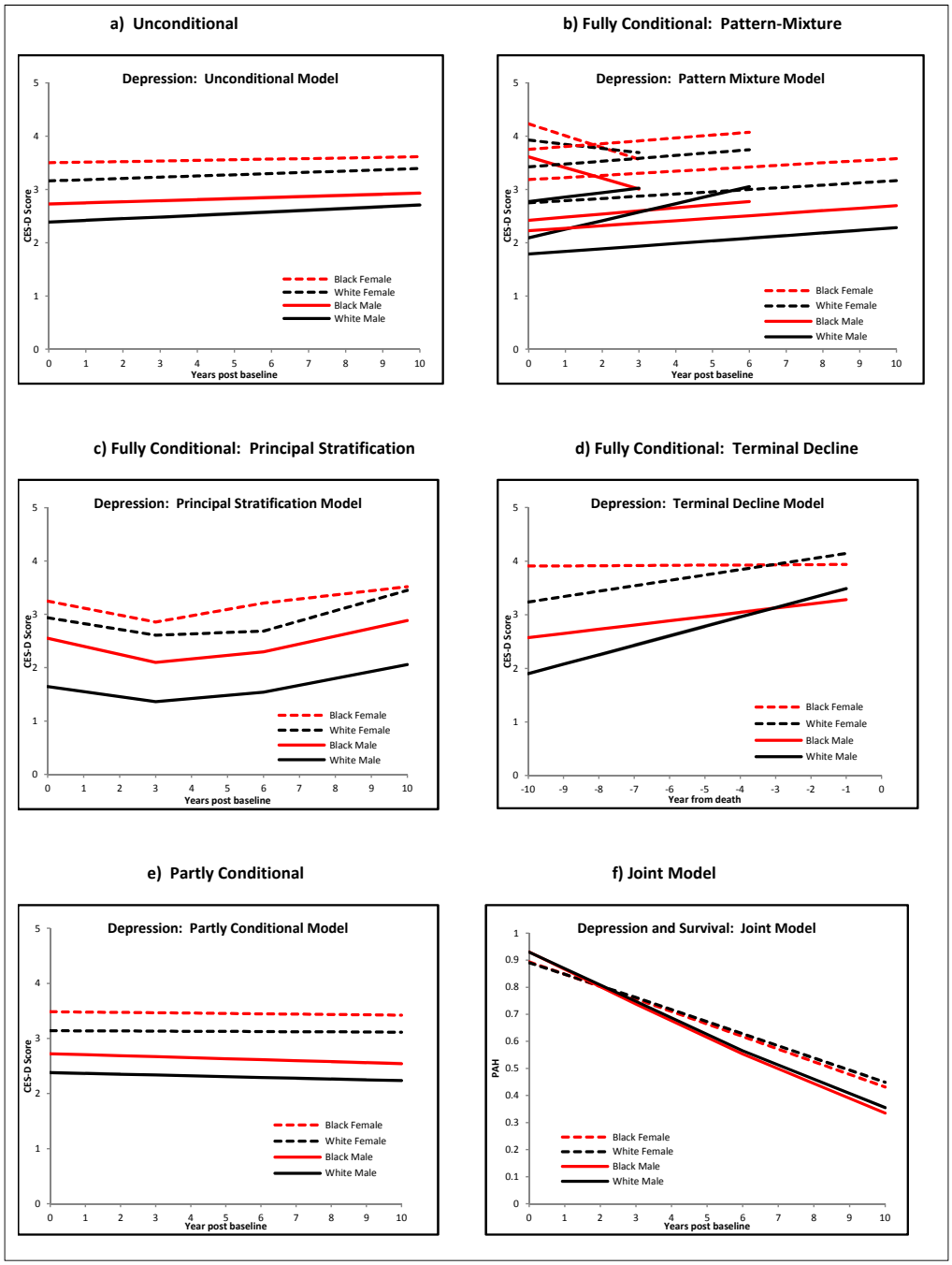


Figure 3.2: Fitted trajectories of systolic blood pressure for EPESE participants

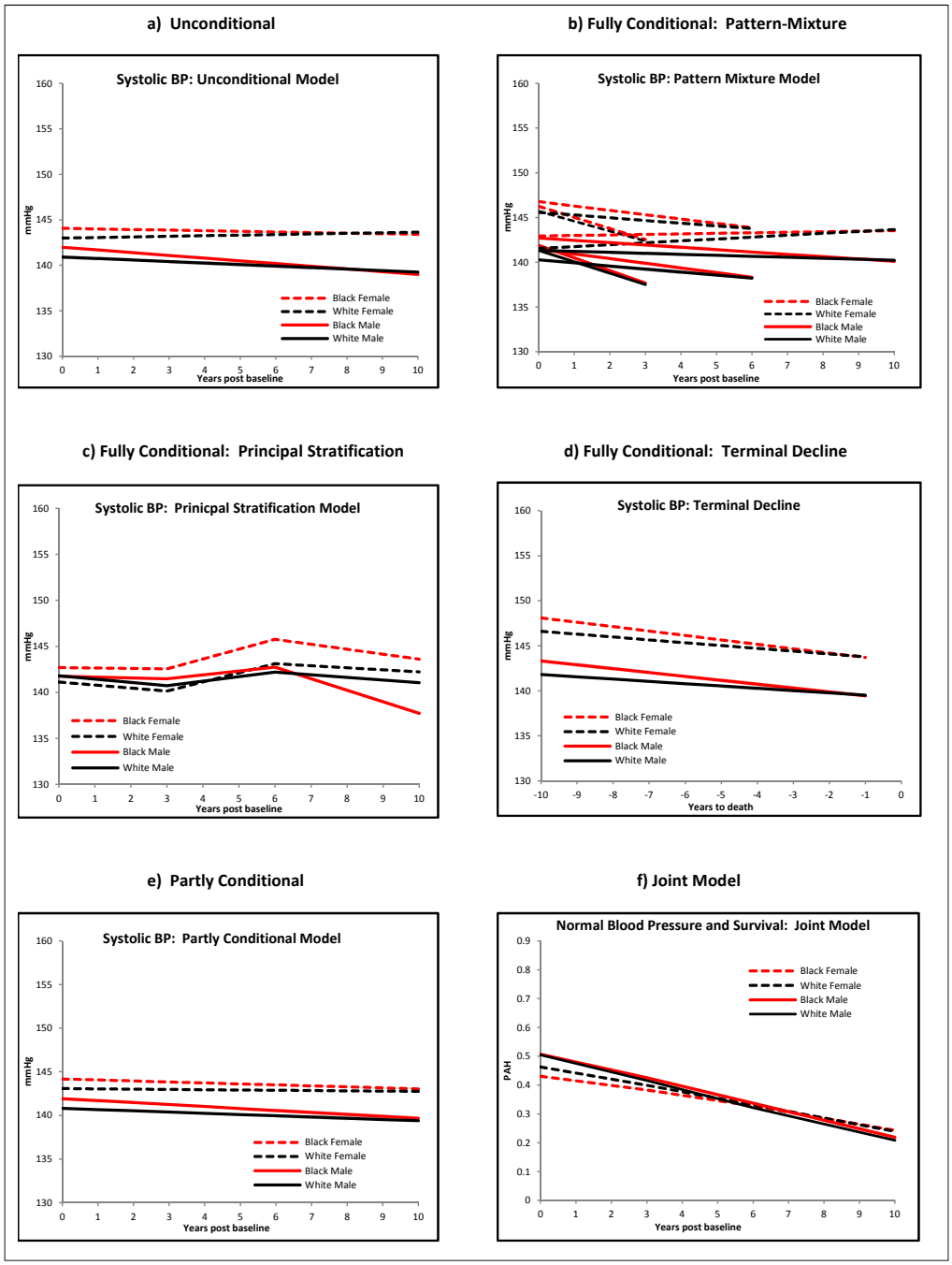


Figure 3.3: Fitted trajectories of diastolic blood pressure for EPESE participants

