Fall 12-2018

Bayesian Based Semiparametric Multidimensional Approaches To Analysis Parkinson'S Disease

Jun Zhang
UTHealth School of Public Health

BAYESIAN BASED SEMIPARAMETRIC MULTIDIMENSIONAL APPROACHES TO ANALYSIS PARKINSON'S DISEASE

by

JUN ZHANG, MS

APPROVED:

_____

HULIN WU, PHD

_____

SHENG LUO, PHD

_____

MOMIAO XIONG, PHD

_____

SUJA RAJAN, PHD

_____

DEAN, THE UNIVERSITY OF TEXAS SCHOOL OF PUBLIC HEALTH

DEDICATION

To my families

BAYESIAN BASED SEMIPARAMETRIC MULTIDIMENSIONAL APPROACHES TO THE ANALYSIS OF PARKINSON'S DISEASE

by

Jun Zhang
BS, DONGHUA UNIVERSITY, 1988
MS, FLORIDA STATE UNIVERSITY, 2004
MPH, UNIVERSITY OF NORTH TEXAS SCHOOL OF PUBLIC HEALTH, 2013

Presented to the Faculty of The University of Texas

School of Public Health

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH
Houston, Texas
December, 2018

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor Sheng Luo for helping me identifying the dissertation topics, for his excellent guidance. I would like to express my sincere gratitude to my committee for their support, encouragement, and guidance in my career and academic endeavors. I would also like to thank my wife Kun Huang for being my mainstay and keeping me on track. Without her support and sacrifices, I can not finish my study.

BAYESIAN BASED SEMIPARAMETRIC MULTIDIMENSIONAL APPROACHES TO THE ANALYSIS OF PARKINSON'S DISEASE

Jun Zhang, MS, PhD
The University of Texas
School of Public Health, 2018

Dissertation Supervisor: Sheng Luo, PhD

In the slow progression of Parkinson's Disease (PD), impairments arise and affect multiple domains (e.g., motor, cognitive, and behavioral). Mixed types, multivariate longitudinal data are commonly used in PD studies. Challenges exist in assessing PD status and investigating disease progression due to lack of biomarkers and ubiquitous impairment in the disease. Collecting disease information from multiple ordinal outcomes (> 10) makes it more difficult and complicated in modeling disease status, disease progression and progressive treatment effects. We proposed a model framework by combining the semiparametric approach and multidimensional framework, and used the proposed model to investigate the heterogeneous disease development and the non-linear treatment effects in the multiple domains predefined in PD.

Furthermore, we extended the semiparametric multidimensional approach to the data with multi-types endpoints. We investigated the multi-type events (competing risks) simultaneously with longitudinal profile in presence of impairment across domains and domain-specific heterogeneous disease progression. Our approach provides an explicit framework for defining and estimating the impaired covariate effects, the association between domain-specific longitudinal profile and multi-type endpoints.

Lastly, we addressed the missing data in PD. We extended the multidimensional joint model to missing data by analyzing two missingness patterns (intermittent and monotone missingness) jointly in domain levels. We provided a statistical method for simultaneous likelihood

inference on missing data in presence of two missingness pattens and two missing mechanisms, missing at random (MAR) and missing not at random (MNAR).

In summary, the studies in this dissertation add to current PD studies by focusing on those ignored or not fully addressed problems in PD. The applications in longitudinal data, survival data and missing data promote this framework usability in public health research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Background

1.1 Literature Review

Parkinson's disease (PD) is a chronic progressive neurodegenerative disorder that places substantial burdens on patients. PD's prevalence is estimated at 0.3% in the general population [50, 75], and affects about 1% of people older than 60 years in the United States along [101]. Because the risk of PD increases with age, the financial and public health burden of PD is expected to increase as the population ages [50]. Currently, the pathogenesis of PD is unknown and there is no cure for PD. Many clinical trials have been conducted to search for effective treatments to slow disease progression (e.g., the completed Derenyl And Tocopherol Antioxidative Therapy of Parkinsonism (DATATOP) study [83], the Neuroprotection Exploratory Trial in PD Long-term Study-1 (LS-1) study [48], and the ongoing Parkinson Progression Marker Initiative (PPMI) study [62]). PD is now considered as a systemic disease because its non-motor symptoms often precede clinical motor signs [11].

PD is characterized by its impairments among multiple domains (e.g., motor, cognitive, and behavioral) [55, 92]. The progressions of disease status have different trajectories, impairments exist within and across domains. Researchers can not rely on single clinical outcome

to define the disease severity and progression. One specific clinical measurement only collects partial PD information from a certain dimension; the information manifested in this measurement could contradict to those from other measurements. For example, we may observe that measurements of non-motor experience of daily living have no change while the measurements of experience of daily living deteriorate fast in the same time. These impairments could happen in overall trend and in any given time horizon. Therefore, no single health outcome reliably reflects the full spectrum of disease severity and progression. All PD clinical studies depend on repeatedly collected multiple health outcomes of mixed types (nominal, ordinal and continuous) to monitor the status of disease and follow the progression for PD patients. In studying PD, all statistical models and analysis have to be based on the data with multivariate longitudinal and clinical outcomes. No robust research and model can be constructed without an effective approach to extricating the disease information from these multivariate longitudinal and clinical outcomes.

To describe and infer the severity of disease and further to evaluate treatment effects, researchers have formulated various frameworks to deal with multivariate and mixed type data. There are two commonly used approaches. One is based on linear combination of several outcomes, and the other approach is to conduct tests using one single primary outcome [5, 25, 74]. However, these approaches either are not able to use all the meaningful information or suffer from other structure flaws. In addition, clinical measurements in PD studies often include dozens of ordinal responses which require effective and efficient statistical approaches. Studies based on sum ranking or adjusted ranking [43, 71] are problematic due to the implied assumption that the scale settings in each item have the same discriminative ability. The magnitude of difference between consecutive ordinal levels in an item is hard to be structured consistently for ordinal outcomes. A recent system review showed that the use of any single biomarker to define disease progression in PD has insufficient evidence [65]. For multiple ordinal outcomes, researchers showed that total score is not a good predictor for implied severity of

a disease [4]. Furthermore the impairment and heterogeneity in PD are hard to be modeled in these approaches. A primary and important prerequisite task to infer disease status and treatment effects from mixed typed data is to construct a working framework to collect the full information from these multivariate outcomes.

In addition to the challenge that there exists no generally accepted biomarker, researchers face with the complication originated from natural features of PD, such as inconsistent information from different domains, heterogeneous disease progression across patients and domains, correlations within, and between outcomes. The existing impairments in PD precludes the conventional analysis approaches which are hard to incorporate the heterogeneity. These PD impairments from various sources make the PD study complicated when trying to analyze study results. Bjomestad *et al.* [7] disclosed that motor complications affected over 50% of PD patients in the first 5 years. A 5-year study conducted by Krack *et al.* [51] showed that there was no consistent clinical improvement for patients across different clinical outcomes. The impairments occur not only across disease domains, but also across time (heterogeneous domain-specific progression), researchers showed that treatment effects over time were not constant in clinical studies [31, 53, 105]. In PD studies, we have no evidence to support that disease progression is constant or linear [38, 97]. Overall, very limited research has been conducted to directly quantify varied covariate effects, temporal PD progression and continuous trajectories of treatment effects.

To address the impairment and overcome the problem that no available biomarker in PD studies, alternative approaches were proposed, low-dimensional interaction model [30] was introduced to model interaction among high dimensional data. Multivariate marginal models [15, 29] provide direct inference for marginal treatment effect but have difficult to handle unbalanced data. While multivariate random effect models [29, 93] have to overcome computation difficulties when number of random effects becomes large. Specifically, multilevel item response theory (MLIRT) model [33, 60, 99] was proposed to analyze the longitudinal scores on the

individual items. Verhagen and Fox [96] extended the MLIRT model by accounting for changes in item characteristics. Schmidt *et al.* [82] introduced a Pretest-Posttest-Posttest multivariate MLIRT model to handle repeated measures. Multivariate MLIRT model can utilize the raw item score based on latent variable model, which provides a better approach compared with sum-score approach. However, the unidimensional latent variable is questionable on the capability of capturing the impairment and heterogeneity from different domains. Other researchers [34, 72] proposed latent variable based multidimensional item response model to assess latent abilities. This cross sectional multidimensional method can not model longitudinal impairment information and correlations, nor is this approach capable to describe the disease progression process. Recently, in order to address the disease impairment in longitudinal item responses, Wang and Luo [98] developed a multidimensional latent trait linear mixed model (MLTLMM). This multidimensional approach provides a framework to deal with the complicated diseases with existing impairment in longitudinal studies. This new approach can effectively utilize a large number of outcomes and is more computational scalable than multivariate marginal and random effects models. Additional advantages include: 1) easy handling of unbalanced data and outcomes of mixed types; 2) explicit accounting for correlation structures using random effects; 3) seamless incorporation of fixed and random effects; 4) capability to capture the domain-specific heterogeneous covariate effects in corresponding domains.

Another challenge in PD is that the temporal heterogeneity in PD such as time-dependent domain level disease progression or treatment effects. In other studies, attempts were made by researchers to describe the natural disease progression with heterogeneity [8]. Auclair-Ouellet *et al.* [3] adressed the cognitive impairment progression over time. Hastie and Tibshirani [39] introduced a general varying coefficient model which is also known as a time-varying coefficient model. To apply to PD, these approaches are not capable to capture the heterogeneity across domains, nor domain-specific continuous temporal effects. In modeling PD progression or progressive treatment effects, we have to avoid structuring the measurements on one or a

few biomarkers which methods were found not appropriate [65], besides we have to construct a method to incorporate the heterogeneous disease progression both across domain and time.

In many clinical studies, participants are monitored longitudinally with respected to the aforementioned dozens ordinal outcomes plus other outcomes, the observations can be stopped by occurrence of terminal events, such as worsening of disease, dropout. Joint models of longitudinal and time-to-event data can be adopted to derive inference about longitudinal profile with informative drop-out (monotone missing) [41]. This has been an active research area for past two decades [15, 52]. However, there are two important reasons calling for additional consideration when modeling these events. First, these events can be viewed as dependent censoring for the initiation of symptomatic treatment which is related to the patient health outcomes. Second, events from disease-related dropout generate non-ignorable missing values in the outcomes. Standard methods for joint modeling of longitudinal and survival data allow for one types of endpoint or events with a single model of failure and have an assumption of independent censoring [23, 40]. In PD studies, more than one possible causes of event or informative censoring typically exist. When there are several causes of event risks, it is known as competing risks. Moreover, treating the outcome-dependent terminal event as independent censoring introduces bias into model estimation. To this end, a joint modeling framework for analyzing all outcomes and events simultaneously is essential, and the requirement to properly handle the ubiquitous heterogeneity in PD call for extra capability for this framework.

Generally, if the endpoints are disease related, it can lead to underestimation of true event time and the drive of events. There can be a great deal of diversity in modeling multiple event risks in PD studies [15, 52]. The extension of classical joint modeling framework were proposed by researchers. Chi and Ibrahim [13] used a joint model for multivariate longitudinal and survival data. Elashoff *et al.* [24] extended the joint model to competing risks data. Dantan *et al.* [14] proposed a joint model with latent state for longitudinal data and event data. However, there is no study addressing the multiple event risks while simultaneously taking into

consideration of the impairments across domains, and domain-specific heterogeneous disease progression.

PD studies can be further complicated by missing data. Missing data in PD clinical trials can seriously undermine the benefits provided by randomization. In PD study, the occurrence of terminal event and other critical events can substantially affect the longitudinal profile and should be analyzed simultaneously. There are two missing data patterns: monotone missing data and intermittent (non-monotone) missing data, based on whether or not the patients will return to the study after the missed visit. Rubin [80] defined three missing data mechanisms. If the missingness is independent of the observed and unobserved data, this missing data mechanism is missing completely at random (MCAR). When missingness is not dependent on unobserved data, it is missing at random (MAR). The missing data belonging to these two missing data mechanisms are treated as 'ignorable' missingness, which do not cause bias in statistical inference for likelihood-based estimation. However, when missingness is associated with the unobserved underlying response process, this missingness is missing not at random (MNAR). For example, patients' dropouts are due to worsening of disease or death. MNAR mechanisms are 'nonignorable'. Under the MNAR assumption, the missing data mechanism needs to be modeled simultaneously with the outcome variables to avoid biased parameter estimates [18].

Estimating parameters with nonignorable missing data is more complex than with ignorable missing data. Recently, modeling longitudinal observations with nonignorable missing data has drawn much attention [56, 57, 106]. Molenberghs *et al.* discussed a selection model for longitudinal ordinal data with nonrandom dropout [68]. Ekholm and Skinner proposed a pattern-mixture model for a longitudinal binary incomplete data set [22]. The full likelihood approach has been used to specify the joint likelihood of outcomes and missing indicators when handling nonmonotone pattern of missing data [44]. For example, the random-coefficient-based selection models were adopted to link dropout time to the longitudinal outcomes through in-

dividual random effects [16, 79, 104]. Alternatively, pseudo likelihood was proposed to provide statistical solutions [90]. Elashoff *et al.* [24] developed the latent random effects model to incorporate effects from nonignorable monotone missing data. Most statistical models focus on one missing data pattern (either monotone or non-monotone missing). Besides, those models are based on one or two outcomes and use the latent traits as predictors for monotone missing and other missing observations. Wu *et al.* [103] proposed a nonlinear mixed-effects model for both monotone and non-monotone patterns of missing data. In PD study, the missed visits are frequently happened during the long follow-up, both patterns of missing data exist in study. How to address the missed responses which consist of dozens ordinal response is an open problem in PD study.

Overall, there are many challenges in studying PD data. Some fundamental problems are still not solved. The study described in this dissertation used a multidimensional latent trait model framework to define PD domain-specific severity and trajectory. In addition, we extended this approach and combined with nonparametric method to jointly analyze PD data in presence of multi-type terminal events and different missingness patterns. In summary, the studies added to current PD studies by addressing those ignored or not fully addressed problems in PD.

1.2 Public Health Significance

In PD studies, researchers and investigators are facing the difficulties to infer true disease status and treatment effects. Defining disease progression and treatment effects in clinical trials with multivariate longitudinal outcome data is an open problem. PD data typically consist a lot of binary/ordinal responses from questionnaire. Unfortunately, there is currently no formal statistical framework can be utilized to define and estimate the treatment effects in clinical trials with multivariate mixed type outcomes. In addition, the traditional unidimensional

latent variable model does not suffice to incorporate the complexity of PD and to model the ubiquitous heterogeneity within and across domains. The heterogeneous nature and unknown pathogenic mechanisms of PD make it impossible to depend on the traditional approaches such as unidimensional measure or unilatent model to define full spectrum of disease severity and progression.

The studies described in this dissertation extend the latent trait model, provides quantified methods for PD impairments, refine the cause-specific competing risk model by incorporating temporal multidimensional disease trajectories, and provide statistical test for missingness mechanism. By providing a comprehensive modeling framework to capture the primary features of PD with various heterogeneities and complexities, the proposed framework in this dissertation has considerable impact on the design and analysis of future clinical studies in neurodegenerative disorders where subtle differences in longitudinally measured multiple outcomes are the primary interest. The development in this dissertation advanced latent trait model and item response model, enriched the investigating tools for public health investigator by shedding light on methods to identify the impairment, heterogeneity and other major features of in PD development. Moreover, we provided an open, general and working methodology for the analysis of PD and can accommodate more sophisticated models.

The overall objectives of the study were to build a more sophisticated class of models that account for known, and currently ignored problems in PD data. Using multidimensional latent trait as started approach, we developed model frameworks for three aims. In Aim 1, the heterogeneous treatment effects with time were defined and investigated based on multivariate mixed type longitudinal outcomes. We proposed a nonlinear multidimensional latent trait model to accommodate the heterogeneity in treatment from time horizon and the impairments across domains. In Aim 2, we extended semiparametric multidimensional latent trait model to data with failure event. A generalized framework for multiple competing dependent censoring events was developed while incorporating heterogeneous disease progression. In Aim 3, we

constructed a model for data in presence of two missingness patterns and MNAR mechanism.