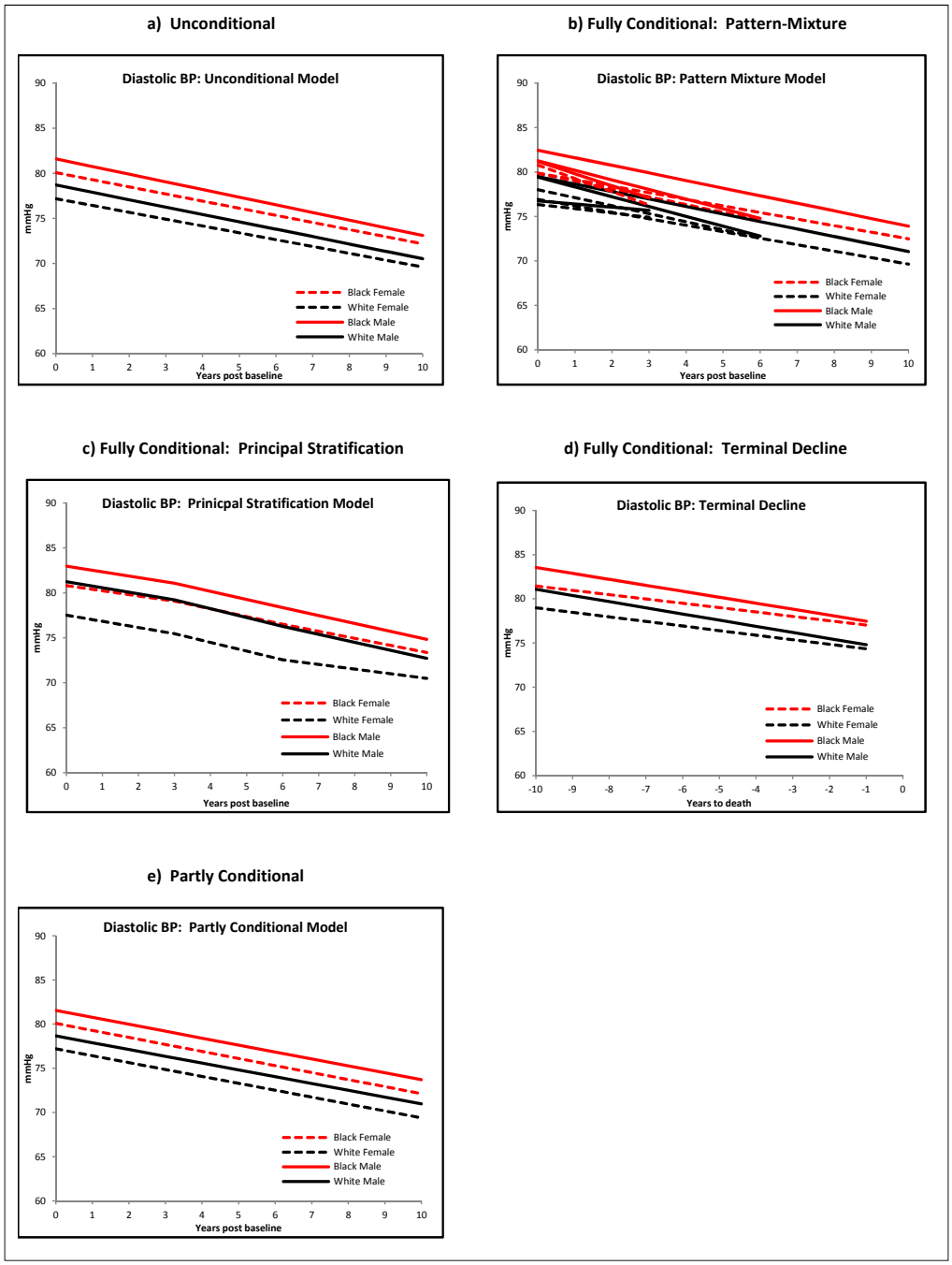


Figure 3.4: Fitted trajectories of ADL Scores for EPESE participants

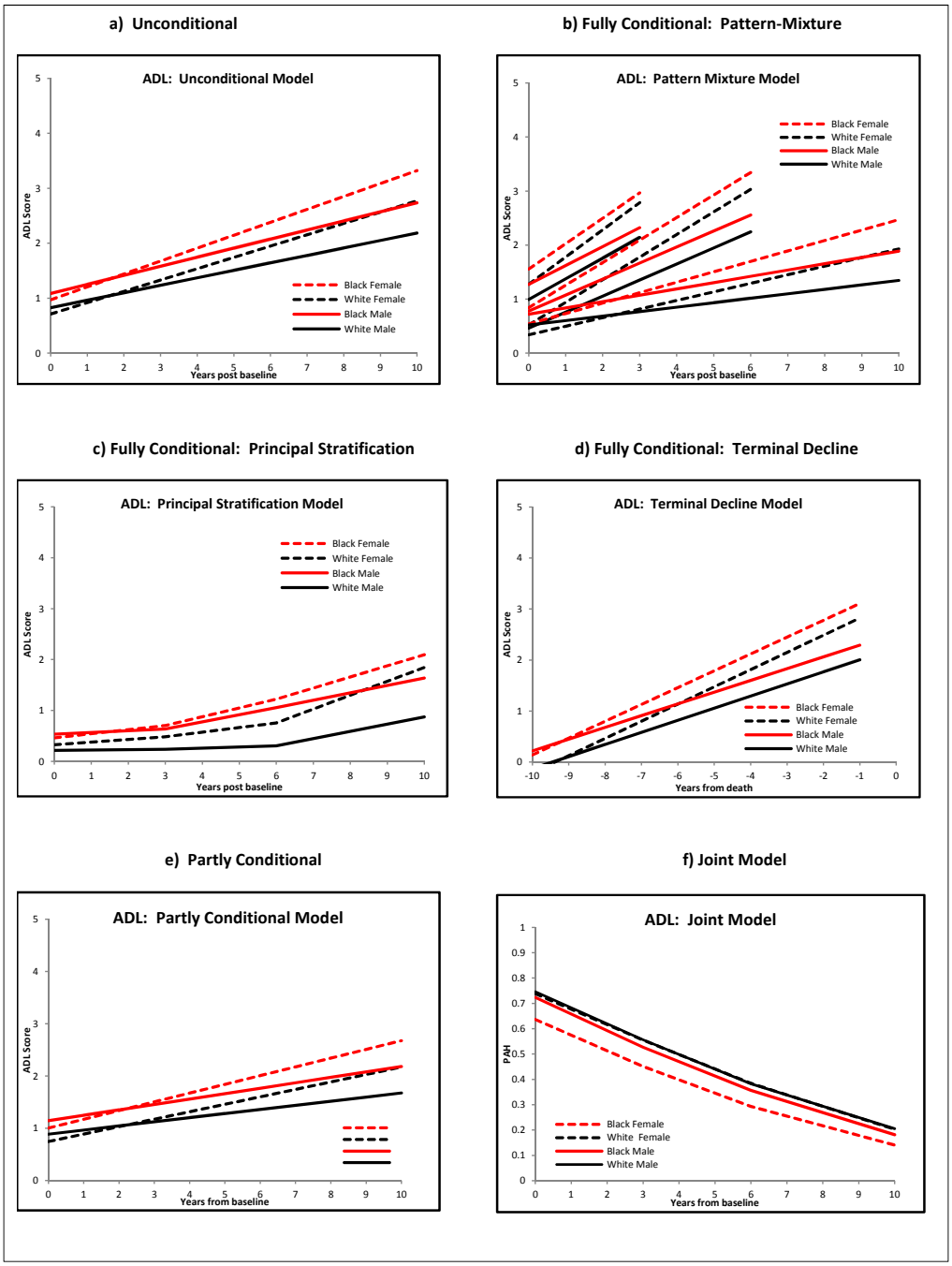


Table 3.1: Monotone dropout in the outcome hierarchy and required IPC weights for directly parameterized RCA models

Dropout pattern	Dropout pattern assumption: $P(R_{ij}=1\|\boldsymbol{y}_i, S_i=s)$ becomes	**IPC weights π_{ij} for IEE quasi-score**
MCAR	$P(R_{ij}=1\|S_i>t_j)$	None: $f(\boldsymbol{R_i})$ is ignorable
MCAR-S	$P(R_{ij}=1\|S_i=s)$	$\pi_{ij}^S = P(R_{ij}=1\|S_i=s, S_i>t_j)$
MAR	$P(R_{ij}=1\|S_i>t_j, y_{i1},\ldots,y_{i(j-1)})$	$\pi_{ij}^A = P(R_{ij}=1\|y_{i1},\ldots,y_{i(j-1)}, S_t>t_j)$
MAR-R	$P(R_{ij}=1\|S_i=s, y_{i1},\ldots,y_{i(j-1)})$	$\pi_{ij}^S = P(R_{ij}=1\|y_{i1},\ldots,y_{i(j-1)}, S_i=s, S_i>t_j)$

Table 3.2: Baseline demographics by drop-out categories for each outcome

Missing Type	**Outcomes**				
	Depression	**Systolic BP**	**Diastolic BP**	**ADL**	**Overall**
Due to Death (N)	1416	1581	1570	1760	1787
Male N (%)	611 (43)	658 (42)	655 (42)	729 (41)	742 (41)
Black N (%)	792 (56)	885 (56)	880 (56)	1000 (57)	1015 (57)
Age Mean (sd)	74.9 (7.03)	75.6 (7.34)	75.6 (7.36)	75.8 (7.4)	75.8 (7.38)
Base Outcome Mean (sd)	0.11 (0.32)	143.93 (21.13)	78.88 (12.44)	1.54 (2.19)	
Other Reasons (N)	1163	806	818	758	712
Male N (%)	334 (29)	245 (30)	250 (31)	232 (31)	211 (30)
Black N (%)	635 (55)	417 (52)	423 (52)	387 (51)	361 (51)
Age Mean (sd)	74.5 (6.60)	73.3 (6.24)	73.3 (6.22)	73.2 (6.30)	73.2 (6.26)
Base Outcome Mean (sd)	0.10 (0.30)	143.92 (20.69)	79.03 (11.81)	0.61 (1.39)	
Completers (N)	1389	1479	1473	1603	1632
Male N (%)	437 (31)	455 (31)	453 (31)	477 (30)	490 (30)
Black N (%)	723 (52)	804 (54)	802 (54)	864 (54)	882 (54)
Age Mean (sd)	70.7 (4.78)	71.2 (5.10)	71.2 (5.08)	71.3 (5.18)	71.3 (5.19)
Base Outcome Mean (sd)	0.73 (0.26)	142.03 (19.47)	79.89 (11.43)	0.45 (1.20)	

Table 3.3: Outcomes' annual rates of change for each model for the non-response due to all reasons dataset and non-response due to death-only dataset

Outcomes	Black Female		White Female		Black Male		White Male	
Models	All	Death only	All	Death only	All	Death only	All	Death only
Depression								
Unconditional	0.01	0.07	0.02	0.09	0.02	0.09	0.03	0.09
Pattern-Mix 1	-0.10	-0.12	-0.10	-0.12	-0.20	-0.03	-0.06	0.04
Pattern-Mix 2	0.05	0.15	0.05	0.15	0.06	0.19	0.16	0.24
Pattern-Mix 3	0.04	0.11	0.04	0.10	0.05	0.12	0.05	0.11
Principal Strat	-0.01	0.03	-0.04	0.00	-0.04	0.02	-0.02	0.06
Terminal Dec	0.00	0.12	0.10	0.17	0.08	0.17	0.18	0.22
Partly Cond	-0.01	0.07	-0.00	0.06	-0.02	0.06	-0.01	0.05
PAH	-0.04	-0.04	-0.04	-0.04	-0.06	-0.06	-0.06	-0.05
Systolic BP								
Unconditional	-0.07	-0.11	0.06	-0.09	-0.30	-0.24	-0.17	-0.22
Pattern-Mix 1	-1.24	-0.94	-1.12	-0.92	-1.40	-1.66	-1.28	-2.97
Pattern-Mix 2	-0.49	-0.58	-0.24	-0.55	-0.52	-0.31	-0.34	-0.29
Pattern-Mix 3	0.06	0.01	0.21	0.06	-0.26	-0.23	-0.11	-0.18
Principal Strat	0.51	0.04	0.33	0.12	0.16	-0.06	0.07	-0.20
Terminal Dec	-0.49	-0.53	-0.31	-0.58	-0.43	-0.38	-0.26	-0.23
Partly Cond	-0.11	-0.12	-0.03	-0.12	-0.22	-0.19	-0.14	-0.19
PAH	-0.02	-0.01	-0.02	-0.02	-0.03	-0.03	-0.03	-0.03
Diastolic BP								
Unconditional	-0.79	-0.80	-0.76	-0.80	-0.85	-0.81	-0.82	-0.81
Pattern-Mix 1	-1.49	-1.27	-0.48	-0.40	-1.37	-1.24	-0.36	-0.37
Pattern-Mix 2	-0.88	-0.85	-0.90	-0.91	-1.08	-1.01	-1.10	-1.08
Pattern-Mix 3	-0.74	-0.77	-0.73	-0.78	-0.85	-0.82	-0.84	-0.83
Principal Strat	-0.72	-0.84	-0.83	-0.77	-0.77	-0.85	-0.83	-0.84
Terminal Dec	-0.49	-0.46	-0.51	-0.49	-0.67	-0.61	-0.70	-0.64
Partly Cond	-0.79	-0.79	-0.78	-0.80	-0.78	-0.77	-0.77	-0.78
ADL								
Unconditional	0.23	0.23	0.21	0.21	0.16	0.17	0.14	0.15
Pattern-Mix 1	0.47	0.46	0.50	0.47	0.35	0.39	0.38	0.40
Pattern-Mix 2	0.42	0.38	0.42	0.40	0.30	0.29	0.30	0.32
Pattern-Mix 3	0.19	0.21	0.16	0.18	0.12	0.14	0.08	0.11
Principal Strat	0.13	0.17	0.07	0.14	0.09	0.15	0.01	0.10
Terminal Dec	0.33	0.31	0.34	0.33	0.23	0.24	0.24	0.26
Partly Cond	0.17	0.18	0.14	0.16	0.10	0.12	0.08	0.05
PAH	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06

Table 3.4: CES-D - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1,000 simulated samples with three follow-up times for MAR and NMAR

		Sample size= 100		**Sample size**= 500		**Sample size**= 1000	
Model	**Parameters**	MAR	NMAR	MAR	NMAR	MAR	NMAR
Unconditional	White Females	-0.56	-1.54	0.28	-1.09	0.11	-1.29
		(0.0069)	(0.0068)	(0.0029)	(0.0030)	(0.0022)	(0.0021)
	White Males	-1.15	-2.16	0.38	-1.50	0.49	-1.05
		(0.0100)	(0.0097)	(0.0445)	(0.0046)	(0.0031	(0.0029)
	Black Females	0.04	-0.143	0.22	-0.69	-0.39	-1.21
		(0.0055)	(0.0056)	(0.0024)	(0.0024)	(0.0018)	(0.0017)
	Black Males	-0.28	-1.87	0.39	0.44	-0.06	-0.89
		(0.0083)	(0.0084)	(0.0041)	(0.0039)	(0.0027)	(0.0026)
Pattern	White Females			-0.82	-1.60	-0.79	-1.72
Mixture				(0.0032)	(0.0032)	(0.0024)	(0.0023)
Completers	White Males			-1.17	-1.91	-0.55	-1.85
				(0.0049)	(0.0049)	(0.0034)	(0.0031)
	Black Females			-0.53	-1.13	-1.16	-1.59
				(0.0026)	(0.0026)	(0.0019)	(0.0018)
	Black Males			-0.70	-1.22	-0.93	-1.58
				(0.0044)	(0.0042)	(0.0029)	(0.0028)
Partly	White Females	-1.22	-1.85	-0.53	-1.55	-0.66	-1.69
Conditional		(0.0071)	(0.0070)	(0.0031)	(0.0031)	(0.0023)	(0.0022)
	White Males	-1.75	-2.79	-0.72	-1.83	-0.37	-1.74
		(0.0101)	(0.0100)	(0.0047)	(0.0047)	(0.0033)	(0.0031)
	Black Females	-0.65	-1.63	-0.4	-1.08	-1.05	-1.63
		(0.0058)	(0.0059)	(0.0025)	(0.0025)	(0.0019)	(0.0018)
	Black Males	-0.92	-2.33	-0.46	-1.14	-0.79	-1.56
		(0.0087)	(0.0087)	(0.0042)	(0.0041)	(0.0028)	(0.0027)

Table 3.5: Systolic BP - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1,000 simulated samples with three follow-up times for MAR and NMAR

		Sample size= 100		**Sample size**= 500		**Sample size**= 1000	
Model	**Parameters**	MAR	NMAR	MAR	NMAR	MAR	NMAR
Unconditional	White Females	-0.14	-0.49	0.00	-0.40	-0.02	-0.44
		(0.0006)	(0.0006)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
	White Males	-0.10	-0.52	-0.01	-0.40	0.02	-0.41
		(0.0007)	(0.0008)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
	Black Females	-0.01	-0.35	0.03	-0.39	-0.03	-0.43
		(0.0006)	(0.0006)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
	Black Males	0.02	-0.38	0.02	-0.40	0.01	-0.39
		(0.0007)	(0.0007)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Pattern	White Females			-0.47	-0.55	-0.47	-0.61
Mixture				(0.0003)	(0.0003)	(0.0002)	(0.0002)
Completers	White Males			-0.51	-0.55	-0.48	-0.58
				(0.0004)	(0.0004)	(0.0002)	(0.0003)
	Black Females			-0.44	-0.55	-0.50	-0.60
				(0.0003)	(0.0003)	(0.0002)	(0.0002)
	Black Males			-0.47	-0.55	-0.51	-0.57
				(0.0004)	(0.0004)	(0.0002)	(0.0002)
Partly	White Females	-0.38	-0.59	-0.25	-0.48	-0.25	-0.53
Conditional		(0.0006)	(0.0006)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
	White Males	-0.36	-0.61	-0.28	-0.49	-0.24	-0.50
		(0.0008)	(0.0008)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
	Black Females	-0.24	-0.44	-0.21	-0.47	-0.27	-0.52
		(0.0006)	(0.0006)	(0.0002)	(0.0003)	(0.0002)	(0.0002)
	Black Males	-0.22	-0.46	-0.23	-0.48	-0.26	-0.49
		(0.0007)	(0.0007)	(0.0003)	(0.0003)	(0.0002)	(0.0002)

Table 3.6: Diastolic BP - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1,000 simulated samples with three follow-up times for MAR and NMAR

		Sample size= 100		**Sample size**= 500		**Sample size**= 1000	
Model	**Parameters**	MAR	NMAR	MAR	NMAR	MAR	NMAR
Unconditional	White Females	-0.05	-0.24	0.03	-0.20	-0.01	-0.23
		(0.0006)	(0.0006)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
	White Males	-0.04	-0.21	-0.01	-0.23	0.01	-0.21
		(0.0007)	(0.0007)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
	Black Females	0.02	-0.22	0.04	-0.18	-0.03	-0.22
		(0.0005)	(0.0005)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
	Black Males	0.03	-0.19	0.00	-0.21	0.00	-0.20
		(0.0006)	(0.0006)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Pattern	White Females			-0.18	-0.29	-0.24	-0.32
Mixture				(0.0003)	(0.0003)	(0.0002)	(0.0002)
Completers	White Males			-0.23	-0.32	-0.22	-0.29
				(0.0003)	(0.0003)	(0.0002)	(0.0002)
	Black Females			-0.20	-0.26	-0.26	-0.30
				(0.0002)	(0.0002)	(0.0002)	(0.0002)
	Black Males			-0.24	-0.29	-0.25	-0.27
				(0.0003)	(0.0003)	(0.0002)	(0.0002)
Partly	White Females	-0.13	-0.28	-0.07	-0.24	-0.12	-0.27
Conditional		(0.0006)	(0.0006)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
	White Males	-0.12	-0.24	-0.11	-0.27	-.0.10	-0.25
		(0.0007)	(0.0007)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
	Black Females	-0.10	-0.26	-0.07	-0.22	-0.14	-0.26
		(0.0005)	(0.0005)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
	Black Males	-0.01	-0.21	-0.11	-0.24	-0.12	-0.24
		(0.0006)	(0.0006)	(0.0003)	(0.0003)	(0.0002)	(0.0002)

Table 3.7: ADL - Relative Bias (×100) and (SE) of mean estimates five years post-baseline based on 1,000 simulated samples with three follow-up times for MAR and NMAR

		Sample size= 100		**Sample size**= 500		**Sample size**= 1000	
Model	**Parameters**	MAR	NMAR	MAR	NMAR	MAR	NMAR
Unconditional	White Females	-2.27	-2.29	-0.61	-1.4	-1.24	-1.58
		(0.0059)	(0.0059)	(0.0025)	(0.0025)	(0.0019)	(0.0018)
	White Males	1.13	0.23	2.00	1.36	2.23	1.61
		(0.0079)	(0.0076)	(0.0034)	(0.0034)	(0.0024)	(0.0024)
	Black Females	0.98	0.22	0.93	0.59	0.27	0.14
		(0.0047)	(0.0047)	(0.0021)	(0.0020)	(0.0014)	(0.0014)
	Black Males	0.50	2.64	3.34	3.19	3.44	3.08
		(0.0063)	(0.0064)	(0.0029)	(0.0029)	(0.0020)	(0.0020)
Pattern	White Females			-1.02	-1.67	-1.64	-1.88
Mixture				(0.0028)	(0.0027)	(0.0020)	(0.0020)
Completers	White Males			1.51	1.10	1.66	1.34
				(0.0037)	(0.0038)	(0.0026)	(0.0036)
	Black Females			0.50	0.30	-0.09	-0.11
				(0.0022)	(0.0023)	(0.0016)	(0.0016)
	Black Males			2.85	2.89	2.92	2.85
				(0.0032)	-0.0031	(0.0021)	(0.0021)
Partly	White Females	-2.46	-2.87	-0.84	-1.65	-0.155	-1.79
Conditional		(0.0062)	(0.0062)	(0.0026)	(0.0026)	(0.0020)	(0.0019)
	White Males	0.83	0.04	1.72	1.15	1.87	1.35
		(0.0081)	(0.0080)	(0.0035)	(0.0036)	(0.0025)	(0.0025)
	Black Females	0.12	-0.06	0.62	0.43	0.01	-0.02
		(0.0049)	(0.0049)	(0.0021)	(0.0021)	(0.0015)	(0.0015)
	Black Males	3.23	2.74	2.99	3.06	3.13	2.87
		(0.0067)	(0.0067)	(0.0030)	(0.0030)	(0.0021)	(0.0020)

CHAPTER 4: EFFICIENCY OF MODELS USED TO ANALYZE LONGITUDINAL OUTCOMES TRUNCATED DUE TO DEATH

4.1 Introduction

In public health and medical research, the interest is usually in estimating the fixed effects and in making inference about the parameters (Gurka et al., 2011). These estimates provide the average change in the response for the population. The nature of the missing data can potentially impact or bias the inference of the estimates of interest (Crouchley and Ganjali, 2002). The focus for this chapter is limited to efficiency concerning the marginal models of the unconditional, pattern-mixture, and partly conditional models.