1.3 Specific Aims

1.3.1 Aim 1: To develop a formal semiparametric multidimensional latent trait linear mixed model to define and estimate the domain-specific time-dependent treatment effects

To define and understand the treatment effects on neurodegerative disorders, PD clinical trials depend on collection of various health outcomes data. PD causes impairment in many aspects, and it progresses heterogeneously across domains (e.g., motor, cognitive, and behavioral) and in time. Without a valid biomarker, how to conduct statistical analysis while incorporating existing impairments and heterogeneity across disease domains is an open problem in PD studies. The characteristics of PD determine that treatment effects tend to be time-dependent rather than constant or linear. We combined semiparametric approach and multidimensional approach, and used the proposed model to investigate the development of disease and treatment effects in domain level. Aim 1 of this dissertation provides a set of principled analytic tools and quantitative methods to demonstrate the disease impairment and time-dependent treatment effects in domain levels. In addition, our semiparametric approach is capable to obtain the domain-specific heterogeneous temporal treatment effects.

We are the first to propose the semiparametric multidimensional modeling framework to utilize both continuous and categorical outcomes that provide a high dimensional interpretation of treatment effects and other covariate effects in domain level. The expectation is that by collecting multiple outcomes, clinical studies will provide full spectrum of disease progression, heterogeneous treatment effects, improve the understanding of PD etiology.

1.3.2 Aim 2: To extend the semiparametric multidimensional latent trait model to the data with multi-type terminal events

Joint analysis of longitudinal measurements and time-to-event data is an active area of statistics studies that have received much attention. Joint analysis of the failure times and repeated longitudinal measurements can provide unbiased interpretation for the longitudinal profile interrupted by different disease related events. When applying to PD studies, the conventional joint modeling framework are not capable to incorporate the complicated features of PD, such as impairment within domains, impaired covariate effects, and the heterogeneous disease progression which are common in PD. In addition, the follow-up of PD patients are subject to multi-types of endpoints (e.g., worsening of disease, therapy or dropout), and the dependent censoring can cause the assumption of conventional Cox model being violated. To continue our works in Aim 1 and extend to joint modeling, we proposed a semiparametric multidimensional latent trait joint model for joint analysis of longitudinal multivariate, mixed types outcomes and competing failure time data. Our approach provided an explicit model framework to estimate the domain-specific covariate effects, identify the heterogeneous disease progression and quantify these impacts on event failure time. We allowed multiple latent variables' within-item multidimensionality (one outcome can be a manifestation of more than one latent variable). This approach provides additional clinical insight by quantify the domain level effects from longitudinal profile of disease severity and progression on the risks of terminal event, while the risks are of multiple causes.

We are the first to propose a statistical method to analyze multivariate longitudinal outcomes data and the association with multiple competing dependent censoring events in the presence of impairments across domain and heterogeneous disease progression. This study identified the decomposed effects of PD disease and progression in domain level on the risks of terminal events of various types (e.g., informative dropout, Symptomatic Therapy).

1.3.3 Aim 3: To develop an analysis tool that account for parametric departures from the missing at random (MAR) assumption

Missing data are ubiquitous in longitudinal studies. There are two missing data patterns: monotone missing data and intermittent (non-monotone) missing data. The missed visits can lead bias in PD study if not being treated properly, particularly when these missed visits are missing not at random (MNAR). A number of statistical approaches have been developed to handle missing data. Few studies provide complete inference to process data existing both missing patterns, especially, there is no study addressing the missingness in data that consist of dozens of ordinal responses, no robust test is available to test the missing mechanisms (MNAR or MAR).

In this aim, we attempted to develop an analysis model for both monotone and non-monotone missing data in multivariate longitudinal outcomes of mixed types. We proposed a generalized approach to the analysis of the longitudinal data in the presence of two missing data patterns along with both missing data mechanisms. We extended multidimensional latent trait methods to model and test missing data mechanisms based on the responses from dozens of ordinal outcome. We jointly analyzed the data without excluding MNAR assumption, and assessed missing data mechanisms under the impacts from heterogeneous disease development in multiple domains. One objective of this aim is to build a method capable to handle the data consisting dozens of ordinal responses with existing impairment in multiple domains (e.g. motor, non-motor in Parkinson's Disease). This is the first study addressing the multiple ordinal responses carrying impairment information from multiple domains, and in presence of missing data with different missingness patterns.

Chapter 2

JOURNAL ARTICLE 1: Semiparametric multidimensional latent trait linear mixed model and application in Parkinson's Disease study

2.1 Introduction

Parkinson's disease (PD), is the second most common neurodegenerative progressive movement disorders [75]. In the Unite States, approximately 60,000 individuals are diagnosed with PD each year [50]. The deterioration of PD is irreversible. Thus, diagnosis of PD and accurate assessment of disease progression are critical in treating PD.

PD causes impairment in multiple domains (e.g., motor, cognitive, and behavioral). The complicated nature of PD and limited knowledge on etiology of the disease make it impossible to describe severity of PD and disease progression based on one single or a few clinical measurements. United Parkinson's Disease Rating Scale (UPDRS), which is based on questionnaires and tests is a relatively effective diagnostic measure. The UPDRS has four parts with 55 (each with 5 categories) measuring motor and non-motor symptoms, while each item has 5 different categories. The 44 items in Part I, II and III are most widely used in PD diagnosis and progres-

sion monitoring. UPDRS Part I has 13 items and it measures mentation, behavior and mood (MBM), UPDRS Part II has 13 items and it measures activities of daily living (ADL), UPDRS Part III has 27 items and is used to measure motor examination. More details about UPDRS can be found in Appendix. Many PD studies and clinical trials have adopted UPDRS as primary measure. One of the recent studies is Neuroprotective Exploratory Trials in Parkinson's Disease Long-Term Study 1 (LS-1) study (n=1741), which was the largest cohort of patients with early treated PD. This trial was terminated in August 2013 due to futility of creatine, the targeted medicine.

To describe and infer the severity of disease and further to evaluate treatment effects, researchers have formulated various frameworks to handle multivariate and mixed type data. One is based on linear combination of several outcomes. Another approach is to conduct tests using one single primary outcome [25, 74]. However, these approaches either are not able to use all the meaningful information or suffer from other structure flaws. The sum-score of ordinal responses from questionnaires (e.g., UPDRS) is commonly used to provide an alternative method to deal with ordinal scores. This method leads to loss of information by ignoring differences between response pattens [36]. Furthermore, the development of disease varies among different disease domains. For example, non-motor symptoms can occur much earlier than other symptoms [11]. Both the sum score analysis and linear combination approach are not able to model the impairment and heterogeneity in PD.

To address these issues in the traditional models and utilize multiple outcomes effectively, alternative approaches were proposed, low-dimensional interaction model [30] was introduced to model interaction among high dimensional data. And multilevel item response theory (MLIRT) model [33, 60, 99] was proposed to analyze the longitudinal scores on the individual items. Verhagen and Fox [96] extended the MLIRT model by accounting for changes in item characteristics. Schmidt *et al.* [82] introduced a Pretest-Posttest-Posttest multivariate MLIRT model to handle repeated measures. Multivariate MLIRT model can utilize the raw item score

based on latent variable model, which provides a better approach compared with sum-score approach, the unidimensional latent variable is questionable on the capability of capturing the impairment and heterogeneity from different domains. Other researchers [34, 72] proposed a latent variable based multidimensional item response model to assess latent abilities. However, this cross sectional multidimensional method can not model longitudinal impairment information and correlations. Recently, in order to address the disease impairment in longitudinal item responses, Wang and Luo [98] developed a multidimensional latent trait linear mixed model (MLTLMM). This multidimensional approach provides a framework to deal with the complicated diseases with impairment in longitudinal studies. This new approach can effectively utilize a large number of outcomes and is more computational scalable than multivariate marginal and random effects models. Additional advantages include: 1) easy handling of unbalanced data and outcomes of mixed types; 2) explicit accounting for correlation structures using random effects; 3) seamless incorporation of fixed and random effects; 4) capability to capture the domain-specific heterogeneous covariate effects in corresponding domains.

Furthermore, the heterogeneity in PD is not limited to the impairment across domains. The disease progression and temporal covariate effects are not necessary to be linear or unchanged over the follow-up period. Researchers demonstrated that complications or side effects could happen for treatments in clinical studies. For example, Salat *et al.* [81] disclosed the impaired treatment effects in different stages in addition to the impaired effects across domains. Recently, Auclair-Ouellet *et al.* [3] addressed the cognitive impairment progression over time. All these studies reveal that it will lead to bias by assessing the treatment effects with the assumptions of unchanged or linear treatment effects, especially for chronic and progressive diseases like PD. Attempts were made by researchers to describe the natural PD progression with heterogeneity [8]. However, previous studies mainly addressed the heterogeneous symptoms or focused on assessment of comparisons at several fixed time intervals. To the best of our knowledge, the issues on the full range heterogeneity of disease progression and covariate effects

with time, the different heterogeneous PD trajectories and covariate effects across domains over the entire PD trials are not addressed. In addition, no study addressed the difference between short-term and long-term treatment effects in LS-1 study.

In this study, we extended MLTLMM model by proposing a semiparametric multidimensional latent trait linear mixed model that allows domain-specific trajectories of non-linear disease progression, and within-item multidimensionality (allowing inputs from more than one latent variables). This model has two levels. The first level multivariate latent trait model defines the relationship between a patient's multidimensional unobserved disease severity scores and the observed multivariate outcomes, while the second level semiparametric multidimensional linear mixed model (SMLMM) connects the high-order latent disease scores (incorporating the continuous change rate) to covariates, time and subject-specific random effects. Our model allows the cross dimensional correlated effects (e.g., correlation between MBM and ADL, ADL and motor, etc.), allows the different temporal covariate effects in different domains. Because the number of latent disease traits is much smaller than the number of observed outcomes, models are quite parsimonious, which can improves the computational feasibility and model interpretability.

The remainder of this article proceeds as follows. In section 2, we discuss the proposed model, Bayesian inference and model selection. Section 3 presents studies to assess the performance of the proposed models. In section 4, we apply our method to the motivating studies and compare our semiparametric multilatent model to other parametric models. Section 5 provides concluding remarks and discussion.

2.2 Model and estimation

2.2.1 Semiparametric multidimensional latent trait linear mixed model

Let $y_{ik}(t)$ be the observed outcome k from subject i at time t, where $i = 1, \ldots, N$, $k = 1, \ldots, K$, and $t = t_1, \ldots, t_{J_i}$. All outcomes are properly coded so that larger values are worse clinical conditions. We assume that there are P (with $P < K$) latent variables (LVs), each of which represents the underlying disease severity in a specific domain. We use $\boldsymbol{\theta}_i(t) = (\theta_i^{(1)}(t), \ldots, \theta_i^{(p)}(t), \ldots, \theta_i^{(P)}(t))'$ to denote the domain-specific unobserved disease status for subject i at time t, where the superscript (p) $(p = 1, \ldots, P)$ denotes the latent variable in the pth domain. To incorporate mixed types, multivariate outcomes, we construct the first level generalized linear model as,

$$G(y_{ik}(t)) = a_k + \boldsymbol{b_k'\theta_i(t)}, \tag{2.1}$$

where $G(\cdot)$ is link function which depends on the types of outcome. In linear regression, it is identity transformation, while dealing with binary or ordinal responses, it usually takes the logit transformation. More specifically, for continuous responses, we have $y_{ik}(t) = a_k + \boldsymbol{b_k'\theta_i(t)} + \varepsilon_{ik}(t)$, where a_k and $\boldsymbol{b_k}$ are the outcome-specific parameters, while the random errors $\varepsilon_{ik} \sim N(0, \sigma_{\epsilon_k})$. For the ordinal outcomes (e.g., item responses) we use an extended two-parameter model, $\text{logit}\big\{p(y_{ik}(t) \leq l|\theta_i(t))\big\} = a_{kl} - \boldsymbol{b_k'\theta_i(t)}$, where $l = 1, 2, \ldots, n_k - 1$ is the lth level of the kth random variable, which is ordinal with n_k levels, while a_{kl} and $\boldsymbol{b_k}$ are the difficulty parameter and discrimination parameter (vector) correspondingly. The negative sign for $\boldsymbol{b_k}$ in the ordinal model is to ensure that worse disease severity (higher $\theta_i(t)$ value) is associated with a more severe outcome (higher $y_{ik}(t)$). The model enables loading of latent variables from all domains. To model the dependence of severity scores $\boldsymbol{\theta}_i(t)$ on covariates, we

propose the second level semiparametric multivariate linear mixed model (SMLMM)

$$\theta_i^{(p)}(t) = \boldsymbol{X}_i^{(p)}(t)\boldsymbol{\beta}^{(p)} + f^{(p)}(t)w_i + \boldsymbol{Z}_i^{(p)}(t)\boldsymbol{u}_i^{(p)} + e_i^{(p)}(t), \tag{2.2}$$

where $\boldsymbol{X}_i^{(p)}(t)$ and $\boldsymbol{Z}_i^{(p)}(t)$ are the covariates corresponding to the fixed and random effects respectively. The latent variable $\theta_i^{(p)}(t)$ denotes ith subject's unobserved disease severity in the pth domain at time t. We use $f^{(p)}(t)$ to model the time-dependent effects for the covariate w in the pth domain, which can be easily extended to multiple covariates. The latent variables are continuous, higher value indicating worse severity of disease. The vector $\boldsymbol{u}_i = (\boldsymbol{u}_i^{(1)'}, \ldots, \boldsymbol{u}_i^{(P)'})'$ contains the random effects for the ith subject, it follows a multidimensional normal distribution, $\boldsymbol{N(0, \Sigma)}$, where $\boldsymbol{\Sigma}$ is the covariance matrix with dimension equal to the number of random effects incorporated. There are several ways to model random effects. For example, when we incorporate fully correlated random intercepts and random slopes into the framework, this covariance matrix will have the dimension of $2p \times 2p$. The residual part $e_i^{(p)}(t)$ is assumed to be mutually independent, and $e_i^{(p)}(t) \sim N(0, \sigma_e^{(p)})$.

In this study, we try to investigate the decomposed (domain-specific), time-dependent treatment effects. We use $\theta_i^{(p)}(t) = \boldsymbol{X}_i^{(p)}(t)\boldsymbol{\beta}^{(p)} + \beta_1^{(p)}t + f^{(p)}(t)trt_i + u_{i0}^{(p)} + u_{i1}^{(p)}t + e_i^{(p)}(t)$ to model the domain-specific non-linear treatment effects (can extend to other covariates), where $\beta_1^{(p)}$ is the average disease progression rate (positive for getting worse) in the pth domain for participants in control group. The $f^{(p)}(t)$ is the add-on time-dependent treatment effects in the pth underlying disease domain, denoting the average heterogeneous treatment effects with time for those in treatment group, adjusted for fixed covariates and subject-specific random effects. The nonparametric function is structured as $f^{(p)}(t) = \sum_{n=1}^{N} c_n^{(p)} B_{n,q}(t)$, where $B_{n,q}(t)$ is order q spline basis for the nth knot at time t, while $c_n^{(p)}$ is the penalized coefficient for corresponding basis in the pth domain, and $c_n^{(p)} \sim N(0, \sigma_c^{(p)})$.

The major feature of this two-level modeling is that all outcomes incorporate the whole

dimensional $\boldsymbol{\theta_i(t)}$ or disease information from all domains, and multi-type measurements can be effectively utilized. First, the latent variable of each domain models the impact from other domains using cross-domain random effects. Besides, this model allows both between-item multidimensionality and within-item multidimensionality (i.e. one outcome can be a manifestation of more than one latent variable). Information on the impairment of disease is captured by dimensional latent variables and domain specific heterogeneous temporal effects. Other researchers discussed the cross-loading concept in the factor analysis and applied in independent cluster structure [64, 78]. We extend and generalize this approach by conceptualizing the cross-domain disease severity interaction with impaired effects, and providing a framework to incorporate the heterogeneous temporal covariate effects in different domains. Generally, the domain impact on the clinical outcomes can be regressed by the loading intensity of domain-specific latent variable in the cross-loading vector. When there is one latent variable ($P = 1$), our model reduces to the univariate latent variable model as a special case.