As a reminder, the general linear mixed model for the combined N subjects is given as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{d} + \boldsymbol{e}, \tag{4.1}$$

with,

$\boldsymbol{y}$ is the $\sum_{i=1}^{N} p_i \times 1$ stacked vector of the response vectors, y_i for all i

$\boldsymbol{X}$ is the $\sum_{i=1}^{N} p_i \times q$ stacked matrix of known design matrices, $\boldsymbol{X}_i$ for each subject i

$\boldsymbol{\beta}$ is the $q \times 1$ vector of unknown population parameters

$\boldsymbol{Z}$ is the $\sum_{i=1}^{N} p_i \times mN$ block-diagonal matrix with the $p_i \times m$ design matrix, $\boldsymbol{Z}_i$ of the $m \times 1$ random effects, $\boldsymbol{d}_i$, for each subject i on the main diagonal

$\boldsymbol{d}$ is the $mN \times 1$ stacked vector of subject-specific unknown parameters, $\boldsymbol{d}_i$ for each subject i

$\boldsymbol{\Delta}$ is the $mN \times mN$ block-diagonal covariance matrix with the $(m \times m)$ covariance matrix, $\boldsymbol{D}$ of random effects, $\boldsymbol{d}_i$, on the main diagonal

$\boldsymbol{e}$ is the $\sum_{i=1}^{N} p_i \times 1$ stacked vector of residual errors for each subject i

$\boldsymbol{\Sigma}_e$ is the $\sum_{i=1}^{N} p_i \times \sum_{i=1}^{N} p_i$ block-diagonal covariance matrix with the covariances, $\boldsymbol{\Sigma}_{e_i}$ for the random errors, $\boldsymbol{e}_i$ on the main diagonal

and $Var(\boldsymbol{y}) = \boldsymbol{V}$ which is a block-diagonal matrix with blocks of $\boldsymbol{V}_i = \boldsymbol{Z}_i \boldsymbol{D} \boldsymbol{Z}_i' + \boldsymbol{\Sigma}_{e_i}$ on the main diagonal and zeroes elsewhere.

Marginally, $\boldsymbol{y} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \boldsymbol{V})$. When the fixed effects are of primary interest, the maximum likelihood (ML) estimate for $\boldsymbol{\beta}$ conditional on the estimates of the variance components of $\boldsymbol{V}$ is given as $\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}) = (\boldsymbol{X}^T \hat{\boldsymbol{V}}^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{X})^{-1} \boldsymbol{X}^T \hat{\boldsymbol{V}}^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{y} = (\sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{V}}_i^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{X}_i)^{-1} \sum_{i=1}^{N} \boldsymbol{X}_i^T \hat{\boldsymbol{V}}_i^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{y}_i$, where $\hat{\boldsymbol{\theta}}$ is the estimate of the covariance components. The ML estimate of the parameters is normally distributed with mean $\boldsymbol{\beta}$ and covariance, $var(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})) = (\boldsymbol{X}^T \boldsymbol{V}^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{X})^{-1}$. Under maximum likelihood estimation, the $var(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}))$ is underestimated because the estimation neglects the variability induced from the estimation of $\hat{\boldsymbol{\theta}}$. For this reason, generally, residual (or restricted) maximum likelihood (REML) estimation, which was introduced by

Patterson and Thompson (1971), is preferable to estimate the covariance components. Verbeke and Molenberghs (2000) describes REML's estimation of the variance components as the process of maximizing the likelihood function of a collection of residual contrasts, $\boldsymbol{U} = \boldsymbol{A}^T\boldsymbol{y}$, where $\boldsymbol{A}$ is an $(\sum_{i=1}^{N} p_i \times (\sum_{i=1}^{N} p_i - q))$ full-column rank matrix with columns orthogonal of the design matrix $\boldsymbol{X}$. The error contrast, $\boldsymbol{U}$, follows a normal distribution with $E(\boldsymbol{U}) = \boldsymbol{0}$ and $var(\boldsymbol{U}) = \boldsymbol{A}^T\boldsymbol{V}_i\boldsymbol{A}$. REML estimates of $\boldsymbol{\theta}$ are computed by optimizing the REML log-likelihood function Harville (1977), which takes the form

$$\begin{aligned} l_{REML}(\boldsymbol{\theta}) = &-\frac{1}{2}\left[(\sum_{i=1}^{N}(p_i) - q)\right]\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{N}\ln(|\boldsymbol{V}_i|) \\ &- \frac{1}{2}\sum i = 1^N(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}})^T\boldsymbol{V}_i(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}) \\ &- \frac{1}{2}\sum_{i=1}^{N}\ln(|\boldsymbol{X}_i^T\boldsymbol{V}_i^{-1}\boldsymbol{X}_i|) \end{aligned} \tag{4.2}$$

where $\hat{\boldsymbol{\beta}}$ is the expression given previously. By accounting for the loss of degrees of freedom from the estimation of the fixed effects, $\boldsymbol{\beta}$, REML estimation produces unbiased estimates for the covariance parameter components.

In the preceding chapters, the computational E-M algorithm was discussed. For incomplete data, the E-M algorithm's expectation step conceptually creates a complete dataset by assuming the data is balanced and that the dependent variable is complete. One of the most problematic issues of the E-M algorithm is that it is computationally heavy and slow to reach convergence. This estimation process can also overpromise because likelihood is maximized over the complete data rather than the observed data.

An alternative numerical computational method is the Newton-Raphson algorithm (N-R). The N-R algorithm is the most commonly used algorithm for either maximum likelihood or restricted maximum likelihood estimation for the linear mixed model. This numerical optimization procedure minimizes (-2) times the ML profile log-likelihood

function for ML estimation given below and REML log-likelihood functions for REML estimation, given in (4.2).

$$l_{ML}(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i=1}^{N}(p_i)\ln(2\pi) - \frac{1}{2}\sum_{i=1}^{N}\ln(|\boldsymbol{V}_i) - 12\sum_{i=1}(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}})^T\boldsymbol{V}_i(\boldsymbol{y}_i - \boldsymbol{X}_i\hat{\boldsymbol{\beta}}) \quad (4.3)$$

where $\hat{\boldsymbol{\beta}}$ is the expression given previously. For each iteration, the N-R algorithm requires computing the vector of partial derivatives and second derivative matrix with respect to the covariance parameters, which is given as

$$\begin{bmatrix} \tilde{\boldsymbol{\beta}} \\ \tilde{\boldsymbol{\theta}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta}^{\circ} \\ \boldsymbol{\theta}^{\circ} \end{bmatrix} - \begin{bmatrix} \boldsymbol{H}_{\beta\beta} & \boldsymbol{H}_{\beta\theta} \\ \boldsymbol{H}_{\theta\beta} & \boldsymbol{H}_{\theta\theta} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{s}_{\beta} \\ \boldsymbol{s}_{\theta} \end{bmatrix} \quad (4.4)$$

with

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_{\beta\beta} & \boldsymbol{H}_{\beta\theta} \\ \boldsymbol{H}_{\theta\beta} & \boldsymbol{H}_{\theta\theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 l}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}} & \frac{\partial^2 l}{\partial\boldsymbol{\beta}\partial\boldsymbol{\theta}} \\ \frac{\partial^2 l}{\partial\boldsymbol{\theta}\partial\boldsymbol{\beta}} & \frac{\partial^2 l}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}} \end{bmatrix} \quad (4.5)$$

and

$$\boldsymbol{s} = \begin{bmatrix} \boldsymbol{s}_{\beta} \\ \boldsymbol{s}_{\theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial l}{\partial\boldsymbol{\beta}} \\ \frac{\partial l}{\partial\boldsymbol{\theta}} \end{bmatrix}. \quad (4.6)$$

$\boldsymbol{H}$ is referred to as the Hessian matrix and $\boldsymbol{s}$ is often described as the gradient or score vector. During the computation algorithm these values are evaluated using the current values of the parameters. Details for this numerical method have been published by Jennrich and Schluchter (1986).

For the partly conditional model (Kurland et al., 2009; Kurland and Heagerty, 2005), we estimate the fixed effects using independent estimating equations. Under this approach, an estimator $\hat{\boldsymbol{\beta}}_I^A$ of $\boldsymbol{\beta}^A$, where $\boldsymbol{\beta}^A$ is the direct parameterization of those who are alive at the time of measurement, is the solution of the score equations of the

form:

$$U(\beta^A) = \sum_{i=1}^{N} \sum_{j=1}^{p_i} A_{ij} \frac{\partial \mu_{ij}^A}{\partial \beta} (Y_{ij} - \mu_{ij}^A) = 0, \tag{4.7}$$

where A_{ij} is an indicator variable that equals 1 if the individual i survives beyond the current survey wave and 0 otherwise. Variance of $\hat{\boldsymbol{\beta}}_I^A$ can be consistently estimated by the sandwich estimator corresponding to the independent estimating equations:

$$\left[\boldsymbol{X}^T \hat{\boldsymbol{W}} \boldsymbol{X}\right]^{-1} \left[\sum_{i=1}^{N} \boldsymbol{X}_i^T (\boldsymbol{y}_i - \hat{\boldsymbol{\mu}}_i^A)(\boldsymbol{y}_i - \hat{\boldsymbol{\mu}}_i^A)^T \boldsymbol{X}_i\right] \left[\boldsymbol{X}^T \hat{\boldsymbol{W}} \boldsymbol{X}\right]^{-1} \tag{4.8}$$

where $\boldsymbol{X}$ is a matrix of stacked $\boldsymbol{X}_i$'s and $\hat{\boldsymbol{W}}$ is a diagonal matrix of final weights if any exist. The sandwich estimator was first proposed by Huber (1967) and White (2007) and later applied to longitudinal data by Liang and Zeger (1986). In large samples, this estimator provides an appropriate estimator of $var(\hat{\boldsymbol{\beta}}_I^A)$ regardless of the true variance structure of $\boldsymbol{y}_i$. Zeger et al. (1988) asserted that the sandwich estimator is highly efficient when the within-subject correlation is weak.