To model cross-loading in the model, we let $\boldsymbol{a} = (\boldsymbol{a}_1', \ldots, \boldsymbol{a}_k', \ldots, \boldsymbol{a}_K')'$, and $\boldsymbol{a}_k = (a_{k,1}, \ldots, a_{k,n_k-1})'$ for the kth ordinal outcome with n_k categories. We let $\boldsymbol{b} = \boldsymbol{(b_1, \ldots, b_K)'}$, a K by P matrix, where $\boldsymbol{b_k} = (b_k^{(1)}, \ldots, b_k^{(p)})'$. Because the model is over-parameterized, additional constraints are required to make it identifiable. The indeterminacy between the latent variable loadings $\boldsymbol{b}_k$ and the scales of the latent variables $\boldsymbol{\theta}_i(t)$ can be fixed by either setting one element in each column of $\boldsymbol{b}$ to be 1, or letting $\sigma_e^{(p)} = 1$ for $p = 1, \ldots, P$ with at least one of the loadings constrained to be positive for each factor [20]. Finally, to identify parameters $\boldsymbol{a}$ and intercepts in regression coefficients, we set the constraints on one selected item in each domain, we let $a_{p,1} = 0$ (or other constant) for $p = 1, \ldots, P$ ordinal outcomes and the order constraint $a_{k,1} < \ldots < a_{k,l} < \ldots < a_{k,n_{k-1}}$ must be satisfied. Besides, we set identifiability constraints on vectors $\boldsymbol{b}$, for example, when $P = 3$ and the constraints are put on the first three items, we let $b_1^{(1)} = b_2^{(2)} = b_3^{(3)} = 1$, all other elements are 0. In real data analysis, in order to achieve the better domain calibration and locate the three bases, we have to carefully select the item to

put constraints for each domain.

2.2.2 Nonparametric function and Likelihood

We use nonparametric functions to model non-linear covariate effects. The smoothness of nonparametric functions can be adjusted by the number of knots while penalized method can use a small number of knots. Wu and Zhang [102] suggested using cubic smoothing spline to achieve better smoothness. Among different spline approaches, the local basis cubic B-spline has many advantages in numerical computation. Generally, cubic B-spline with 10 to 20 basis can achieve ideal results with little advantage gained for more than 20 knots [37]. In this study, we use cubic B-spline with 8 inner knots (total 12 knots including intercept). To avoid over-fitting, we adopt penalized cubic B-spline function in our model. Penalized approach is based on Eilers method [21].

Let the parameter vector $\boldsymbol{\Theta} = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \sigma_{\epsilon_k}, \sigma_e^{(p)}\}$. Conditional on the random effects $\boldsymbol{u}_i$ and nonparametric part, all measurements of each subject are assumed to be independent. We have the full likelihood of subject i as follows:

$$L(\boldsymbol{\Theta}, \boldsymbol{u}_i; \boldsymbol{y}_i) = \left[\prod_{j=1}^{J_i}\prod_{k=1}^{K} p(y_{ik}(t_{ij})|\boldsymbol{u}_i, \boldsymbol{c})\right] p(\boldsymbol{u}_i|\boldsymbol{\Sigma})p(\boldsymbol{c})p(\boldsymbol{\Sigma}), \tag{2.3}$$

where $p(\boldsymbol{u}_i|\boldsymbol{\Sigma})$ is the density function of random effects vector $\boldsymbol{u}_i$.

2.2.3 Bayesian inference

To make inference on the parameter vector $\boldsymbol{\Theta}$, we use Bayesian methods based on Markov chain Monte Carlo (MCMC) posterior simulations. We use vague priors on all elements in $\boldsymbol{\Theta}$, except for the aforementioned constrained parameters, i.e., $a_{p,1} = 0$ (or other constant) for $p = 1, \ldots, P$ and $b_p^{(p)} = 1$, for all p. Specifically, the prior distributions of unconstrained

parameters a_k of the continuous outcomes is $a_k \sim N(0, 10,000)$ (SD=100). To obtain the prior distributions for the threshold parameters of ordinal outcome k, we let $a_{k,1} \sim N(0, 400)$ (SD=20), and $a_{k,l} = a_{k,l-1} + \Delta_l$ for $l = 2, \ldots, n_k - 1$, with $\Delta_l \sim N(0, 10,000)I(> 0)$ (SD=100), i.e., normal distribution left truncated at 0. The setting of high SD here is considering the scenarios that there are rare responses to some top level of items (some items have few responses to the level 5). Prior distributions for unconstrained elements in $\boldsymbol{b}$ and $\boldsymbol{\beta}$ are $N(0, 400)$ (SD=20). We sample the first nonparametric coefficient $c_1^{(p)}$ in each domain from $N(0, \sigma_c^{(p)})$, while the remained coefficients are sampled using random walk, $c_k^{(p)} \sim N(c_{k-1}, \sigma_c^{(p)})$. We use the Cholesky factorization method to sample the correlation coefficients, the random effects covariance matrix is modeled as $\boldsymbol{\Sigma} = \boldsymbol{\sigma_U} \boldsymbol{\Sigma_U} \boldsymbol{\sigma_U'}$, where $\boldsymbol{\Sigma_U}$ is the correlation matrix, while $\boldsymbol{\sigma_U}$ is the diagonal matrix of standard deviation of random effects. All variances terms are from Inverse-Gamma$(0.01, 0.01)$. We have investigated other selections of vague prior distributions with various hyper-parameters and obtained very similar results.

The posterior samples are obtained from the full conditional of each unknown parameter using Hamiltonian Monte Carlo (HMC) [19] and No-U-Turn Sampler (NUTS) [42]. Both HMC and NUTS samplers are implemented in `Stan`, which is a probabilistic programming language implementing statistical inference. The model fitting is performed in `Stan` (version 2.17.0) [87] by specifying the full likelihood function and the prior distributions of all unknown parameters. For large datasets, `Stan` may be more efficient than `BUGS` language [59] in achieving faster convergence and requiring smaller number of samples [42]. To monitor Markov chain convergence, we use the trace plots and view the absence of apparent trends in the plot as evidence of convergence. In addition, we use the Gelman-Rubin diagnostic to ensure the scale reduction $\hat{R}$ of all parameters are smaller than 1.1 as well as a suite of convergence diagnosis criteria to ensure convergence [32]. To facilitate reading and implementation of the proposed model, `Stan` codes are posted in the supplement part.

2.2.4 Model selection

To compare the proposed model to other models, we implement model selection in this study. Among the various model selection methods available in Bayesian inference, we select the deviance information criterion (DIC), and Watanabe-Akaike information criterion (WAIC).

The deviance information criterion (DIC) assesses model fit based on the posterior mean of the deviance and a penalty on the model complexity [86]. Due to mixture framework applied in our model, we choose the DIC_3 measurement [10]. The DIC_3 is defined as $DIC_3 = \overline{D(\boldsymbol{\theta})} + \tau_D$, where $\overline{D(\boldsymbol{\theta})} = -2\text{E}_{\boldsymbol{\theta}|\mathcal{D}}\{\log[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\}$ is the posterior mean deviance, $\tau_D = \overline{D(\boldsymbol{\theta})} + 2\log\{\text{E}_{\boldsymbol{\theta}|\mathcal{D}}[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\}$ is a measure of the effective number of parameters in the model, and $\text{E}_{\boldsymbol{\theta}|\mathcal{D}}(.)$ is the expectation with respect to the joint posterior distribution $\pi(\boldsymbol{\theta}|\mathcal{D})$. Thus, we have $DIC_3 = -4\text{E}_{\boldsymbol{\theta}|\mathcal{D}}\{\log[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\} + 2\log\{\text{E}_{\boldsymbol{\theta}|\mathcal{D}}[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\}$. Applying Monte Carlo approximation, the expression of DIC_3 is

$$\widehat{DIC_3} = -\frac{4}{M}\sum_{m=1}^{M}\sum_{i=1}^{I}\log\left\{f(\boldsymbol{y}_{ij}|\boldsymbol{\theta}^{(m)})\right\} + 2\log\left\{\frac{1}{M}\sum_{m=1}^{M}\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta}^{(m)})\right\}.$$

A smaller value of DIC_3 indicates a better-fitting model.

WAIC [100] can be viewed as an improvement on the DIC for Bayesian models. DIC has gained popularity through its implementation in the graphical modeling package BUGS [59]. WAIC is fully Bayesian and closely approximates Bayesian cross-validation. Unlike DIC, WAIC is invariant to parametrization and even works for singular models [91]. The WAIC is defined as

$\sum_{i=1}^{I}\left(\frac{1}{M}\sum_{m=1}^{M} f(y_i|\theta^{(m)})\right) - \sum_{i=1}^{I} var_{post}\left(logf(y_i|\theta)\right)$, where M is total sampling times. A small WAIC value denotes a better model.

2.3 Simulation studies

In this section, we conduct simulations studies to investigate the identifiability of the proposed model. We incorporate 3 domains in simulation. We generate 240 datasets with $N = 800$ subjects and eleven visits (baseline and ten follow-up visits) for each subject. The data have ten ordinal outcomes ($K = 10$) and each has 5 categories. Data are simulated using the 3 LVs model: $\theta_i^{(p)}(t) = \beta_0^{(p)} + \beta_1^{(p)}t + f^{(p)}(t)x_i + u_i^{(p)}$, where $p = 1, 2, 3$, the covariate x_i takes value 0 or 1, each with probability 1/2 to mimic the treatment assignment. The time vector is set as $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{i11})' = (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10)'$. We set the regression coefficients to be $\boldsymbol{\beta}^{(1)} = (\beta_0^{(1)}, \beta_1^{(1)})' = (-0.5, 0.5)'$, $\boldsymbol{\beta}^{(2)} = (\beta_0^{(2)}, \beta_1^{(2)})' = (-0.8, 0.8)'$, $\boldsymbol{\beta}^{(3)} = (\beta_0^{(3)}, \beta_1^{(3)})' = (-0.2, 1.2)$. We are using the following functions, $-1.4sin(x/2) - x^{1.2}/5$, $x^{0.8}/250 - 1.5sin(x/3)$ and $-x^3/400 + 0.1x$ to mimic non-linear treatment effects in three domains. We use 3×3 covariance matrix to simulate random effects, we set $\boldsymbol{\sigma_u} = (\sigma_u^{(1)}, \sigma_u^{(2)}, \sigma_u^{(3)}) = (0.7, 0.8, 1)$, and $\boldsymbol{\rho} = (0.4, 0.6, -0.1)$. For this 3LV model, we assign true value to the constrained items, which are the first three items, $a_{1,1}$=$a_{2,1}$=$a_{3,1} = 0$, $b_1^{(1)} = b_2^{(2)} = b_3^{(3)} = 1$. Other parameters' setting are presented in Appendix.

2.3.1 Model performance

Table 2.31 displays the simulation results. In the table, bias (the average of posterior means minus the true parameter values), standard deviation (SD, for the posterior means), and coverage probabilities (CP) are presented. The summarized parameter estimates for ordinal items are presented in Appendix. The estimated non-linear functions are visualized in Figure 2.31, which demonstrates the fitting between estimated time-dependent functions and true non-linear functions. Furthermore, point-wise coverage rates for estimated three domains' non-linear functions are presented in Figure 2.32 to assess the performance of the proposed nonparametric

framework.

Table 2.31: Simulation results based on three domains setting.

	Domain 1				Domain 2				Domain 3			
	EST	Bias	SD	CP	EST	Bias	SD	CP	EST	Bias	SD	CP
Latent variable												
β_0	-0.501	-0.001	0.070	0.946	-0.796	0.004	0.103	0.908	-0.184	0.016	0.153	0.929
β_1	0.499	-0.001	0.011	0.950	0.799	-0.001	0.017	0.912	1.197	-0.003	0.026	0.912
Random effects												
σ	0.699	-0.001	0.031	0.971	0.793	-0.007	0.039	0.950	0.992	-0.008	0.060	0.950
ρ^*	0.391	-0.009	0.056	0.954	0.592	-0.008	0.046	0.967	-0.112	-0.012	0.071	0.938

* $\hat{\rho^{12}} = 0.391$, $\hat{\rho^{23}} = 0.0.592$, $\hat{\rho^{13}} = -0.112$.

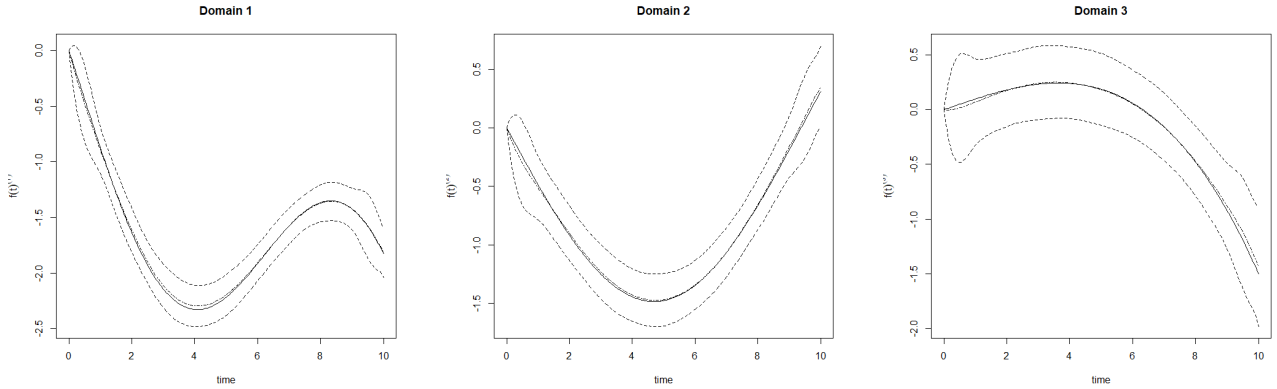

Figure 2.31: Comparison of true functions(solid) and estimated nonparametric functions (dashed) with 95% credible intervals (dot dash). Left panel: Domain 1. Middle panel: Domain 2. Right panel: Domain 3.

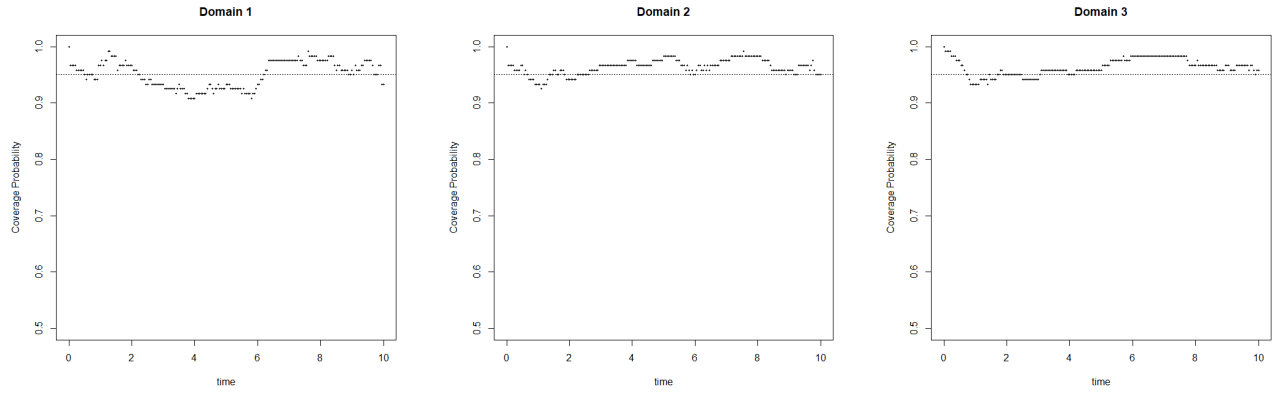

Figure 2.32: Point-wise coverage probabilities with reference lines (dotted) at 0.95. Left panel: Domain 1. Middle panel: Domain 2. Right panel: Domain 3.

In summary, the estimated parameters have low bias, and the estimated non-linear

functions provide good fit for the true non-linear temporal effects with coverage probabilities around 95%. We conclude that the proposed model is identifiable, in presence of non-linear covariate effects.