Given the current estimates of the nuisance parameters, $\boldsymbol{\phi}$ and $\boldsymbol{\alpha}$, the estimate $\hat{\boldsymbol{\beta}^{\boldsymbol{A}}}$ can be computed by the following iterative procedure:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{j+1}^A = \hat{\boldsymbol{\beta}}_j^A - & \left\{\sum_{i=1}^{N} \frac{\partial \boldsymbol{\mu}_i^{AT}(\hat{\boldsymbol{\beta}}_j)}{\partial \boldsymbol{\beta}^A} \tilde{\boldsymbol{V}}_i^{-1}(\hat{\boldsymbol{\beta}}_j) \frac{\partial \boldsymbol{\mu}_i^A(\hat{\boldsymbol{\beta}}_j^A)}{\partial \boldsymbol{\beta}^A}\right\}^{-1} \\ & \times \left\{\sum_{i=1}^{N} \frac{\partial \boldsymbol{\mu}_i^{AT}(\hat{\boldsymbol{\beta}}_j)}{\partial \boldsymbol{\beta}^A} \tilde{\boldsymbol{V}}_i^{-1}(\hat{\boldsymbol{\beta}}_j)(\boldsymbol{y}_i - \boldsymbol{\mu}_i^A(\hat{\boldsymbol{\beta}}_j))\right\} \end{aligned} \tag{4.9}$$

where $\tilde{\boldsymbol{V}}_i^{-1}(\hat{\boldsymbol{\beta}}_j)$ is the variance of $\boldsymbol{y}_i$ with the estimates of the nuisance parameters. This procedure is a modification of Fisher's scoring method. The computing process for $\hat{\boldsymbol{\beta}}$ oscillates between a modified Fisher scoring for $\boldsymbol{\beta}$ and the moment estimation of the nuisance parameters.

4.1.1 Inference of the Fixed Effects

Making generalizations about the average change of any outcome in the populations requires testing of the fixed effects or a linear combination of the fixed effects. In the linear mixed model, the inference about $\boldsymbol{\beta}$ is established by constructing Wald-like tests using the estimated standard errors. For any known matrix $\boldsymbol{L}$ a test of the hypothesis

$$\mathrm{H}_0 : \boldsymbol{L}\boldsymbol{\beta} = \mathbf{0} \text{ versus } \mathrm{H}_A : \boldsymbol{L}\boldsymbol{\beta} \neq \mathbf{0} \tag{4.10}$$

is conducted from the fact that

$$(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T \boldsymbol{L} \left[\boldsymbol{L} \left(\sum_{i=1}^{N} \boldsymbol{X}_i^T \boldsymbol{V}_i^{-1}(\hat{\boldsymbol{\theta}}) \boldsymbol{X} \right)^{-1} \boldsymbol{L}^T \right]^{-1} \boldsymbol{L}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \tag{4.11}$$

asymptotically follows a chi-squared distribution with rank($\boldsymbol{L}$).

Because of the variability introduced from the estimation of the covariance parameters, the chi-squared reference distribution is replaced by an approximate F-distribution. A scaled-Wald statistic was introduced by Kenward and Roger (1997) that adjusts the covariance estimate to account for the additional introduced variability. Interests for most gerontologists and public health scientists lie in the population mean–that is, the fixed effects–and the specification of the covariance structure will impact the results. Gurka et al. (2011) postulated that inference for the fixed effects is not robust to the misspecification of the covariance in the linear mixed model. Furthermore, Gurka (2006) demonstrated that no one method will reliably identify the best covariance model, especially for small samples.

The sandwich estimator developed by Liang and Zeger (1986) for the $var(\hat{\boldsymbol{\beta}})$ has been shown to be robust to covariance model misspecification, but is not as efficient as the true covariance model (Gurka et al., 2011). For the linear mixed model with

assumed Gaussian errors, the sandwich estimator, also referred to as the robust or empirical variance estimator, is specified as

$$var(\hat{\boldsymbol{\beta}}) = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}(\boldsymbol{X}^T\boldsymbol{V}^{-1}(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})(\boldsymbol{y}-\boldsymbol{X}\hat{\boldsymbol{\beta}})^T\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1} \quad (4.12)$$

In their study of the sensitivity of inference for the fixed effects in linear mixed models to misspecification of the error distribution, Jacqmin-Gadda et al. (1980) found that inference was not compromised when Gaussian errors were assumed but the true distribution was either non-Gaussian or heteroscedastic. Additionally, they concluded that the mixed model with random intercept and slope is more robust to misspecification of the covariance structure than the compound symmetrical model with a random intercept only.

Thus far, we have examined the bias in a subset of the models proposed to incorporate survival in the estimates of non-mortality outcomes. Bias in the regression models is represented by the difference of the true value of the regression parameter $\boldsymbol{\beta}$ and the expected value of its estimate. That is,

$$\text{Bias } = B\left(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\right) = \boldsymbol{\beta}(\hat{\boldsymbol{\theta}}) - E\left(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\right).$$

Bias of the regression parameters indicates if the estimator is under- or over-estimating the value of the parameter.

Efficiency is based on the mean squared error (MSE) of the estimator. The mean squared error is a characteristic of the estimator that combines the variance of the estimator and its bias. For the ML unbiased estimator for $\boldsymbol{\beta}$ shown previously, the MSE of $\boldsymbol{\beta}$ is given as

$$\text{MSE}\left[\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\right] = var\left(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\right) + B\left(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\right) = var\left(\hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}})\right).$$

The variance of an unbiased estimator for a parameter, say $\boldsymbol{\theta}$ is bounded below by the Fisher information matrix as stated by the Cramér-Rao inequality: $var(\boldsymbol{\theta}) \geq \mathcal{I}^{-1}(\boldsymbol{\theta})$, where $\mathcal{I}(\boldsymbol{\theta})$ is the Fisher information. The Fisher information is the second moment of the partial derivative with respect to $\boldsymbol{\theta}$ of the log-likelihood, such that

$$\mathcal{I}(\boldsymbol{\theta}) = E\left[\left(\frac{\partial l}{\partial \boldsymbol{\theta}}\right)^2 |\boldsymbol{\theta}\right]. \tag{4.13}$$

Efficiency of an estimator is a measurement of the optimality of an estimator. Because estimators with small variance are more precise, a more efficient estimator essentially requires fewer samples than an inefficient estimator.

4.1.2 Efficiency in Models with Truncated Outcomes due to Death

To examine the efficiency of estimators for parameters of the unconditional, pattern-mixture, and the partly conditional models, we reexamined the previous simulations. Each of the four outcomes were treated as continuous outcomes and were generated from the mixed model with a random intercept and slope as described in equation (2.1). Four waves of longitudinal outcomes were simulated from a normal distribution with mean $\boldsymbol{X_i\beta}$ and covariance $\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i\boldsymbol{D}_i\boldsymbol{Z}_i^T + \sigma^2\boldsymbol{I}_i$, where $\boldsymbol{D}$ was allowed to be unstructured. For each outcome, the design matrix $\boldsymbol{X}_i$ consisted of a column vector of ones for the intercept, a column vector for time of measurements post-baseline (0, 3, 6, 10 years), a column vector of baseline age centered about the mean, a column vector of indicators for identifying as black, a column vector of indicators for identifying as male, a column vector indicating time by race, and a column vector indicating time by sex. The design matrix of the random effects, $(\boldsymbol{Z}_i)$, was constructed as a column vector of ones for the intercept and a column vector for the time of measurements post-baseline (0, 3, 6, 10 years). The race-gender combinations values were treated as multinomial random variables and were generated accordingly: white males ($\pi = 0.16$), black males

($\pi = 0.19$), and white females ($\pi = 0.29$). Age was simulated assuming it was from a normal distribution dictated by the mean and standard deviation in each race-gender group from the NC EPESE. Similarly, time of measurements mirrored the NC EPESE, and we therefore assume measurements were only possible at baseline, and 3, 6, and 10 years post-baseline. The values of the parameters $\boldsymbol{\beta}, \sigma^2$, and $\boldsymbol{D}$ used in the simulation are given for each outcome below.

CES-D score: $\boldsymbol{\beta}^T = \begin{pmatrix} 3.243 & 0.086 & 0.065 & 0.356 & -0.686 & 0.001 & 0.003 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 5.32 & -0.14 \\ -0.14 & 0.39 \end{pmatrix} \boldsymbol{Z}_i^T + 7.16\boldsymbol{I}_i$$

Systolic BP: $\boldsymbol{\beta}^T = \begin{pmatrix} 143.29 & 0.088 & 0.002 & 0.756 & -2.326 & -0.017 & -0.136 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 139.69 & -3.38 \\ -3.38 & 0.91 \end{pmatrix} \boldsymbol{Z}_i^T + 314.05\boldsymbol{I}_i$$

Diastolic BP: $\boldsymbol{\beta}^T = \begin{pmatrix} 77.258 & 0.799 & -0.289 & 2.775 & 1.529 - 0.0006 - 0.010 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z}_i \begin{pmatrix} 47.66 & -2.26 \\ -2.26 & 0.39 \end{pmatrix} \boldsymbol{Z}_i^T + 92.94 \boldsymbol{I}_i$$

ADL score: $\boldsymbol{\beta}^T = \begin{pmatrix} 0.742 & 0.211 & 0.114 & 0.243 & 0.133 & 0.001 & -0.063 \end{pmatrix}$

$$\boldsymbol{\Sigma_i} = \boldsymbol{Z} \begin{pmatrix} 1.52 & -0.002 \\ -0.002 & 0.04 \end{pmatrix} \boldsymbol{Z}^T + 1.78 \boldsymbol{I}_i$$

A 1,000 samples of complete data were generated for each sample size– $N = 100$, $N = 500$, and $N = 1000$. After the complete datasets were generated, missing indicators were created to represent participants leaving the study because of death or non-participation. Subjects in the simulated datasets became at risk of death or dropout following baseline responses.

In each dataset, a death indicator was created from a Bernoulli random generating function to simulate participants leaving the study because of death at a rate of 10% per survey wave following baseline for an overall death rate of approximately 27%. For convenience, an ordinal ranking for the missing categories, such as death ($k = 1$), non-response ($k = 2$), and completers was assumed. Following this assumption, a cumulative regression model was used to simulate missing categories with known probabilities for each wave for each outcome with dependence on the baseline covariates

and previous observed response values. Missing at random due to death and non-response was modeled using the cumulative regression model described below:

$$logit(\gamma_{ik}) = logit[P(y_i \leq k)] = \eta_k + \boldsymbol{x}^T\boldsymbol{\beta} \tag{4.14}$$

with $\boldsymbol{x}_i$ representing a vector of the covariates – sex (male=1, 0 otherwise), race (black=1, 0 otherwise), baseline age centered about baseline mean, lag measurement occasion, $t_{i(j-1)}$ (lag wave), the additional effects on the probability over the lag measurement occasion given one identifies as male (male-lag wave interaction) and given one identifies as black (black-lag wave interaction) and the previous (lag) response outcome value ($y_{i(j-1)}$) – and $\boldsymbol{\beta}$ representing their corresponding regression parameters: $\boldsymbol{\beta}^T = \begin{pmatrix} 0.1 & 0.05 & 0.001 & 0.001 & -0.02 & -0.02 & .01 \end{pmatrix}$. η_k represents the intercept for the k cumulative logit. The η_k by outcome is given as

CES-D Score: $\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -1.6 \\ -1.2 \end{pmatrix}$ Systolic BP: $\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -3.0 \\ -2.6 \end{pmatrix}$

Diastolic BP: $\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -2.4 \\ -2.0 \end{pmatrix}$ ADL: $\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} -1.6 \\ -1.2 \end{pmatrix}$

Using the estimates of the probabilities as a parameter in a random Bernoulli generator, these models led to 31% missing due to death and 13% missing due to non-response for each outcome. The linear mixed model was fitted by evoking the REML estimation and the partly conditional model was fitted as described previously. All analyses were performed in SAS v9.2.

4.2 Results

The outcome with the most efficient fixed-effects estimator proved to be the ADL (Figure 4.4) for each of the models–unconditional, pattern-mixture, and partly conditional models. The efficiency of the estimation of the effect of the centered baseline age and the effect of the time of measurement was similar for each outcome and model by sample size. As expected, the increase in sample size from 500 to 1,000 increased the efficiency of the estimation for all three models for every outcome.

4.3 Discussion

Gurka et al. (2011) performed a simulation that showed that if the true variance-covariance structure of a linear mixed model is compound symmetric, but a structure allowing the random effects to be correlated with different variance and the within error to have homogenous variance is modeled, then the standard errors can underestimated. The distribution used to generate the ADL values were nearly from a normal distribution with a compound symmetric variance-covariance structure. This could be one explanation of the ADL outcome behaving much differently than the other dependent variables. The graphs in Figures 4.1-4.4 support that the models considered are less efficient in estimation categorical parameters. Similarly, each model for each outcome has difficulty estimating intercepts. Unconditional models and partly conditional models were the most efficient.

Figure 4.1: Parameter's Root Mean Squared Errors for Depression by Missing Pattern and Sample Size

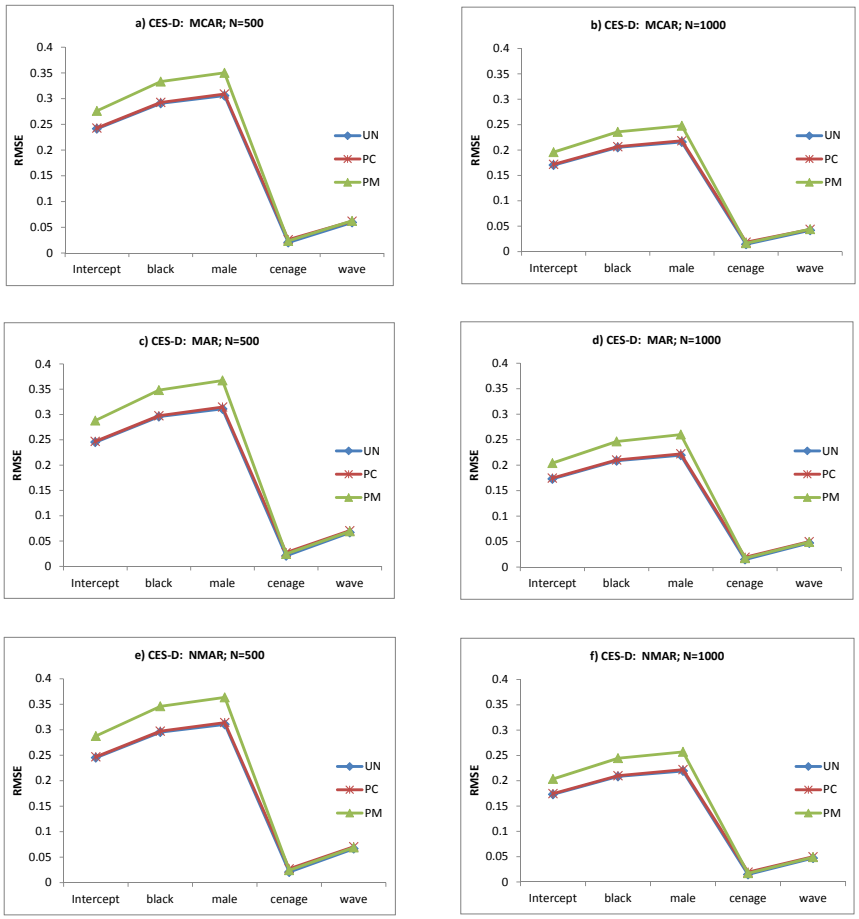


Figure 4.2: Parameter's Root Mean Squared Errors for Systolic BP by Missing Pattern and Sample Size

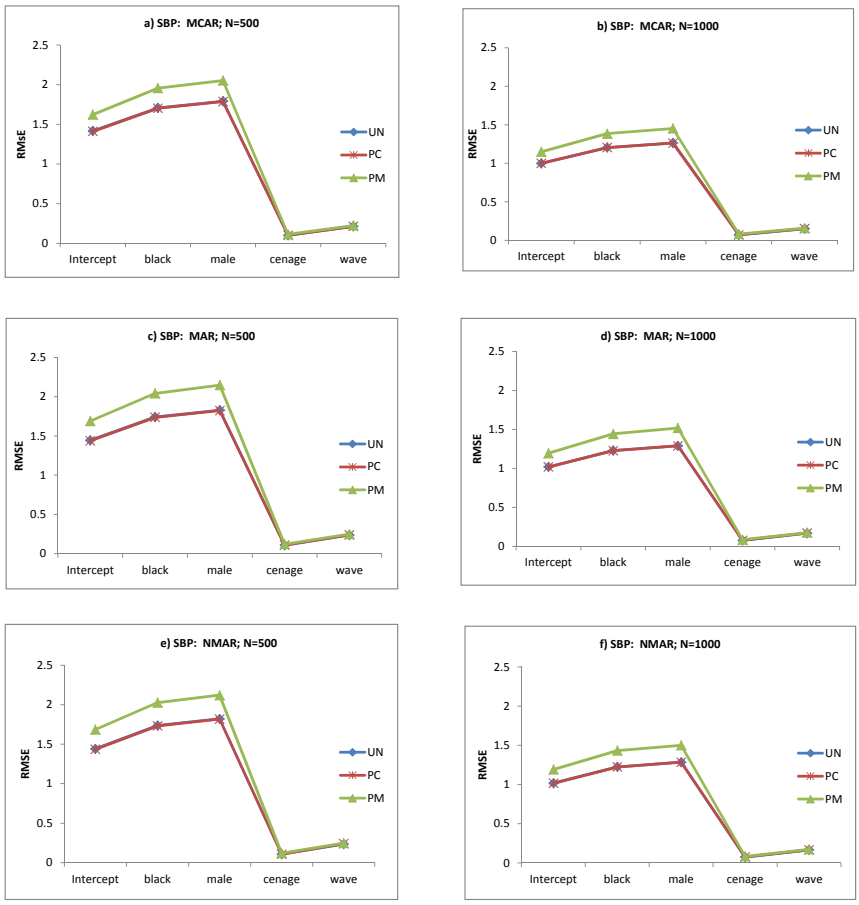


Figure 4.3: Parameter's Root Mean Squared Errors for Diastolic BP by Missing Pattern and Sample Size