2.4 Application to LS-1 study

We apply the proposed semiparametric multidimensional latent variable model and Bayesian framework to LS-1 study. A total of 1741 participants are included in the study. We use the 44 item responses in UPDRS part I, II and III as outcome responses. Based on guideline of UPDRS, the 4 items in part I are targeting for MBM, part II's 13 items are for ADL, part III's 27 items are for motor examination. The structured questionnaire design (part I, II, and III) confines the information manifested by those ordinal responses in each part to the corresponding disease domains. To fit to the data structure, we refine our models assuming that item responses in each part manifest the unobserved status of corresponding disease domains. We add dependency across domains by incorporating fully correlated random effects (both between and cross domains). Considering the data structure, we only incorporate between-item multidimensionality in each domain based on UPDRS structured parts. As both age and gender are important risk factors for PD [26, 70], we include age and gender covariates into model. Hence, the latent trait model is updated as $\theta_i^{(p)}(t) = \beta_0^{(p)} + \beta_1^{(p)} age_i + \beta_2^{(p)} gender_i + \beta_3^{(p)} t + f^{(p)}(t) trt_i + u_{i0}^{(p)} + u_{i1}^{(p)} t + e_i^{(p)}(t)$, where $p = 1, 2, 3$ corresponding to disease domains for MBM, ADL and Motor examination, while $f^{(p)}(t)$ is the domain-specific non-linear treatment effects over time, the covariate trt_i is treatment indicator, indicating if patient i in treatment group or not (1 for creatine group, 0 for placebo), $u_{i0}^{(p)}$ and $u_{i1}^{(p)}$ are the random effects on the pth domain. The $2P$ dimensional vector $\boldsymbol{U_i} = (u_{i0}^{(1)}, u_{i1}^{(1)}, \ldots, u_{i0}^{(P)}, u_{i1}^{(P)})'$ is set to follow multivariate normally distribution $N_{2p}(\mathbf{0}, \boldsymbol{\Sigma})$. We impose identifiability constraints on one selected item in each domain, so that these three

items from different domains construct the bases of three dimensions, and calibrate each latent variable's measure. We use factor analysis to select these three items. In each UPDRS part (MBM, ADL and motor), we conduct factor analysis to select the item which has the largest factor loading. After confirmatory factor analysis, we choose item 1 (Mentation) from UPDRS Part I, item 7 (Hygiene) from UPDRS Part II and item 18 (Hand grips) from UPDRS Part III to put constraints. For these selected items, we set $a_{k,1} = 0$ (can be other constant) and $b_k^{(p)} = 1$.

Table 2.41 displays the coefficient estimates for the fixed covariates. For participants in placebo group (not using creatine as therapy), disease is significantly getting worse in all three domains. The domain-specific time effects show disease progresses the fastest in ADL, and the slowest in motor examination. Specifically, the participants without creatine therapy have average 0.437 (95%CI: [0.398, 0.473]) units yearly worsening rate in ADL, and average 0.227 (95%CI: [0.202, 0.253]) units yearly worsening rate in motor, while the average worsening rate in MBM is 0.263 (95%CI: [0.223, 0.301]) units per year. Other studies show non-motor symptoms can occur much earlier than other symptoms [11], our study discovers that at the current stage (during follow-up) non-motor disease progression has the second fastest deterioration rate among three domains. Besides the significant time effects, we also identify both significant age and gender (male) effects in all three disease domains. Our model shows that the age factor has the highest risk intensity in motor examination, and has the least risk intensity in MBM. Specifically, every one year increase in age associates with 0.106 (95%CI: [0.020, 0.189]) units increase in MBM severity measure (getting worse), 0.146 (95%CI: [0.044, 0.241]) units increase in ADL severity measure, and 0.297 (95%CI: [0.235, 0.359]) units increase in motor examination. Compared with female, Male would have 0.186 (95%CI: [0.008,0.356]) units of increased MBM severity measurement (getting worse), 0.310 (95%CI: [0.124, 0.504]) units of increased ADL measurement, and 0.183 (95%CI: [0.050, 0.312]) units of increased motor examination measurement. These results are consistent with other researchers' conclusion [9,

70, 77]. The different covariate risk loadings across domains imply the impairment in different disease domains, and are explicitly displayed in those three-domain coefficient estimations. In summary, increased ages associates with worse disease severity, and this positive association is highest in motor examination, least in MBM. The gender effects show that male has significantly higher risk in disease deterioration compared to female. And this gender related deteriorated disease effects are the highest in ADL, the least in Motor examination.

Table 2.41: Parameter estimates by domains.

	Mentation, behavior and mood				Activities of daily living				Motor examination			
	Mean	SD	95% CI		Mean	SD	95% CI		Mean	SD	95% CI	
Int.	−0.769	0.093	−0.951	−0.593	−1.164	0.103	−1.379	−0.978	−0.146	0.065	−0.262	−0.022
Age (yr)	0.106	0.044	0.020	0.189	0.146	0.049	0.044	0.241	0.297	0.033	0.235	0.359
Male	0.186	0.090	0.008	0.356	0.310	0.100	0.124	0.504	0.183	0.069	0.050	0.312
Time (yr)	0.263	0.020	0.223	0.301	0.437	0.020	0.398	0.473	0.227	0.013	0.202	0.253

Figure 2.41 displays domain-specific treatment effects trajectories over the follow-up period. Both treatment effects in MBM and ADL progress to the wrong direction, and these deterioration effects turn to significant at the end of year 5. There is no any treatment benefit gained in MBM and ADL domains after first year. In a very short time window, there is some treatment benefit at the start of trial, but the effects are not significant. As regards to motor examination, there are some insignificant treatment benefits at the beginning, but generally there is no significant desired treatment effects over the whole follow-up period. The domain-specific trends show the divergence of short-term (prior year 1) and long-term treatment effects. This finding is consistent with the conclusion of the exploratory trial on 2001 sponsored by National Institute of Neurological Disorders and Stroke (NINDS). In that trial, NINDS found that creatine monohydrate was the only one which pass the futility analysis of 2 clinical trials [45, 46]. This potential early treatment effects was not being studied furthermore as majority of studies based on (LS-1) dataset focused on 5-year change from baseline and used this as criteria to assess long-term effects. Indeed, our finding reaches a partial similar conclusion made by Fine *et al.* [27], which disclosed that there was no sustained improvement for patients with advanced PD undergone unilateral posteroventral medical pallidotomy though there existed

significant early improvements. The graphic illustration of impaired domain-specific treatment effects is presented in Appendix A.01. In the end, the three different trajectories show that the undesired treatment effects in MBM and ADL contribute most for the failure of trial. This finding is first reported and discovered using our model.

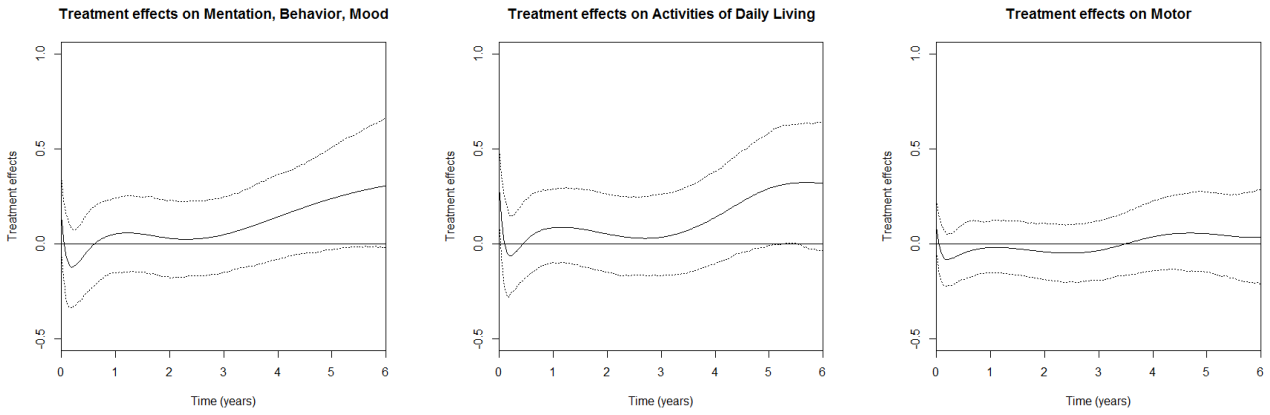

Figure 2.41: Estimated time-dependent treatment effects by domains, estimated functions: dashed lines; 95% point-wise credible intervals: solid lines. Left panel: Mentation, behavior and mood. Middle panel: Activities of daily living. Right panel: Motor examination.

Random effects related posterior parameters are presented in Appendix Tables A.03 and A.04. Table A.03 shows similar random intercept variabilities ($\sigma_0^{(1)}$=1.646, 95% CI: [1.537, 1.773], $\sigma_0^{(2)}$=1.844, 95% CI: [1.749, 1.943], and $\sigma_0^{(3)}$=1.250, 95% CI: [1.190, 1.315]) across domains, same findings are with random slope ($\sigma_1^{(1)}$=0.337, 95% CI: [0.307, 0.349], $\sigma_1^{(2)}$=0.474, 95% CI: [0.439, 0.507], $\sigma_1^{(3)}$=0.327, 95% CI: [0.307, 0.351]). The correlation matrix in Appendix Table A.04 provides additional information of inter-dependency of diseases and correlated (in high dimensional) subject-specific characteristics between and within domains. Every individual has 3 pairs of random effect terms sampled from a 6×6 covariance matrix, denoting the random effects in 3 dimensions or domains, each domain has both random intercept and random slope. For each random effects vector, the 1st and 2nd elements are for MBM domain, the 3rd and 4th are for domain in ADL, while the last two are for motor examination domain. The only significant within-domain correlations (between random intercept and random slope in same

domain, $\rho = -0.093$, 95% CI: $[-0.158, -0.021]$) resides in ADL, explained as: if disease status is worse at the start of the trial, the disease progression is slow in ADL during follow-up. The intercept between-domain correlations (random intercept vs random intercept) are all positive and significant. These associations reveal that in the trial if any initial disease status is worse, the other two are also worse, for example, if a participant's initial disease status in MBM is worse compared to the average level at the start of trial, the disease severity levels in ADL and motor are also worse. Same finding is discovered in the pairwise correlations of within-domain's random slopes (random slope vs random slope), if a the participant has faster disease progression rate compared to the average rate in any domain, his or her disease development in other domains also progresses faster compared to the average, for example, disease develops fast in ADL, it also deteriorates fast in MBM and motor examination.

The difficulty and discrimination parameter estimations are presented in Appendix Tables A.06, A.07 and A.08. All parameter estimates are significant for these ordinal responses. The results indicate that the unobserved domain-specific disease status is significantly manifested by these 44 ordinal outcomes.

To assess the performance of proposed nonparametric model compared with other parametric models, we conduct model comparison using the criteria which are discussed in Section 2.2.4. We test both linear and quadric multidimensional latent trait models. For linear setting, the multi-dimensional latent trait is $\theta_i^{(p)}(t) = \beta_0^{(p)} + \beta_1^{(p)} age_i + \beta_2^{(p)} gender_i + \beta_3^{(p)} t + \beta_4^{(p)} trt_i + \beta_5^{(p)}(t \times trt_i) + u_{i0}^{(p)} + u_{i1}^{(p)} t + e_i^{(p)}(t)$, while the quadric setting is $\theta_i^{(p)}(t) = \beta_0^{(p)} + \beta_1^{(p)} age_i + \beta_2^{(p)} gender_i + \beta_3^{(p)} t + \beta_4^{(p)} trt_i + \beta_5^{(p)}(t \times trt_i) + \beta_6^{(p)} t^2 + \beta_7^{(p)}(t^2 \times trt_i) + u_{i0}^{(p)} + u_{i1}^{(p)} t + e_i^{(p)}(t)$. Table 2.42 displays the DIC_3 and WAIC values for these three models. The results suggest nonparametric approach outperforms the other two models for either selection criteria.

Table 2.42: Model selection criteria for LS-1 study.

	Nonparametric	Linear Model	Quadric Model
DIC_3	731124.6	731203.1	731156.7
WAIC	730780.4	730867.3	730794.6

2.5 Discussion

In this study, we propose a semiparametric multidimensional latent trait model to assess domain-wise time-dependent treatment effects and use it to the LS-1 study which was the longest clinical trial in PD study and included the largest cohort of PD patients. This model provides a general approach for researchers to investigating impaired covariate effects and heterogeneous disease progression in domain levels. In this model, the domain-specific latent variable is served as underlying severity of disease in a certain dimension, manifested by assorted clinical outcomes. We adopt a Bayesian inference framework based on Markov chain Monte Carlo (MCMC) to identify the time-dependent treatment effects and trajectories. The extensive simulation studies show that our model can accurately estimate the domain-specific non-linear covariate effects in the presence of multidimensionality.

A number of studies on LS-1 evaluated the long-term treatment effects by assessing change in functional performance for a fixed interval such as 5-year duration [5, 48]. However, the effects comparison based on fixed time interval does not provide the whole picture of the performance of the targeted treatment. In addition, those studies can not distinguish the domain-specific treatment effects (i.e. treatment on motor or non-motor, etc.). Compared to other studies on the LS-1 trial, our semiparametric multidimensional framework discloses the treatment performance of creatine in depth, such as when the treatment effects become deteriorated, in which domain the treatment deteriorates and to what extend. Moreover, we demonstrate the different treatment trajectories in different domains and different time lines for the wrong treatment effect in sub-domain.

There are some limitations in our model that we will address in our future studies. In this article, the domains are pre-specified by the structured UPDRS questionnaires. While in many multivariate studies, the domains and the number of domains to be used in model will be hard to define. Using factor analysis can facilitate in determining domains. On the other hand, how to explain the structured domains clinically, and map responses to these domains are not straightforward and hard to be evidence-based. In the future study, we will continue to investigate the method to address these issues. Besides, in modeling random effects, we chosen a full correlated multivariate normal distribution, it is because the flexibility and interoperability of the structure while accommodating correlation of both within and between multiple latent variables. Generally, misspecification of random effects distribution has little impact on the parameters [63]. In the future, we will investigate the mixture normal distributions based on Dirichlet process [49].

Chapter 3

JOURNAL ARTICLE 2: A semiparametric multidimensional latent trait model to joint analysis of longitudinal outcomes and competing risks

3.1 Introduction

Parkinson's disease (PD), is the second most common neurodegenerative disorder. It is diagnosed in about 1% of individuals over the age of 65 in the United States [6]. PD is an incurable and progressive disorder. Current understanding of PD suggests that it is a multiorgan disorder presenting with heterogeneous clinical conditions [47, 67]. PD is not just a complex motor disorder, it is now considered as a systemic disease due to its non-motor symptoms in addition to the motor symptoms [11].

The lack of validated biomarkers for PD is the major barrier in PD study [67]. Currently, the diagnosis method mainly depends on clinical information provided by patients, for example self-reported motor sign and symptoms (rigidity, tremor, etc.,). The Unified Parkinson's Disease Rating Scale (UPDRS) is constructed as a scale measure of neuronal impairment in PD [76]. There are a lot of PD studies and clinical trials using UPDRS to follow the longitudinal course

of PD [11, 88, 89]. However, there are two major limitations for this version of test: lack of consistent anchor among subscales and not sufficient emphasis on the non-motor features of PD. In 2007, a revised UPDRS, the Movement Disorder Society-UPDRS (MDS-UPDRS) was introduced to provide more comprehensive and more accurate tests than the original UPDRS [35]. The MDS-UPDRS consists of 65 items, all items are anchored with five categories, from 0 to 4, the higher the score the worse the disease status. The first three parts of MDS-UPDRS which include 59 items are commonly used in PD studies. Statistical methods are required to extract useful information from these 59 item responses to define disease status and its progression.