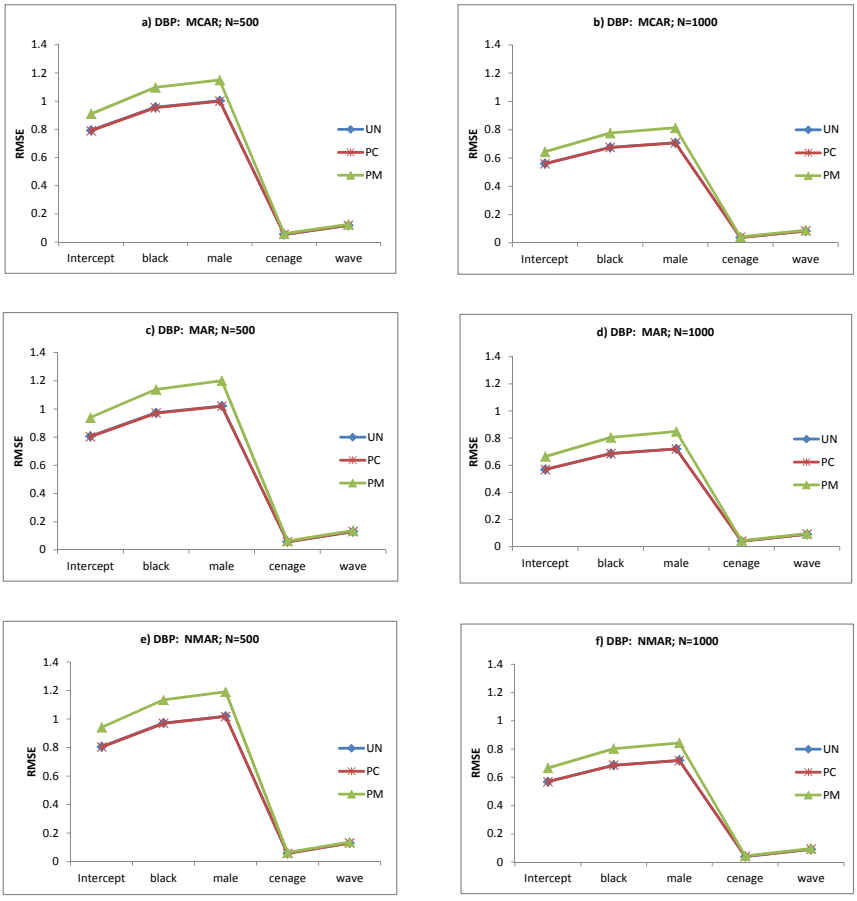
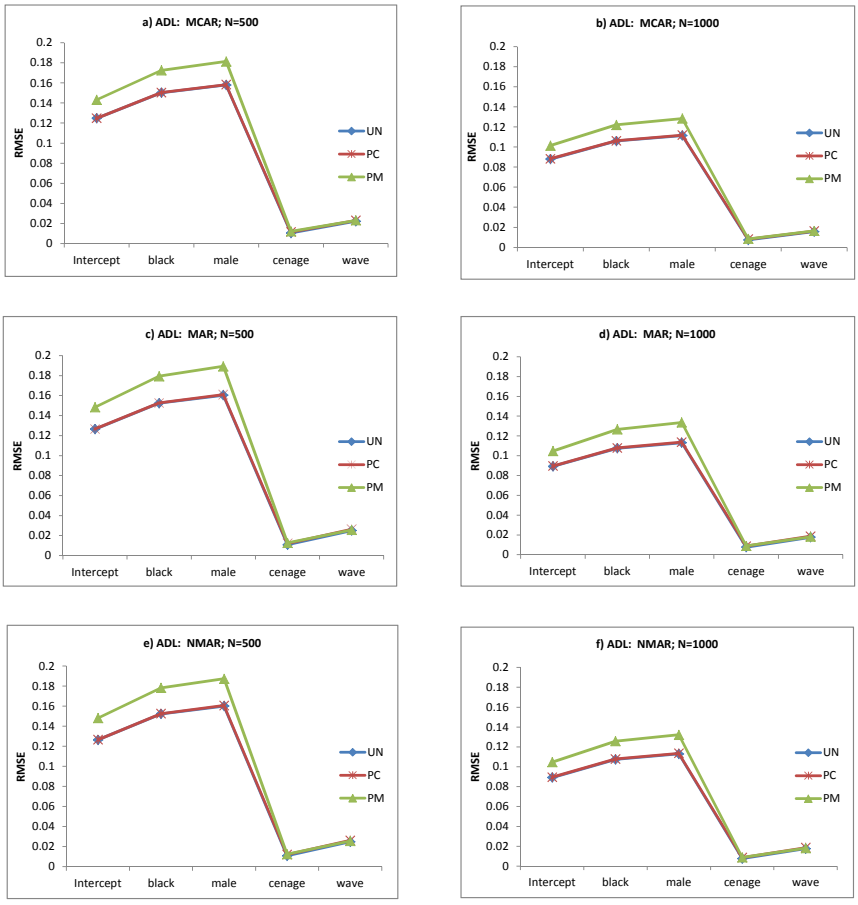


Figure 4.4: Parameter's Root Mean Squared Errors for ADL by Missing Pattern and Sample Size



CHAPTER 5: CONCLUSIONS AND FURTHER RESEARCH

5.1 Summary

This dissertation has contributed to the discussion of the effects of mortality on rates of change and variability in unconditional, pattern-mixture, and partly conditional models to analyze longitudinal outcomes truncated by deaths. By applying the proposed models to an established dataset we were able to broadly compare previous conclusions and the conclusions resulting from the proposed models. Next, we compared the performance of the models in the presence of death and non-participation to results when response truncation was only due to death. Further, using data generated from a theoretical distribution, we evaluated the proposed models for fitting longitudinal outcomes truncated by deaths on their ability to avoid bias in the parameters and in the variance of the parameters when analyzing datasets with different missing data burdens.

Our first objective was to apply the proposed methods to an established dataset with death data and allowing missing to only be due to death. Although most of the results were similar to the previously published results of the NC EPESE, there were a few surprises. For instance, the direct relationship of the CES-D score and age was not supported in previous analysis. Unsurprisingly, the simulations with different percentages of MCAR deaths produced minimal biased mean values for the race-gender groups due to being asymptotically unbiased.

After examining the models' performance in estimating means from incomplete data with MCAR death, we wanted to assess the methods ability to accurately

analyze data with a mixture of reasons for truncation of outcome variables–death and non-participation. The NC EPESE dataset was analyzed again without imputing incomplete data not due to death and compared to the previous results of the imputed NC EPESE data. The results showed that for some outcomes the results where nearly identical. For example, the diastolic blood pressure and activities of daily living (ADL) score rates of change did not differ significantly. Additionally, the probability of being alive and healthy (PAH) joint model was the most stable. Results from the simulation provided evidence that when missing due to NMAR death and non-response the unconditional, pattern-mixture, and partly conditional models underestimated depression and blood pressure means.

We concluded with reviewing the efficiency in the estimation of the marginal models. The models were the most efficient in estimating the regression parameters for the ADL score. Overall, the unconditional model and the partly conditional model were similar in efficiency by sample size and missing assumption. Because the pattern-mixture model for complete lack of information from the non-completers, the pattern-mixture model had higher inefficiency in estimating the fixed effects than the other two models.

Death is a major obstacle in longitudinal studies of older adults. Over the years, studies that have collected longitudinal data from a cohort of older individuals did not always utilize longitudinal techniques to assess the associations and rates of change while controlling for relevant confounding characteristics. Using these techniques to analyze a well-established study of community-dwelling seniors is important to ensure that the most suitable analysis is being conducted, especially when the study is designed to inform policy and best practices. Besides knowing the question of interest to determine the proper method, this research has shown that for certain outcomes the models are able to provide similar results to the initial population even with 60% of the population leaving the study due to death. A closer look at the properties of the

outcomes may be of interest in future research. Additionally, this dissertation has shown that death is correlated to many of the common health outcomes researched in older populations. In order to incorporate the mortality in the analysis of longitudinal responses, we firmly advocate that mortality status be included as a design variable and monitored during the study and, if possible, beyond.

5.1.1 Future Research

This dissertation focused on the performance of the models for their intended purpose and their robustness when other reasons for missingness are present. Each model aims to offer an estimate of the mean of the response. Another component of obtaining a reliable estimate of the population mean is selecting the most parsimonious model. For those models that can be estimated using linear mixed models, Orelien and Edwards (2008) and Edwards et al. (2008) have proposed a R^2 statistic for selecting the fixed effects that contribute to the best model fit. For many social scientists that may consider several variables that are presumed to affect the response, ensuring that power is maximized by not over-fitting is important. Understanding the effectiveness of the R^2 statistic for these models in cohorts with large percentages of deaths would further prescribe the correct usages of these models in longitudinal studies of older populations.

Just as important as correctly specifying the mean model is selecting the most appropriate covariance structure. Although the estimate of $\boldsymbol{\beta}$ remains consistent and asymptotically normal when $\boldsymbol{V} = var(\boldsymbol{y})$ is specified incorrectly, the estimate $\hat{var}(\hat{\boldsymbol{\beta}}) = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}$ is no longer valid nor completely efficient. In this dissertation, the methods were compared assuming a set covariance model, but effort was not made to assess if the assumed covariance was supportive of the data or the most parsimonious. Verbeke and Molenberghs (2000) stated that the sandwich estimator that is employed

to estimate many generalized linear models is less efficient than specifying the correct covariance model.

In our assessment of the proposed models, we only considered modeling data with assumed normal errors. Assessing the models' performance in estimating outcomes with non-normal errors and including the other three models – terminal decline, principal stratification and the joint models – would contribute to the completion of the discussion of these models and their strengths and limitations in estimating rates of change and modeling the variability of longitudinal data with outcomes truncated to death.

One of the many purposes of the North Carolina Established Populations for Epidemiological Studies of the Elderly (NC EPESE) study was to measure the changes in chronic conditions, impairments, and general function in older community-dwelling adults. Nonetheless, some measurements of the chronic conditions and impairments were scheduled very sparsely (e.g., three years for blood pressure measurements). This design weakness could have been accommodated by using other indicators because many illnesses affect the progression of other conditions. Future research should examine the outcomes measured or simulated with smaller gaps in time of measurements. Efforts should also be given to the consideration of the effect of modeling an outcome that has been shown to be highly associated or predictive to other disorders.

As mentioned in the summary, mortality status information could add valuable strength to the analysis of data with truncation due to death. The extent of this strength has yet to be quantified. Further, the circumstances to reach an optimal strength have not been described (e.g., the number of years after the observational period to monitor participants mortality status).

BIBLIOGRAPHY

Alwin, D. F. and Campbell, R. T. (2001). Quantitative approaches: Longitudinal methods in the study of human development and aging. In Binstock, R. H. and George, L. K., editors, *Handbook of Aging and Social Sciences (5th ed)*, pages 22–43. Academic Press, San Diego, CA.

Blazer, D., Burchett, B., Service, C., and George, L. K. (1991). The association of age and depression among the elderly: An epidemiologic exploration. *Journal of Gerontology: Medical Sciences*, 46:210–215.

Blazer, D. G., Landerman, L. R., Hays, J. C., Grady, T. A., Havlik, R., and Corti, M. C. (2001). Blood pressure and mortality risk in older people: Comparison between african americans and whites. *Journal of the American Geriatrics Society*, 49:375–381.

Branch, L. G., Katz, S., Kniepmann, K., and Papsidero, J. A. (1984). A prospective study of functional status among community elders. *American Journal of Public Health*, 74(3):266 – 268.

Crouchley, R. and Ganjali, M. (2002). The common structure of several recent statistical models for droppout in repeated continuous responses. *Statistical Modeling*, 2:39–62.

Crowder, M. (1986). On consistency and inconsistency of estimating equations. *Econometric Theory*, 2:305–330.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39:1–38.

Dempster, A. P., Rubin, D. B., and Tsutakawa, R. K. (1981a). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 76(374):341–353.

Dempster, A. P., Rubin, D. B., and Tsutakawa, R. K. (1981b). Estimation in covariance components models. *Journal of the American Statistical Association*, 76:341–353.

Diehr, P., Patrick, D., Hedrick, S., Rothman, M., Grembowski, D., Raghunathan, T. E., and Beresford, S. (1995). Deaths when measuring health status over time. *Medical Care*, 33:AS164–AS172.

Diehr, P., Williamson, J., Burke, G. L., and Psaty, B. M. (2002). The age and dying processes and the health of older adults. *Journal of Clinical Epidemiology*, 55:269–278.

Diggle, P. J. and Kenward, M. G. (1994). Informative drop-out in longitudinal data analysis. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 43:49–93.

Dufouil, C., Brayne, C., and Clayton, D. (2004). Analaysis of longitudinal studies with death and drop-out: a case study. *Statistics in Medicine*, 23:2215–2226.