In PD studies, the items response in UPDRS or MDS-UPDRS are often summed up to obtain total score, which is treated as a continuous outcome. It is easy to implement but leads to loss of information by ignoring differences between response item pattens [36]. Alternatively, multilevel item response theory (MLIRT) model was utilized to analyze the longitudinal scores. This model links the multiple items to an unobserved disease status structured as a univariate latent variable. However, the unidimensional framework limits the application of the model in analyzing PD due to the complication of the disease. PD is characterized by existing impairments across domains, for instance non-motor symptoms often occur decade before the clinical motor signs[11]. The unidimensional framework does not suffice to define the impairments in motor and non-motor. To address these issues in the traditional models, multidimensional item response model was introduced [72]. Though cross-sectional impairments were addressed in this model, the longitudinal impairment information and correlations are still not fully assessed in this cross-sectional multidimensional model. Recently, Wang and Luo [98] proposed a new multidimensional latent trait linear mixed model (MLTLMM) to address the disease impairment in longitudinal study. This new multidimensional latent trait model allows multiple latent variables and within-item multidimensionality. By adopting latent disease score to reduce the number of observed outcomes, MLTLMM is more computational scalable than multivariate

marginal and random effects models.

Though impairments across domains can be modeled in this multidimensional latent trait model, the temporal impairment or heterogeneous disease progression is not taken into account in the model framework. Researchers [8, 27] have shown that development of PD is gradual and not necessary linear, and this heterogeneous progression has different trajectories in different domains. PD has domain-specific temporal patterns that may be dependent on PD clinical and/or pathological stages. Overall, to the best of our knowledge, the domain-specific heterogeneous disease progression is still not adequately addressed in previous studies. In order to fully assess the temporal patterns of PD in longitudinal study and allow for a more flexible description of disease development, we propose a semiparametric multidimensional latent trait model which incorporates domain-specific high-order temporal effects. This semiparametric MLTLMM provides a coherent and explicit framework to facilitate clinicians and researchers to evaluate the temporal patterns of disease change [62] and impairment across domains, help to design precise personal-specific medicine or treatment.

In PD clinical trials, participants are monitored longitudinally with respected to the aforementioned dozens ordinal outcomes plus other outcomes, the observations can be stopped by occurrence of terminal events, such as worsening of disease, dropout. Joint analysis of the failure times and repeated longitudinal measurements can provide solutions to this issue. However, there are two important reasons calling for additional consideration when modeling these events. First, they can be viewed as dependent censoring for the initiation of symptomatic treatment which is related to the patient health outcomes. Second, disease-related dropout events generate non-ignorable missing values in the outcomes. Standard methods for joint modeling of longitudinal and survival data allow for one types of endpoint or events with a single model of failure and have an assumption of independent censoring [23, 40]. When there are several causes of event risks, it is known as competing risks. Moreover, treating the outcome-dependent terminal event as independent censoring introduces bias into model estimation. In

PD studies, participants are exposed to the risks of the terminal events, while the events are not independent. For example, patients can drop out or receive therapy, any occurrence of these events can censor the potential of other risks (dependent censoring). Moreover, if the endpoints are disease related, it can lead to underestimation of true event time of other types and the drive of events. In clinical trials, researchers expect to gain more efficiency in statistical inferences with a joint model and utilize multiple types endpoints.

There can be a great deal of diversity in modeling multiple event risks in PD studies, The extension of classical joint modeling framework were proposed by researchers. Chi and Ibrahim [13] used a joint model for multivariate longitudinal and survival data. Elashoff *et al.* [24] extended the joint model to competing risks data. Dantan *et al.* [14] proposed a joint model with latent state for longitudinal data and event data. However, there is no study addressing the multiple event risks while simultaneously taking into consideration of the impairments across domains, and domain-specific time-dependent disease progression. To model PD's longitudinal impacts on events, or vise versa, we go beyond the standard formulation of joint models. We extend the proposed semiparametric multidimensional model by incorporating impairment and heterogeneity to survival sub-models. In addition, we address the cause-specific associations between multiple types event outcomes (competing events), and the domain-specific disease status with non-linear disease progression.

The joint model presented in this study differs in several aspects to the previous approaches. First, our model obtains and utilizes the information in heterogeneous disease progression manifested in dozens of ordinal responses, which approach helps to ensure the full range and non-linear longitudinal trajectories to be fully evaluated in the model. Second, our model decomposes the impairments from different domains and quantify these domain-specific associations with different types of events. Overall, our model simultaneously takes into account the impaired covariate effects across domains, heterogeneous domain-specific disease progression and multiple failure risks. The remainder of this article proceeds as follows. In section

2 we describe the motivating study and the data structure. Section 3 discusses the proposed model, Bayesian inference. Section 4 presents simulation studies to assess the performance of the proposed models. In section 5, we apply our method to the motivating study. Section 6 provides concluding remarks and discussion.

3.2 Motivating clinical study

This methodological development is motivated by Parkinson's Progression Markers Initiative (PPMI) study. PPMI is an ongoing longitudinal observational study that aims to identify one or more markers of progression for PD. The study was launched in 2010. All participants were grouped into several cohorts, including Parkinson Disease (PD), scans without evidence of dopaminergic degeneration (SWEDD) and healthy control (HC) etc. At baseline, patients were not expected to require PD medications within at least 6 months. PD medications without any restriction on number or type might be initiated at any time based on discretion of the patients or treating physicians. PD cohort includes 423 subjects. After excluding those having only one visit, there are 415 subjects in study. Among these 415 subjects, total 40 dropped out early for different reasons, and 197 individuals underwent Symptomatic Therapy (ST). The disease progression was mainly assessed using MDS-UPDRS scales plus other measurements. According to the MDS, the 13 items in Part I are used to measure the disease information in non-motor aspects of experiences of daily living (nM-EDL), the 13 items in Part II are used to measure information in motor aspects of experiences of daily living (M-EDL), while the 18 grouped items (some items have setting for right, left or other body parts' sub-items, total 33 items) in Part III collect information related to motor examination. Items in the different parts are assumed to be the manifestation of different disease domains (motor, cognitive and behavior). Due to impairments, the disease progresses heterogeneously both in dimensions and with time. No study was conducted to identify and define disease progression in domain level,

such as disease development in motor or non-motor, no continuous comparison was carried for disease development by domains, though it is known that domain-specific disease trajectory varies with time, for example, non-motor symptoms precede motor's [12].

In PPMI study, there exist two types of events which can affect the longitudinal observations. One is Symptomatic Therapy, the other one is dropout. Though participants have scheduled visits for more than 6 years, the dropouts stop the planned visits. Besides dropouts, ST could cause longitudinal observations deviated from the original trajectories. Generally, researchers treat ST as one type of terminal event [84]. Figure 3.21 shows these two types of events. The patient 3179 underwent ST at end of year 3, which censored the possibility for dropout. Patient 3023 dropped out from study before the end of year 5, the longitudinal observations stopped. When modeling these endpoints in the study, we are trying to answer these questions, how these events associate with disease development, which domain's disease progression has more impacts on the events for this multi-domain disease. Furthermore, we are going to investigate which domain's disease status constitutes the primary causes for the events.

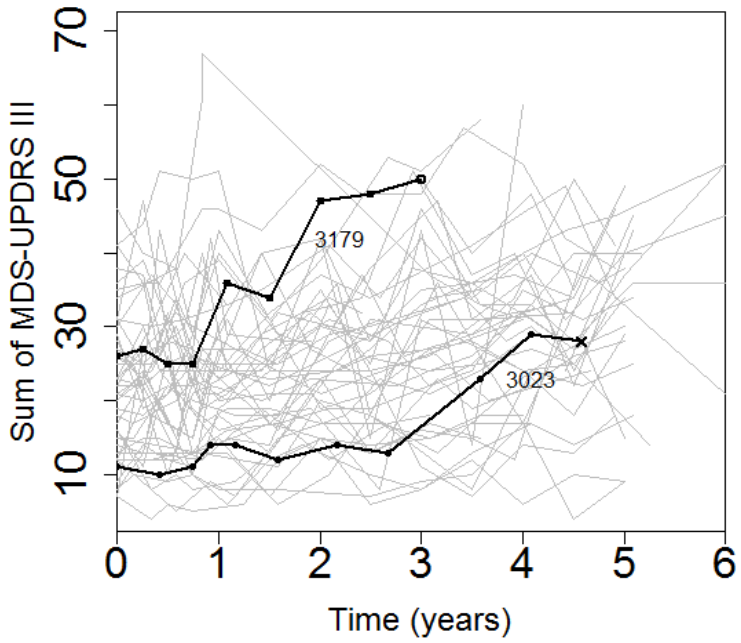

Figure 3.21: MDS-UPDRS III Trajectories of fifty randomly selected patients in PPMI study. Patient 3179 underwent ST at year 3 (empty circle), patient 3023 dropped out before the end of year 5 (cross).

3.3 Model formulation

3.3.1 Multidimensional latent trait linear mixed model (MLTLMM)

Let $y_{ik}(t)$ be the observed outcome k for subject i at time t, where $i = 1, \ldots, N$, $k = 1, \ldots, K$, and $t = t_{i1}, \ldots, t_{iJ_i}$. All outcomes are coded so that larger values are worse clinical conditions. To start building the MLTLMM modeling framework, we assume that there are P (with $P < K$) latent variables (LVs) representing the underlying disease severity scores and denote them as $\boldsymbol{\theta}_i(t) = (\theta_i^{(1)}(t), \ldots, \theta_i^{(p)}(t), \ldots, \theta_i^{(P)}(t))'$ for subject i at time t, where the superscript (p) $(p = 1, \ldots, P)$ denotes the pth latent variable. From a clinical perspective, each latent variable denotes the severity of a PD domain (e.g., non-motor and motor). We introduce

the first level MLTLMM model for continuous outcomes.

$$y_{ik}(t) = a_k + \boldsymbol{b}_k'\boldsymbol{\theta}_i(t) + \varepsilon_{ik}(t), \tag{3.1}$$

where a_k and $\boldsymbol{b}_k = (b_k^{(1)}, \ldots, b_k^{(P)})'$ are the outcome-specific parameters, and the random errors $\varepsilon_{ik}(t) \sim N(0, \sigma_{\varepsilon_k})$ are independent and identically distributed. Note that $a_k = E[y_{ik}(t)|\boldsymbol{\theta}_i(t) = \mathbf{0}]$ is the mean of the kth outcome if the disease severity scores are 0 and $b_k^{(p)} = [y_{ik}(t') - y_{ik}(t)]/\theta_i^{(p)}(t)$ is the expected increase in the kth outcome for one unit increase in the pth disease severity score while holding other disease severity scores unchanged. The parameter $b_k^{(p)}$ also plays the role of bringing up the pth disease severity score to the scale of the kth outcome. When vector $\boldsymbol{b}_k$ has different entries, the disease severity of different domains (latent scores) in $\boldsymbol{\theta}_i(t)$ have varied manifestations in the kth outcome. We model the binary outcomes (e.g., yes/no in questionnaire) and ordinal outcomes (e.g., each item of MDS-UPDRS) by using a two-parameter model [58] as follows:

$$\text{logit}\big\{p(y_{ik}(t) = 1|\boldsymbol{\theta}_i(t))\big\} = a_k + \boldsymbol{b}_k'\boldsymbol{\theta}_i(t), \tag{3.2}$$

$$\text{logit}\big\{p(y_{ik}(t) \leq l|\boldsymbol{\theta}_i(t))\big\} = a_{kl} - \boldsymbol{b}_k'\boldsymbol{\theta}_i(t), \tag{3.3}$$

where $l = 1, 2, \ldots, n_k - 1$ is the lth level of the kth random variable, which is ordinal with n_k levels. The probability of being in a particular category is $p(y_{ik}(t) = l) = p(y_{ik}(t) \leq l|\boldsymbol{\theta_i(t)}) - p(y_{ik}(t) \leq l-1|\boldsymbol{\theta_i(t)})$. Interpretation of parameters is similar to continuous outcomes, except that modeling is on the log-odds, not the native scale of the data. Note that the negative sign for $\boldsymbol{b}_k$ in the ordinal outcome model is to ensure that worse disease severity (higher $\boldsymbol{\theta}_i(t)$) is associated with a more severe outcome (higher $y_{ik}(t)$). A major feature of this model is that $\boldsymbol{b_k}$ plays the role of incorporating $\boldsymbol{\theta}_i(t)$ or explicitly bringing up P dimensional disease severities to the outcomes. This is one of the simplest ways to conceptualize the disease severity scores that allows to define the overall treatment effects. To model the dependence of severity scores

$\boldsymbol{\theta}_i(t)$ on covariates, we propose the second level multivariate linear mixed model (LMM)

$$\theta_i^{(p)}(t) = \boldsymbol{X}_i^{(p)}(t)\boldsymbol{\beta}^{(p)} + f^{(p)}(t) + \boldsymbol{Z}_i^{(p)}(t)\boldsymbol{u}_i^{(p)} + e_i^{(p)}(t), \tag{3.4}$$

where $\boldsymbol{X}_i^{(p)}(t)$ and $\boldsymbol{Z}_i^{(p)}(t)$ are the covariates corresponding to fixed and random effects, respectively, for each latent variable $\theta_i^{(p)}(t)$. The model can include covariates of interest such as treatment and temporal effects. The vector $\boldsymbol{u}_i = (\boldsymbol{u}_i^{(1)'}, \ldots, \boldsymbol{u}_i^{(P)'})'$ contains all the random effects from multi-domains for subject i, which are assumed to be normally distributed as $N_P(0, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is the covariance matrix. The method of modeling random effects can take other formats, for example, normal mixture [94], t-distribution [73] and Laplace distributions [17]. The correlation among the latent variables and domains are accounted for by modeling the correlation among the elements in $\boldsymbol{u}_i$. The residuals $e_i^{(p)}(t)$ are assumed to be independent to $\boldsymbol{u}_i$ and $e_i^{(p)}(t) \sim N(0, \sigma_e^{(p)})$. The non-linear disease progression is modeled using nonparametric formulation as $f^{(p)}(t) = \sum_{n=1}^{N} c_n^{(p)} B_{n,q}(t)$, where $B_{n,q}(t)$ is order q spline basis for the nth knot at time t, while $c_n^{(p)}$ is the penalized coefficient for corresponding basis in the pth domain, and $c_n^{(p)} \sim N(0, \sigma_c^{(p)})$. Indeed, we have domain-specific structure, $\theta_i^{(p)}(t) = \boldsymbol{\beta^{(p)}}\boldsymbol{x_i} + f^{(p)}(t) + u_{i0}^{(p)} + u_{i1}^{(p)}t + e_i^{(p)}(t)$, where $\boldsymbol{\beta^{(p)}}$ is the covariate effects on the pth underlying disease's domain, and is domain-specific. The null hypothesis of no covariate effects is $H_0 : \beta_1^{(1)} = \ldots = \beta_1^{(p)} = \ldots = \beta_1^{(P)} = 0$. This framework allows not only different covariate effects on multidimensional disease domains to facilitate varied disease progression and prognosis, but also the combined covariate effects and overall disease progression to be interpreted on the scales of the observed outcomes. In addition, this model allows both between-item multidimensionality and within-item multidimensionality (some of the items require input from more than one latent variables). This provides a method to conceptualize the non-linear disease progression in different domains. Because the number of outcomes (K) has been reduced to a smaller number of latent disease severity scores (P, with $P < K$), models are quite parsi-

monious in terms of number of random effects, which improves computational feasibility and model interpretability.

For notational convenience, we let $\boldsymbol{a} = (\boldsymbol{a}_1', \ldots, \boldsymbol{a}_k', \ldots, \boldsymbol{a}_K')'$, and $\boldsymbol{a}_k = (a_{k,1}, \ldots, a_{k,n_k-1})'$ for the kth ordinal outcome with n_k categories. We let $\boldsymbol{b} = \boldsymbol{(b_1, \ldots, b_K)'}$, a K by P matrix, where $\boldsymbol{b_k} = (b_k^{(1)}, \ldots, b_k^{(p)})'$. Because the model is over-parameterized, additional constraints are required to make it identifiable. The indeterminacy between the latent variable loadings $\boldsymbol{b}_k$ and the scales of the latent variables $\boldsymbol{\theta}_i(t)$ can be fixed by either setting one element in each column of $\boldsymbol{b}$ to be 1, or letting $\sigma_e^{(p)} = 1$ for $p = 1, \ldots, P$ with at least one of the loadings constrained to be positive for each factor [20]. Finally, to identify parameters $\boldsymbol{a}$ and intercepts in regression coefficients, we set the constraints on one selected item in each domain, we let $a_{p,1} = 0$ (or other constant) for $p = 1, \ldots, P$ ordinal outcomes and the order constraint $a_{k,1} < \ldots < a_{k,l} < \ldots < a_{k,n_{k-1}}$ must be satisfied. Besides, we set identifiability constraints on p orthogonal vectors $\boldsymbol{b}$, for example, when $P = 3$ and the constraints are put on the first three items, we let $b_1^{(1)} = b_2^{(2)} = b_3^{(3)} = 1$, all other elements are 0. In real data analysis, in order to achieve the better domain calibration, we have to carefully select the item to put constraints for each domain.

3.3.2 Non-parametric approach

We use nonparametric functions to model time-dependent covariate effects in each domain. The smoothness of nonparametric functions can be adjusted by changing the number of knots, while penalized method can use a small number of knots. Wu and Zhang [102] suggested using cubic smoothing spline to achieve better smoothness. Among different spline approaches, the local basis cubic B-spline has many advantages in numerical computation. Gray [37] suggested that cubic B-spline with 10 to 20 basis can achieve ideal results with little advantage gained for more than 20 knots. In this study, we use cubic B-spline with 8 knots. To avoid

over-fitting, we adopt penalized cubic B-spline function in our model. Penalized approach is based on Eilers method [21].

3.3.3 Cause-specific proportional hazard model

In survival studies, cause-specific proportional hazard model is widely used in modeling competing risk [23]. There are other competing risk models, such a cumulative incidence function (CIF) model proposed by Fine and Gray [28] and multi-state model [2]. In this study, we use cause-specific proportional hazard model. To model the disease-related endpoints, we link the events to longitudinal observations by assuming that the occurrences of G competing events dependent on the unobserved disease status and the progression (disease related). Let survival observation for ith subject be $\boldsymbol{C_i} = (T_i, \delta_i)$, where δ_i is the censoring indicator, and T_i is the failure time. The observed data structure is augmented to $(\boldsymbol{Y, T, \delta, U})$ while combining with longitudinal observations. We construct cause-specific proportional hazard functions as,

$$\begin{aligned} \lambda_{gi} &= \lim_{h\to 0} \tfrac{P(T_i<t+h|T_i\geq t, U_i, X(t))}{h}, \\ &= \lambda_{0g}(t)\exp\{\boldsymbol{W_i'\gamma_g} + \boldsymbol{\nu_g'\phi(\theta_i)}, \end{aligned} \tag{3.5}$$

where λ_{0g} $(g = 1, \ldots, G)$ denotes the baseline hazard for gth type risk, and the parameter γ_g modulates covariate effect of $\boldsymbol{W_i}$ for ith subject $(i = 1, \ldots, I)$ on the gth risk of terminating follow-up, it can be same or different from the covariate vector $\boldsymbol{X}$ in latent trait. We use $\boldsymbol{\phi(\theta_i)} = (\phi_1(\theta_1), \ldots, \phi_P(\theta_P))'$ for functional form of latent traits. The association parameter vector is defined as $\boldsymbol{\nu_g} = (\nu_g^{(1)}, \ldots, \nu_g^{(m)}, \ldots, \nu_g^{(P)})'$, while $\nu_g^{(m)}$ denotes the effects of disease severity and progression in the mth latent score on the hazard of gth type survival outcomes.

This model assumes that the instantaneous hazard is associated with the current expected disease status at time t. Different to other joint models, each hazard risk incorporates the full impacts from underlying different disease development which is domain-specific in this

model. The cross loading of domain-specific impact is conceptualized by parameter vector $\boldsymbol{\nu_g}$, which can impose each type's hazard to receive the influence of P domains longitudinal profile. Specifically, we incorporate the domain-specific heterogeneous disease progression into cause-specific propositional hazard model in addition to the impairment across domains. Generally, a positive value of $\nu_g^{(p)}$ indicates the patient with worse disease status or deteriorate rate in pth domain is going to undergo event g earlier.

With $\boldsymbol{C_i} = (T_i, \delta_i)$, $Ls_i = \{\prod_{g=1}^{G} \lambda_g(T_i)^{I(\delta_i=g)}\} \exp\Big[-\int_0^{T_i} \sum_{g=1}^{G} \lambda_g(s)ds\Big]$ is the conditional likelihood for ith subject. We define the piecewise constant hazard function as $\lambda_{0g}(t) = \sum_{l=1}^{L} h_l I_l(t)$, where $I_l(t) = 1$ if $(\tau_l < t \leq \tau_{l+1})$, and 0 otherwise.

3.3.4 Numerical approach & likelihood

For simplicity, we use the simple form of disease severity function for θ_i, then the above hazard function changes to $\lambda_{0g} exp(\boldsymbol{W_i'\gamma_g} + \sum_1^P \nu_g^{(p)}(\boldsymbol{\omega_i'\beta^{(p)}} + f(t_j)^{(p)} + u_{i0}^{(p)} + u_{i1}^{(p)}(t_j))$. The likelihood for survival observations $C_i = (T_i, \delta_i)$ contributed by ith subject takes the form:

$$
\begin{aligned}
&L_i = p(T_i, \delta_{ik} | \boldsymbol{\Theta_s, u_i, b}), \\
&= \prod_{g=1}^{G} \lambda_g(T_i)^{I(d_i=g)} exp[-\int_0^{T_i} \sum_{g=1}^{G} \lambda_g(s)ds], \\
&= \prod_{g=1}^{G} \lambda_{0g} exp(\boldsymbol{W_i'\gamma_g} + \sum_1^P \nu_g^{(p)}(\boldsymbol{\omega_i'\beta^{(p)}} + f(T_i)^{(p)} + u_{i0}^{(p)} + u_{i1}^{(p)} T_i)^{I(d_i=g)} \\
&\times exp[-\int_0^{T_i} \sum_{g=1}^{G} \lambda_{0g} exp(\boldsymbol{W_i'\gamma_g} + \sum_1^P \nu_g^{(p)}(\boldsymbol{\omega_i'\beta^{(p)}} + f(s)^{(p)} + u_{i0}^{(p)} + u_{i1}^{(p)} s)ds.
\end{aligned}
\tag{3.6}
$$

There is no explicit form for integration of the high order nonparametric functions (e.g. cubic nonparametric) to estimate time dependent disease progression in the competing sub-

model. Our proposed approaches require numerical integration to approximate the calculation. We use Gauss-Kronrod quadrature method which can provide a very good approximation in most applications. In practice, the numerical grids and non-parametric knots have to be processed together in our algorithm to approximate the integration numerically, and acquire desirable precision.

Finally the full likelihood contribution for the ith subject, conditional on the parameters and random effects takes the form:

$$\begin{aligned} &L(\boldsymbol{y_i}, T_i, \delta_{ik}|\boldsymbol{\Theta}, \boldsymbol{u_i}) = \textstyle\prod_{k=1}^{K} p(\boldsymbol{y_{ik}}|\boldsymbol{\Theta_y}, \boldsymbol{u_i}) \prod_{g=1}^{G} p(T_i, \delta_{ig}|\boldsymbol{\Theta_s}, \boldsymbol{u_i}), \\ &= \textstyle\prod_{k=1}^{K} p(\boldsymbol{y_{ik}}|\boldsymbol{\Theta_y}, \boldsymbol{u_i}) \prod_{g=1}^{G} \{\lambda_g(T_i)^{I(\delta_i=g)} exp[-\int_0^{T_i} \sum_{g=1}^{G} \lambda_g(s)ds]|\boldsymbol{\Theta_s}, \boldsymbol{u_i}\}, \end{aligned} \tag{3.7}$$

where $\boldsymbol{\Theta} = (\boldsymbol{\Theta_y}, \boldsymbol{\Theta_s})$ denotes the parameter vector for both longitudinal parameter vector $\boldsymbol{\Theta_y}$, and survival parameter vector $\boldsymbol{\Theta_s}$.

3.3.5 Bayesian inference

To make inference on the parameter vector $\boldsymbol{\Theta}$, we use Bayesian methods based on Markov chain Monte Carlo (MCMC). We use vague priors on all elements in $\boldsymbol{\Theta}$, except for the aforementioned constrained parameters, i.e., $a_{p,1} = 0$ (or other constant) for $p = 1, \ldots, P$ and $b_1^{(p)} = 1$, for all p. Specifically, the prior distributions of unconstrained parameters a_k of the continuous outcomes is $a_k \sim N(0, 10,000)$ $(sd = 100)$. To obtain the prior distributions for the threshold parameters of ordinal outcome k, we let $a_{k,1} \sim N(0, 400)$, and $a_{k,l} = a_{k,l-1} + \Delta_l$ for $l = 2, \ldots, n_k - 1$, with $\Delta_l \sim N(0, 10,000)I(> 0)$, i.e., normal distribution left truncated at 0. The setting of high SD is considering the scenarios of rare responses to some top level of items (some items have few responses for the level 5). Prior distributions for unconstrained elements in $\boldsymbol{b}$ and $\boldsymbol{\beta}$ are $N(0, 400)$. We sample the first nonparametric coefficient $c_1^{(p)}$ in each domain from $N(0, \sigma_c^{(p)})$, while the remained coefficients are sampled using random walk, $c_k^{(p)} \sim N(c_{k-1}, \sigma_c^{(p)})$.

We use the Cholesky factorization to estimate the correlation coefficients, the random effects covariance matrix is expressed as $\boldsymbol{\Sigma} = \boldsymbol{\sigma}'_{\boldsymbol{u}} \boldsymbol{\Sigma}_{\boldsymbol{U}} \boldsymbol{\sigma}_{\boldsymbol{u}}$, where Σ_U is the correlation matrix, while $\boldsymbol{\sigma}'_{\boldsymbol{u}}$ is the diagonal matrix of standard deviation of random effects. All variance parameters are from Inverse-Gamma$(0.001, 0.001)$. We sample the association parameter $\nu_g^{(p)}$ using $\nu_g^{(p)} \sim N(0, 400)$. The covariate coefficients in the sub-models are using $\gamma_g \sim N(0, 400)$. For piecewise hazards in the sub-models, we sample them from Gamma $(0.01, 0.01)$. We have investigated other selections of vague prior distributions with various hyper-parameters and obtained very similar results.

The posterior samples are obtained from the full conditional of each unknown parameter using Hamiltonian Monte Carlo (HMC) [19] and No-U-Turn Sampler (NUTS) [42]. Both HMC and NUTS samplers are implemented in `Stan`, which is a probabilistic programming language implementing statistical inference. The model fitting is performed in `Stan` (version 2.17.0 [87]) by specifying the full likelihood function and the prior distributions of all unknown parameters. For large datasets, `Stan` may be more efficient than `BUGS` language [59] in achieving faster convergence and requiring smaller number of samples [42]. To monitor Markov chain convergence, we use the trace plots and view the absence of apparent trends in the plot as evidence of convergence. In addition, we use the Gelman-Rubin diagnostic to ensure the scale reduction $\hat{R}$ of all parameters are smaller than 1.1 as well as a suite of convergence diagnosis criteria to ensure convergence [32].

3.3.6 Model selection

There are a wide variety of model selection criteria in Bayesian inference. Among the various model selection methods available in Bayesian inference, we use the deviance information criterion (DIC), and Watanabe-Akaike information criterion (WAIC).

The deviance information criterion (DIC) assesses model fit based on the posterior mean

of the deviance and a penalty on the model complexity [86]. Due to mixture framework applied in our model, we choose the DIC_3 measurement [10]. The DIC_3 is defined as $DIC_3 = \overline{D(\boldsymbol{\theta})} + \tau_D$, where $\overline{D(\boldsymbol{\theta})} = -2\text{E}_{\boldsymbol{\theta}|\mathcal{D}}\{\log[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\}$ is the posterior mean deviance, $\tau_D = \overline{D(\boldsymbol{\theta})} + 2\log\{\text{E}_{\boldsymbol{\theta}|\mathcal{D}}[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\}$ is a measure of the effective number of parameters in the model, and $\text{E}_{\boldsymbol{\theta}|\mathcal{D}}(.)$ is the expectation with respect to the joint posterior distribution $\pi(\boldsymbol{\theta}|\mathcal{D})$. Thus, we have $DIC_3 = -4\text{E}_{\boldsymbol{\theta}|\mathcal{D}}\{\log[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\} + 2\log\{\text{E}_{\boldsymbol{\theta}|\mathcal{D}}[\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta})]\}$. Applying Monte Carlo approximation, the Bayesian based DIC_3 is

$$\widehat{DIC_3} = -\frac{4}{M}\sum_{m=1}^{M}\sum_{i=1}^{I}\log\left\{f(\boldsymbol{y}_{ij}|\boldsymbol{\theta}^{(m)})\right\} + 2\log\left\{\frac{1}{M}\sum_{m=1}^{M}\prod_{i=1}^{I} f(\boldsymbol{y}_{ij}|\boldsymbol{\theta}^{(m)})\right\}.$$

A smaller value of DIC_3 indicates a better-fitting model.

WAIC [100] can be viewed as an improvement on the DIC for Bayesian models. DIC has gained popularity through its implementation in the graphical modeling package BUGS [59]. WAIC is fully Bayesian based and closely approximates Bayesian cross-validation. Unlike DIC, WAIC is invariant to parametrization and even works for singular models [91]. The WAIC is defined as $\sum_{i=1}^{I}\left(\frac{1}{M}\sum_{m=1}^{M}\boldsymbol{f(y_{ij})}|\boldsymbol{\theta}^{(m)}\right) - \sum_{i=1}^{I} Var_{post}\left(log\boldsymbol{f(y_{ij})}|\boldsymbol{\theta}^{(m)}\right)$, where M is total sampling times. A small WAIC value denotes a better model.

3.4 Simulation

We conduct simulation studies to investigate the identifiability and performance of the proposed model. In longitudinal part, we simulate both continuous and ordinal outcomes, we generate eight ordinal responses, each has 5 levels, and three continuous outcomes. All these responses are predicted by disease status from two domains. Two cause-specific terminal events are generated. Specifically, we have $\theta_i^{(p)}(t) = \beta_0^{(p)} + \beta_1^{(p)} age_i + f^{(p)}(t) + u_{i0}^{(p)} + u_{i1}^{(p)} t + e_i^{(p)}(t)$, and the survival part as $\lambda_{gi}(t) = \lambda_{0g}(t) exp[\boldsymbol{\nu_g'\theta(t)}]$, where $g = 1, 2$ and $p = 1, 2$ denoting two competing

risks and two latent domains. We simulate datasets for 1300 subjects. Each subject could have 17 longitudinal observations in maximum. For two competing events we set the parameters for hazards as $\lambda_{01} = 0.010,\ \lambda_{02} = 0.008$, and linking parameters as $\boldsymbol{\nu_1} = (0.65, 0.2)'$, while $\boldsymbol{\nu_2} = (0.3, 0.5)'$. The censoring time is generated from exponential distribution with mean 50 in additional to the administrating censoring time 10. The approximate censoring rate is 35% of total events. The latent traits are simulated with $\beta_0 = (1.5, 1)'$, $\beta_1 = (-0.4, 0.2)'$. and the random errors from $N(0, 0.64)$ and $N(0, 0.36)$. The heterogeneous disease progression is modeled using $f^{(1)}(t) = t^{1.5}/5$ and $f^{(2)}(t) = log(1 + 2t)/2$, one concave and one convex function. The three continuous longitudinal responses follow normal distributions with variance at $(4, 16, 25)$. We set the random effects vector's distribution as $\boldsymbol{U}_i \overset{iid}{\sim} N_4(0, \boldsymbol{\Sigma})$ where the covariance matrix $\boldsymbol{\Sigma} = \{(1, 0.04, 0, 0)', (0.04, 0.01, 0, 0)', (0, 0, 0.64, -0.008)', (0, 0, -0.008, 0.01)'\}$. The simulation is conducted using Bayesian approaches via Markov Chain Monte Carlo, and implemented in `STAN`. We run 240 replications with two chains, each has 2700 iterations with 1700 burn in.

Table 3.41 presents the simulation results for the parametric coefficients, random effects and association parameters. The biases (the average of posterior means minus the true values) are small, and empirical coverage probabilities are around 95%. The estimates of ordinal parameters and continuous parameters are presented in Appendix B.01.