Edwards, L. J. (2000). Modern statistical techniques for the analysis of longitudinal data in biomedical research. *Pediatric Pulmonology*, 30:330–344.

Edwards, L. J., Muller, K. E., Wolfinger, R. D., Qaqish, B. F., and Schabenberger, O. (2008). An R^2 statistic for fixed effects in the linear mixed model. *Statistics in Medicine*, 27:6137–6157.

Egleston, B. L., Scharfstein, D. O., Freeman, E. E., and West, S. K. (2007). Causal inference for non-mortality outcomes in the presene of death. *Biostatistics*, 8:526–545.

Egleston, B. L., Scharfstein, D. O., and MacKenzie, E. (2009). On estimation of the survivor average causal effect in observational studies when important confounders are missing due to death. *Biometrics*, 65:497–504.

Elias, M. F. and Robbins, M. A. (1991). Where have all the subjects gone? Longitudinal studies of disease and cognitive function. In Collins, L. M. and Horn, J. L., editors, *Best Methods for the analysis of change: Recent Advances, Unanswered Questions, Future Directions*, pages 264–275. American Psychological Association, Washington, DC.

Ferraro, K. F. and Kelley-Moore, J. A. (2003). A half century of longitudinal methods in social gerontology: Evidence of change in the journal. *Journals of Gerontology: Series B Pyschological Sciences and Social Sciences*, 58:S264–S270.

Frangakis, C. E. and Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, 58:21–29.

Frangakis, C. E., Rubin, D. B., An, M., and MacKenzie, E. (2007). Principal stratification designs to estimate inpute data missing due to death. *Biometrics*, 63:641–662.

Gelfand, A. E. and Smith, A. R. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409.

Gilbert, P. B., Bosch, R. J., and Hudgens, M. G. (2003). Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials. *Biometrics*, 59:531–541.

Gold, D. T., Pieper, C. F., Westlund, R. E., and Blazer, D. G. (1996). Do racial differences in hypertension persist in successful agers? Findings from the MacArthur study of successful aging. *Journal of Aging and Health*, 8:207–219.

Gurka, M. J. (2006). Selecting the best linear mixed model under REML. *The American Statistician*, 60:19–26.

Gurka, M. J., Edwards, L. J., and Muller, K. E. (2011). Avoiding bias in mixed model inference for fixed effects. *Statistics in Medicine*, 30:2696–2707.

Harville, D. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, 72:320–338.

Hayden, D., Pauler, D. K., and Schoenfeld, D. (2005). An estimator for treatment comparisons among survivors in randomized trials. *Biometrics*, 61:305–310.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81:945–960.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47(260):663–685.

Howard, D. L., Carson, A. P., and Kaufman, J. S. (2009). Consistency of care and blood pressure control among elderly African Americans and whites with hypertension. *Journal of the American Board of Family Medicine*, 3:307–315.

Huber, P. J. (1967). The behavior of maximum likelihood estimation under nonstandard conditions. In LeCam, L. M. and Neyman, J., editors, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pages 221–233. University of California Press.

Hughes, J. P. (1999). Mixed effects models with censored data with application to HIV RNA levels. *Biometrics*, 55(2):625–629.

Hypertension Detection and Follow-up Program Cooperative Group (1978). Variability of blood pressure and results of screening in HDFP program. *Journal Chronic Diseases*, 31:651–667.

Jacqmin-Gadda, H., Sibillot, S., Proust, C., Molina, J.-M., and Thiebaut, R. (1980). Robustness of the linear mixed model to misspecified error distribution. *Computational Statistical & Data Analysis*, 48:817–838.

Jennrich, R. L. and Schluchter, M. D. (1986). Unbalanced repeated-measures models with structured covariance matrices. *Biometrics*, 42:805–820.

Johnson, L. L. (2002). Incorporating death into the statistical analysis of categorical longitudinal health status data. PhD dissertation, University of Washington, Department of Biostatistics.

Katz, S., Downs, T. D., Cash, H. R., and Grotz, R. C. (1970). Progress in development of the index of ADL. *The Gerontologist*, 10(1 Part 1):20–30.

Kenward, M. G. and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 53:983–997.

Kuchibhatla, M. N., Fillenbaum, G. G., Hybels, C. F., and Blazer, D. G. (2012). Trajectory classes of depressive symptoms in a community sample of older adults. *Acta Psychiatrica Scandinavica*, 125(6):492 – 501.

Kurland, B. F. and Heagerty, P. J. (2005). Directly parameterized regression conditioning on being alive: Analysis of longitudinal data truncatd by deaths. *Biostatistics*, 6:241–258.

Kurland, B. F., Johnson, L. L., Egleston, B. L., and Diehr, P. H. (2009). Longitudinal data with follow-up truncated by death: Match the analysis method to research aims. *Statistical Science*, 24:211–222.

Laird, N. M. (1988). Missing data in longitudinal studies. *Statistics in Medicine*, 7:305–315.

Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4):963–974.

li, L. W. (2005). Trajectory of ADL disability among community-dwelling frail older persons. *Research on Aging*, 27:56–79.

Liang, K. and Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73:13–22.

Little, R. J. A. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Staistical Association*, 88:125–134.

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, 90:1112–1121.

Little, R. J. A. and Rubin, D. A. (2002). *Statistical Analysis with Missing Data*. Wiley, New York, NY.

Markides, K. S., Dickson, H. D., and Papas, C. (1982). Characteristics of dropouts in longitudinal research on aging: A study of mexican americans and anglos. *Experimental Aging Research*, 8:163–167.

Mcardle, J. J. and Hamagami, F. (1992). Modeling incomplete longitudinal and cross-sectional data using latent growth structural models. *Experimental Aging Research*, 18:145–166.

Miller, R. B. and Wright, D. W. (1995). Detecting and correcting attrition bias in longitudinal family research. *Journal of Marriage and Family*, 57:921–929.

Murphy, T. E., Han, L., Allore, H. G., Peduzzi, P. N., Gill, T. M., and Lin, H. (2011). Treatment of death in the analysis of longitudinal studies of gerontological outcomes. *Journals of Gerontology A Biological Sciences and Medical Sciences*, 66A:109–114.

Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society . Series A*, 135:370–384.

Norris, F. H. (1985). Characteristics of older nonrespondents over five waves of a panel study. *Journal of Gerontology*, 40:627–636.

Orelien, J. G. and Edwards, L. J. (2008). Fixed effect variable seletion in linear mixed models using R^2 statistics. *Computational Statistics and Data Analysis*, 52:1896–1907.

Patterson, D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58:545–554.

Pauler, D. K., McCoy, S., and Moinpour, C. (2003). Pattern-mixture models for longitudinal quality of life studies in advanced stage disease. *Statistics in Medicine*, 22:795–809.

Rabbit, P., Waston, P., Donlan, C., Bent, N., and McInnis, L. (1994). Subject attrition in a longitudinal study of cognitive performance in a community-based elderly people. In Vellas, B. J., Aldarede, J. C., and Garry, P. J., editors, *Facts and research in gerontology*, pages 29–34. Springer Publishing Co. Inc., NY.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1:385–401.

Rajan, K. B. and Leurgans, S. E. (2010). Joint modeling of missing data due to non-participation and death in longitudinal aging studies. *Statistics in Medicine*, 29:2260–2268.

Rathouz, P. J. and Preisser, Jr, J. S. (2013). Missing data: weighting and imputation. In Culyer, T., editor, *Encyclopedia of Health Economics*. Elsevier Ltd, Oxford.

Rhodes, A. R. (2005). Attrition in longitudinal studies using older adults: a meta-analysis. MS thesis, University of North Texas, Department of Psychology.

Ribaudo, H. J., Thompson, S. G., and Allen-Mersh, T. G. (2000). A joint analaysis of quality of life and survival using a random effect selection model. *Statistics in Medicine*, 19:3237–3250.

Robins, J. M. and Greenland, S. (2000). Comments to causal inference without counterfactuals. *Journal of the American Statistical Association*, 95:431–435.

Robins, J. M., Rotnitzky, A., and Zhao, L. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *Journal of the American Statistical Association*, 90(429):106–121.

Rosenbaum, P. R. (1987). Model-based direct adjustment. *Journal of the American Statistical Association*, 82:387–394.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70:41–55.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66:688–701.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63:581–592.

Rubin, D. B. (1980). Comments to randomization analysis of experimental data: the Fisher randomization test, by d. basu. *Journal of the American Statistical Association*, 75:591–593.

Rubin, D. B. (2000). Comments to causal inference without counterfactuals. *Journal of the American Statistical Association*, 95:435–438.

Schaie, K. W. (1996). *Intellectual Development in Adulthood: The Seattle Longitudinal Study*. Cambridge University Press, New York, NY.

Shardell, M. and Miller, R. R. (2008). Weighted estimating equations for longitudinal studies with death and non-monotone missing time-dependent covariates and outcomes. *Statistics in Medicine*, 27:1008–1025.

Siegler, I. C. (1975). The terminal drop hypthesis: Fact or artifact? *Experimental Aging Research*, 1:169–185.

Svetkey, L. P., George, Linda K.and Burchett, B. M., and Morgan, P. A.and Blazer, D. (1993). Consistency of care and blood pressure control among elderly African Americans and whites with hypertension. *American Journal of Epidemiology*, 137:64–73.

Verbeke, G. and Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. Springer, New York, NY.

Wei, G. C. G. and Tanner, M. A. (1990). A monte carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. *Journal of the American Statistical Association*, 85(411):699–704.

White, H. (2007). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 51:5142–5154.

Wilson, R. S., Beckett, L. A., Bienias, J. L., Evans, D. A., and Bennett, D. A. (2003). Terminal decline in cognitive function. *Neurology*, 60:1782–1787.

Yang, Y. C. (2002). Multiple imputation for missing data: concepts and new development (version 9.0). http://support.sas.com/rnd/app/papers/multipleimputation.pdf.

Yao, L. and Robert, S. A. (2011). Examining the racial crossover in mortality between Afrian American and White older adults: A multilevel survival analysis of race, individual socioeconomic status, and neighborhood socioeconomic context. *Journal of Aging Research*, 2011:1–8.

Zeger, S. L., Liang, K., and Albert, P. S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, 44:1049–1060.

Zhang, J. L. and Rubin, D. B. (2003). Estimation of causal effects via principal stratification when some outcomes are truncated by death. *Journal of Educational and Behavioral Statistics*, 28(4):353–368.