Table 3.41: Results of semiparametric multidimensional latent trait model in simulation setting.

	Domain 1					Domain 2				
	EST	BIAS	SE	SD	CP	EST	BIAS	SE	SD	CP
Latent variables										
β_0	1.495	−0.005	0.044	0.044	0.938	1.000	0.000	0.036	0.036	0.938
β_1	−0.400	0.000	0.034	0.033	0.924	0.201	0.001	0.025	0.025	0.952
ϵ_e	0.635	−0.005	0.040	0.036	0.910	0.358	−0.002	0.025	0.023	0.914
Random effects										
σ_0	0.999	−0.001	0.066	0.060	0.933	0.635	−0.005	0.042	0.040	0.933
σ_1	0.010	0.000	0.002	0.002	0.957	0.010	0.000	0.001	0.001	0.967
Cause-specific										
ν_1	0.659	0.009	0.044	0.044	0.957	0.197	−0.003	0.053	0.055	0.962
ν_2	0.297	−0.003	0.052	0.051	0.962	0.504	0.004	0.065	0.066	0.957

The nonparametric part results are presented in Figure 3.41. In the plots, the solid line is the true function, the middle dotted line is the mean of estimated function over the 200 replicates, the two boundary lines are the 95% confidence bands.

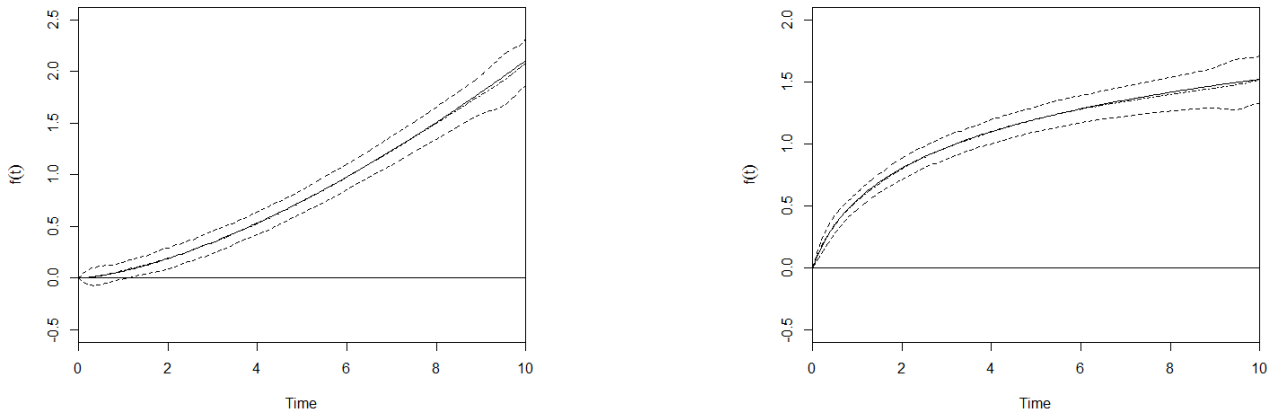

Figure 3.41: The joint model's estimates of nonparametric functions in simulation setting, solid: true functions, dash: mean estimated functions, dot dash: 95% point-wise credible intervals. Left panel: Domain 1. Right Panel Domain 2.

In addition, we present the point-wise coverage plots in Figure 3.42. The empirical coverage probabilities are calculated at each equally spaced grid points. The coverage probabilities are around 95%.

In summary, the biases of the parameter estimates for the nonparametric function and parametric coefficients are small. The confidence intervals based on the proposed standard error estimates have appropriate coverage probabilities. The proposed semiparametric MLTLMM joint competing risk model is identifiable and provides satisfactory performance.

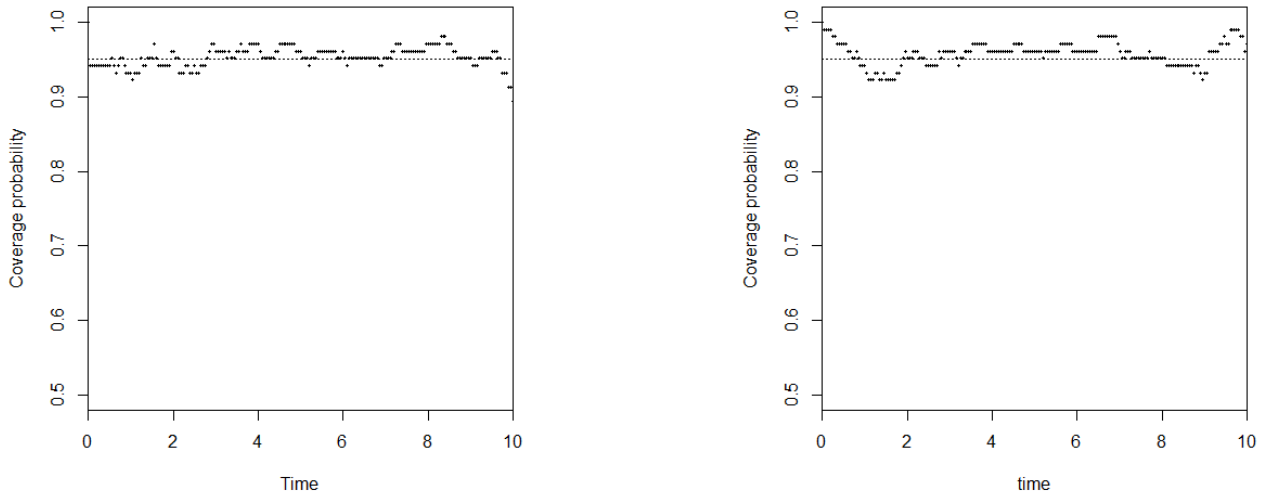

Figure 3.42: Point-wise coverage probabilities with reference lines (dotted horizontal lines). Left panel: coverage probabilities for $f^{(1)}(t)$. Right panel: coverage probabilities for $f^{(2)}(t)$.

3.5 Application to the PPMI study

We applied our semiparametric multidimensional time-dependent latent trait model to PPMI study. The data used in this study were downloaded on Nov. 28, 2017. We use the first three parts of MDS-UPDRS as ordinal responses. In addition, we use Symbol Digit Modalities test score (SDM) as the continuous outcome. SDM score serves as a cognitive measure in PD studies [1], with large value reflecting better clinical outcomes in the original scale (we keep the original scale). Overall, we include 59 items, each has 5 levels responses in longitudinal part. In addition, one continuous response (SDM) is incorporated. Each individual could have two types of endpoints, one is dropout, the other one is ST. These two types of endpoints compete each other, or either one can censor another type.

All questionnaires in MDS-UPDRS (part I, II, and III) are structured (grouped) to collect the domain-specific information manifested by those ordinal responses in each part. We

refine our models to accommodate these structured responses, and adjust latent traits model with the updated assumption that item responses in each of the three predefined parts manifest the unobserved status of corresponding disease domain. We add dependency across domains by incorporating fully correlated random effects. Because of data structure, we only incorporate between-item multidimensionality based on MDS-UPDRS sub-scale parts. While modeling two cause-specific failure risks and the longitudinal continuous outcomes, we incorporate within-domain multidimensionality in survival models by cross loading these three-dimension disease status and progression (each outcome captures the effects of these three latent traits). For simplicity, we use identity form of $\phi(\theta)$ in cox models. Because the model for ordinal part is over-parameterized, we impose identifiability constraints on one selected item in each domain, so that these three items from different domains construct the basis of three dimensions, and calibrate each latent variable. We use factor analysis to select these three items. In each MDS-UPDRS part, we conduct factor analysis to select the item which has the largest factor loading. After confirmatory factor analysis, we choose item 3 (Depressed mood) from nM-EDL part, item 4 (Eating tasks) from M-EDL and item 11 (Hand movement) from Motor Examination to put constraints. For these selected items, we set $(a_{3,1}^{(1)}, a_{4,1}^{(2)}, a_{11,1}^{(3)}) = (1.5, 0.8, -0.4)$ based on observed frequency of the aforementioned item's first level. Besides we preassign all $b_k = 1$.

We compare the different covariate combinations in latent trait model. The general latent trait model is $\theta_i(t)^{(p)} = \boldsymbol{X}^{(\boldsymbol{p})}\boldsymbol{\beta}^{(\boldsymbol{p})} + f^{(p)}(t) + u_{i0}^{(p)} + u_{i1}^{(p)}t + e_i(t)^{(p)}$. We conduct model selection for different fixed covariate $\boldsymbol{X}$ settings, model 1 includes one covariate in latent trait, model 2 has two covariates, while model 3 incorporates three covariates in latent trait model. The fixed covariates to be considered include age, gender, disease duration and Hoehn and Yahr Scale (HY) score. The DIC and WAIC of different covariate combinations are shown in Appendix B.03. The model with one covariate in latent trait has both the lowest DIC and WAIC among all the models. We then use cross loading to incorporate all three domains in competing risk model based on the optimal latent trait model. The estimate results with three

domains are displayed in Appendix Table B.04. The estimated association coefficients show that only disease status and development in M-EDL significantly affect the occurrence of the events. We run reduced model, keeping the domain which has significant association with the terminal events. Table 3.51 displays the estimates of coefficients. We find significant age effects in nM-EDL domain and M-EDL domain. Several studies addressed role of age in PD severity with mixed results [54]. In this study, we find domain-specific age effects, with significance in nM-EDL and M-EDL. The results from competing sub-model show both terminal events are significantly associated with the decomposed disease status and development in M-EDL.

Table 3.51: Posterior estimations for reduced model.

	nM-EDL				M-EDL				Motor examination			
	Mean	SD	95% CI		Mean	SD	95% CI		Mean	SD	95% CI	
Disease Status												
Int.	−0.036	0.088	−0.242	0.116	−0.312	0.104	−0.516	−0.124	−0.180	0.207	−0.598	0.227
Age (yr)	0.105	0.053	0.007	0.219	0.126	0.070	0.001	0.261	0.254	0.156	−0.016	0.559
Competing risks' Linking Parameters												
Dropout	-	-	-	-	0.258	0.054	0.151	0.367	-	-	-	-
ST	-	-	-	-	0.326	0.110	0.119	0.536	-	-	-	-

The Figure 3.51 displays domain-specific disease progression. The identified time-dependent disease progression are all significantly progressing to worse status in three domains (justified by uncertainty bands). The progression of disease varies across domains (varied development trend). Specifically, disease in nM-EDL has approximately linear trend. The disease in M-EDL has one accelerating stage, which happens between year 2 to year 4, while the disease in motor examination also has one accelerating stage between year 2 to year 3.5. These new findings and the aforementioned finding of domain specific age effects signify the unique advantage of our semiparametric multidimensional latent trait framework: ability to capture the domain-specific disease continuous progression and the impaired covariate effects across domains.

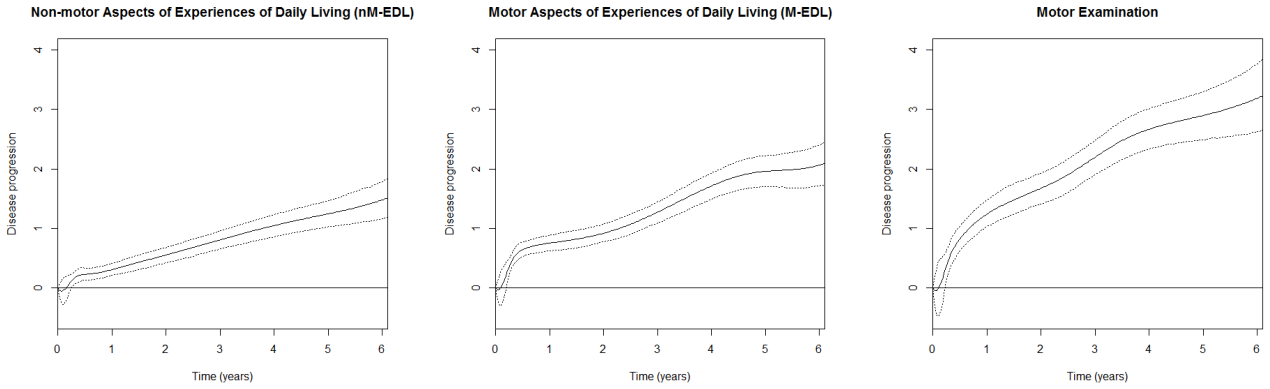

Figure 3.51: Time-dependent disease progression by domains. Dot lines: 95% confident band, Solid lines: mean estimate temporal function. Left panel: nM-EDL . Middle panel: M-EDL. Right panel: motor examination.

To interpret the clinical information carried in domain-specific latent disease measures, we refer to the estimated coefficients in continuous outcomes. Appendix Table B.05 shows the dimensional latent variable effects on SDM. One unit increase in nM-EDL disease status (worsening) associates with 1.248 (Mean=-1.248, CI: $[-1.936, -0.575]$, negative sign for worsening) units decrease in SDM (worsening) while keeping other two sub-domain disease status unchanged; One unit increase in M-EDL disease status (worsening) associates with 1.040 (Mean=-1.040, CI: $[-1.600, -0.475]$) units decrease in SDM (worsening) while other two sub-domain disease status unchanged; One unit increase in Motor Examination (worsening) is associating with 0.190 (Mean=-0.190, CI: $[-0.374, 0.013]$) units decrease in SDM (worsening) while holding other two sub-domain disease status unchanged. Overall, the domain-specific disease statuses have different manifestation on SDM. The influence on SDM from nM-EDL is larger than that from M-EDL, while the influence from Motor Examination is not significant. Our finding provides evidence that this cognitive measure does manifest more information from non-motor domain than from motor domain.

Random effects related posterior parameters are presented in Appendix Tables B.06 and B.07. Every individual has 3 pairs of random effect terms, denoting the 3 dimensional random

effects, each domain has its own domain-specific random intercept and random slope. For each vector of random effects, the 1st and 2nd elements are for nM-EDL domain, the 3rd and 4th are for domain in M-EDL, while the last two are for motor examination domain. The correlation matrix (Appendix Table B.07) provides additional information of inter-dependency of diseases and correlated (in high dimension) subject-specific characteristics between and within domains. The significant within-domain correlations (random intercept vs random slope in same domain) reside in nM-EDL ($\rho_{01} = -0.204$, 95% CI: $[-0.341, -0.043]$), in M-EDL ($\rho_{01} = -0.168$, 95% CI: $[-0.311, -0.028]$) and in motor examination ($\rho_{01} = -0.426$, 95% CI: $[-0.544, -0.295]$), explained as: if a participant's nM-EDL disease status is worse at the start of the study, his or her nM-EDL's disease progression is slower during follow-up. Same interpretation is for disease in other two domains. The intercepts between-domain correlations are all positive and significant (random intercept vs random intercept, $\rho = (0.625, 0.194, 0.375)$ for nM-EDL vs M-EDL, nM-EDL vs Motor and M-EDL vs Motor respectively), these associations reveal that in the study if any initial disease status is worse, the other two are also worse, for example, the initial disease status in nM-EDL is worse, the disease severities in M-EDL and motor examination are also worse. Same finding is identified in the correlations of between-domain random slope (random slope vs random slope, and $\rho = (0.627, 0.200, 0.639)$ for nM-EDL vs M-EDL, nM-EDL vs Motor and M-EDL vs Motor respectively), if disease progresses fast in any domain, it also develops fast in other domains, for example, if disease develops fast in M-EDL, it also deteriorates fast in nM-EDL and motor examination.

The parameter estimates for all 59 items are displayed in Appendix Tables B.09-B.011, and all parameters are significant.

3.6 Discussion

In this study, we address many important features in PD. Our model incorporate the domain-specific covariate effects, time-dependent disease progression and multiple terminal events. There are limited studies addressed these issues all together. In this modeling framework, we use multidimensional latent trait model to identify the impaired covariate effects across domain based on multivariate and mixed type data. In addition, we use nonparametric approach to the domain-specific heterogeneous disease progression (temporal effects) which is one of primary objective of PPMI study. In the end, we simultaneously analyze the longitudinal outcomes and terminal events of multiple types, under the impacts of impaired disease status and heterogeneous disease progression across domains.

Our multidimensional latent variable model are capable to incorporate the domain-specific disease information, obtain the impaired information originated in heterogeneous covariate effects and disease progression across domains. We extend our methods to survival sub-model and assess the association intensities on terminal event in domain level. The competing risk sub-models enable researchers to handle dependent censoring. Moreover, it provides simultaneous inference on both longitudinal observations and survival endpoints for multi-domain diseases, which are typically characterized with impairment and heterogeneity both in domains and time. This is the first study in PD to quantify the impacts of high dimensional disease development in motor and non-motor on terminal events of multiple types. We propose a Bayesian framework for joint modeling of longitudinal and competing events based on time-dependent disease progression. The proposed Bayesian method has appealing features, it can easily handle multi-dimensional (domain-specific) disease status, time-dependent progression which are difficult under the frequentist framework.

Our study has some limitations. We assume multivariate normal distribution of random

effects in the models. Robustness against the departure from the normality assumption were studied by several researchers [85, 107]. There are other methods to model random effects, we will investigate the other forms of random effects in future study, such as mixture normal distributions based on Dirichlet process or Laplace distribution.

Chapter 4

Bayesian joint analysis for longitudinal studies with nonignorable missing data and applications to Parkinson's disease study

4.1 Introduction

Parkinson's disease (PD) is one of the common neurodegenerative disorders [75]. It is diagnosed in about 1% of individuals over the age of 60 worldwide [6]. PD is an incurable, complex and heterogeneous progressive disorder that gradually robs the individual of motor control. It is now considered as a systemic disease due to its non-motor symptoms in addition to the motor symptoms [11].

There is no validated biomarker for PD [66], the diagnosis method mainly depends on clinical information provided by patients, for example self-reported motor sign and symptoms (rigidity, tremor, etc.). The Unified Parkinson's Disease Rating Scale (UPDRS) is one widely used scale for clinical ratings of PD [76]. There are a lot of PD studies and clinical trials using UPDRS to follow the longitudinal course of PD [11, 88, 89]. In 2007, a revised UPDRS, the Movement Disorder Society-UPDRS (MDS-UPDRS) was introduced to provide more compre-

hensive and accurate tests than the original UPDRS [35]. The MDS-UPDRS consists of 65 items, all items are anchored with five categories, from 0 to 4, the higher the score the worse the status. The first three parts of MDS-UPDRS are commonly used in PD studies, which include 59 items. Statistical methods are required to extract useful information to define disease and its progression based on these 59 item responses.

The ordinal response in MDS-UPDRS are often summed up to obtain total score, which is treated as a continuous outcome. It is easy to implement but leads to loss of information by ignoring differences between item pattens [36]. Alternatively, multilevel item response theory (MLIRT) model was utilized to analyze the longitudinal scores. This model links the multiple items to the unobserved disease status structured as a univariate latent variable. However, the unidimensional framework limits the application of the model in analyzing PD due to the complication of disease such as the impairment and heterogeneity. PD is characterized by existing impairment across domains, for instance, non-motor symptoms often occur many years before the clinical motor signs[11]. To address these issues in the traditional models, multidimensional item response model was introduced [72]. Though cross-sectional impairment was addressed in this model, the longitudinal impairment information and correlations are not able to assessed in this cross-sectional multidimensional model. Recently, Wang and Luo [98] proposed a new multidimensional latent trait linear mixed model (MLTLMM) to address the disease impairment in longitudinal study. This new multidimensional latent trait model allows multiple latent variables and within-item multidimensionality. By adopting latent disease score to reduce the number of observed outcomes, MLTLMM is more computational scalable than multivariate marginal and random effects models.

In PD studies, participants are monitored longitudinally with respected to the aforementioned dozens ordinal outcomes plus other outcomes. During the follow-up, outcomes to be collected can be missing due to subjects' non-response, missed visits, dropout and etc. Rubin [80] defined three missing data mechanisms. If the missingness is independent of the observed

and unobserved data, this missing data mechanism is missing completely at random (MCAR). When missingness is not dependent on unobserved data, it is missing at random (MAR). The missing data belonging to these two missing data mechanisms are treated as 'ignorable' missingness, which do not cause bias in statistical inference for likelihood-based estimation. In this study, we use MAR to denote the ignorable missingness, as MCAR is rare and can be handled in same way as MAR in analysis. However, when missingness is associated with the unobserved underlying response process, this missingness is missing not at random (MNAR). For example, patients' dropouts are due to worsening of disease or death. MNAR mechanisms are 'nonignorable'. Under the MNAR assumption, the missing data mechanism needs to be modeled simultaneously with the outcome variables to avoid biased parameter estimates [18].

In addition to these missing data mechanisms, there exist two patterns of missing data. The first missing data pattern is the 'intermittent missing data' or non-monotone missing data, for example, an individual may miss some visits before the last visit. While the other pattern is 'monotone missing data', denoting the data with the pattern that an individual leaves the study and never returns, or the observations are completely disrupted by some events (e.g. dropout or initiation of symptomatic treatment). Generally, the aforementioned two missing data patterns could have varied missing data mechanisms, which are difficult to justify whether the missing data mechanisms are MAR or MNAR. The possibility of missing data being MNAR can hardly be ruled out.

Estimating parameters with nonignorable missing data is more complex than with ignorable missing data. Recently, modeling longitudinal observations with nonignorable missing data has drawn much attention [56, 57, 106]. Many models are proposed, these models can be classified into three types: selection model, pattern-mixture model and shared-parameter model [57]. The selection approach combines the hypothetical complete data together with the missing data process based on likelihood. The pattern-mixture approach models the distribution of the data conditional on the missing data pattern. While the shared-parameter

approach incorporates the dependence between measurements and missingness processes by the means of random effects, it can be extended to latent variable models. Molenberghs *et al.* discussed a selection model for longitudinal ordinal data with nonrandom dropout [68]. Ekholm and Skinner proposed a pattern-mixture model for a longitudinal binary incomplete data set [22]. The full likelihood approach has been used to specify the joint likelihood of outcomes and missing indicators when handling nonmonotone pattern of missing data [44]. For example, the random-coefficient-based selection models were adopted to link dropout time to the longitudinal outcomes through individual random effects [16, 79, 104]. Alternatively, pseudo likelihood was proposed to provide statistical solutions [90]. Elashoff *et al.* [24] developed the latent random effects model to incorporate effects from nonignorable monotone missing data. Most statistical models focus on one missing data pattern (either monotone or non-monotone missing). Besides, those models are based on one or two outcomes and use the latent traits as predictors for monotone missing and other missing observations. Wu *et al.* [103] proposed a nonlinear mixed-effects model for both monotone and non-monotone patterns of missing data. In PD study, the missed visits are frequently happened during the long follow-up, both patterns of missing data exist in study. How to address the missed responses which consist of dozens ordinal response is an open problem in PD study.

In this study, we present a generalized approach to the longitudinal data in the presence of two missing data patterns with both missing data mechanisms. We extend multidimensional latent trait methods to model and test missing data mechanisms based on the responses from dozens of ordinal outcome. We jointly analyze the data without excluding MNAR assumption, and assess missing data mechanisms under the impacts from heterogeneous disease development in multiple domains. This is the first study addressing the multiple ordinal responses (59 ordinal responses) carrying impairment information from multiple domains, and in presence of missing data with different missingness patterns. The remainder of this article proceeds as follows. In section 2, we describe motivating study and the data. Section 3 discusses the proposed model,

and Bayesian inference. Section 4 presents studies to assess the performance of the proposed models. In section 5, we apply our method to the motivating studies. Section 6 provides concluding remarks and discussion.

4.2 Motivating clinical studies

This methodological development is motivated by Parkinson's Progression Markers Initiative (PPMI) study. PPMI is an ongoing longitudinal observational study that aims to identify one or more markers of progression for Parkinson's disease (PD). All participants were grouped into several cohorts. At baseline, patients were not expected to require PD medications within at least 6 months. In PPMI study, MDS-UPDRS scale is used to assess the disease status and progression. In clinical trial, the first three parts of MDS-UPDRS (59 items) are commonly used in study. According to the Movement Disorder Society (MDS), the 13 items in Part I are intended to measure the non-motor aspects of experience of daily living (nM-EDL), the 13 items in Part II are used to measure motor aspects of experiences of daily living (M-EDL), while the 18 grouped items (total 33 items, several with right, left or other body parts' sub-items) in Part III focus on the information from motor examination. Items of the different subscales are assumed to be manifestation of different disease domains (motor, cognitive and behavior). MDS-UPDRS provides a relatively good measure to follow and define PD progression. Fitting MDS-UPDRS longitudinal observations into multidimensional latent trait model, we are able to characterize the natural disease progression of PD patients in PPMI study. In the study, when a participant fails to answer one or more items out of total 59 items, those partial or complete non-responses constitute missing data. Although PD studies are designed to collect complete data on all participants, missing data commonly happen and impact the analysis results.

Besides, during the long course of follow up, the repeated observations are subject to the risks of endpoints, the follow-up of patients might be terminated long before the end of study

for different reasons, which can be treated as aforementioned monotone pattern of missing data. First, the longitudinal observations can be stopped because of dependent censoring (e.g., dropout, death), in addition, the symptomatic therapy (ST) can cause the repeated observations being interrupted (the following observations after ST will not reflect the natural PD progression). In PD studies, ST is generally being treated as one type of endpoint [84]. These events are likely informative for disease progression and status. When modeling these endpoints, we are making statistical inference on whether these events associated with disease status or not. Generally, we are not able to exclude the possibilities that these events as disease-related or the monotone pattern of missing data belongs to MNAR. Figure 4.21 uses sum score of MDS-UPDRS part III to illustrate the two missing data patterns. Patient 3 had the last visit at end of year 2 (monotone missing pattern), while patient 299 has three intermittent missed visits (intermittent missing pattern). Indeed, we do not need an explicit model for the probabilities of missingness if missing data are MAR. However, we can not just conduct analysis based on MAR or MCAR assumption, and we need a statistical framework to test these hypothesis, and model the missing data simultaneously with outcome variables to avoid biased parameter estimates [18].

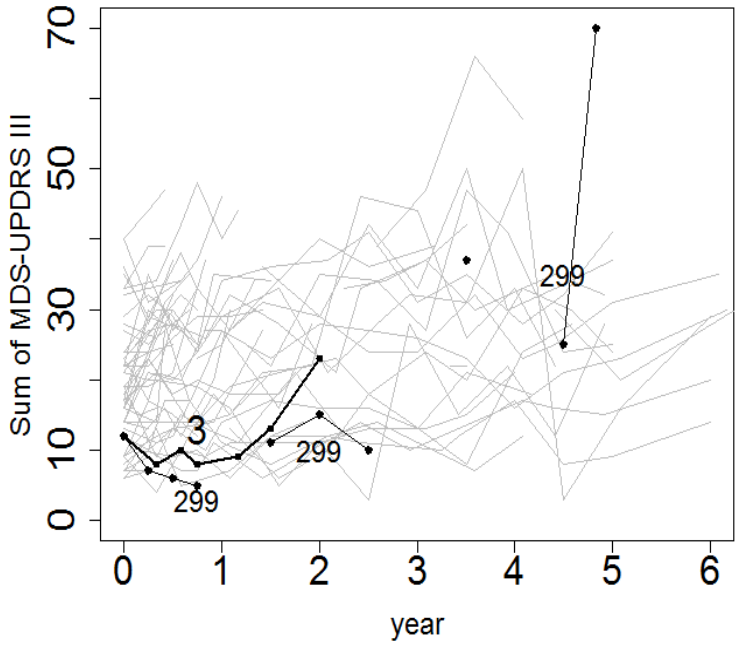

Figure 4.21: 50 randomly selected patients from PPMI study. Patient 299 had three missed visits before the last visit (intermittent pattern of missing data), patient 3 dropped out at end of year two (monotone pattern of missing data).

4.3 Model and estimation

4.3.1 Latent trait model

Let $y_{ik}(t)$ be the observed outcome k from ith subject at time t, where $i = 1, \ldots, N$, $k = 1, \ldots, K$, and $t = t_{i1}, \ldots, t_{iJ_i}$. All outcomes are coded so that larger values are worse clinical conditions.

To start building the MLTLMM modeling framework, we assume that there are P (with $P < K$) latent variables (LVs) representing the underlying disease severity scores and denote them as $\boldsymbol{\theta}_i(t) = (\theta_i^{(1)}(t), \ldots, \theta_i^{(p)}(t), \ldots, \theta_i^{(P)}(t))'$ for subject i at time t, where the superscript ($p =$

$1, \ldots, P$) denotes the pth latent variable. From a clinical perspective, each latent variable denotes the disease severity of a PD domain (e.g., non-motor and motor). We introduce the linear model for continuous outcomes, logistic model for binary outcomes, and ordinal logistic model for ordinal responses.

$$y_{ik}(t) = a_k + \boldsymbol{b_k'\theta_i(t)} + \varepsilon_{ik}(t), \tag{4.1}$$

$$\text{logit}\big\{p(y_{ik}(t) = 1|\theta_i(t))\big\} = a_k + \boldsymbol{b_k'\theta_i(t)}, \tag{4.2}$$

$$\text{logit}\big\{p(y_{ik}(t) \leq l|\theta_i(t))\big\} = a_{kl} - \boldsymbol{b_k'\theta_i(t)}, \tag{4.3}$$

where a_k and $\boldsymbol{b_k}$ are the outcome-specific parameters, while the random errors $\varepsilon_{ik} \sim N(0, \sigma_{\epsilon_k})$, are independent and identically distributed. Note that for continuous outcome, $a_k = E[y_{ik}(t)|\boldsymbol{\theta}_i(t) = \mathbf{0}]$ is the mean of the kth outcome if the disease severity scores are 0. The parameter $\boldsymbol{b_k}$ also plays the role of bringing up disease severity score to the scale of the kth outcome. The negative sign for $\boldsymbol{b_k}$ in the ordinal outcome model is to ensure that worse disease severity (higher $\theta_i(t)$) is associated with more severe outcomes (higher $y_{ik}(t)$). For ordinal responses, the probability of being in a particular category is $p(y_{ik}(t) = l) = p(y_{ik}(t) \leq l|\boldsymbol{\theta_i(t)}) - p(y_{ik}(t) \leq l-1|\boldsymbol{\theta_i(t)})$, where $l = 1, 2, \ldots, n_k - 1$ is the lth level of the kth random variable, which is ordinal with n_k levels. Interpretation of parameters is similar with continuous outcomes, except that modeling is on the log-odds, not the native scale of the data. Because the ordinal model is over-parameterized, additional constraints are required to make model identifiable. Using this model, we are able to explicitly combine information from all outcomes, specifically those dozens of ordinal outcomes. This is one of the simplest ways to conceptualize the disease severity scores that allows to define the disease status and progression when there is no gold standard. To model the severity scores $\boldsymbol{\theta_i(t)}$, we propose the second level multivariate linear mixed model,

$$\theta_i^{(p)}(t) = \boldsymbol{X}_i^{(p)}(t)\boldsymbol{\beta}^{(p)} + \boldsymbol{Z}_i^{(p)}(t)\boldsymbol{u}_i^{(p)} + e_i^{(p)}(t), \tag{4.4}$$

where $\boldsymbol{X}_i^{(p)}(t)$ and $\boldsymbol{Z}_i^{(p)}(t)$ are the covariates corresponding to fixed and random effects respectively, latent variable $\theta_i^{(p)}(t)$ denotes ith subject's unobserved disease severity in the pth domain at time t. The latent variables are continuous, higher value indicating worse disease severity. The vector $\boldsymbol{u}_i = (\boldsymbol{u}_i^{(1)'}, \ldots, \boldsymbol{u}_i^{(P)'})'$ contains the random effects for the ith subject, it follows a multidimensional normal distribution, $\boldsymbol{N(0, \Sigma)}$, where $\boldsymbol{\Sigma}$ is the covariance matrix with dimension equal to the number of random effects incorporated. There are several ways to model random effects. For example, when we incorporate fully correlated random intercepts and random slopes in framework (4.4), this covariance matrix will have the dimension of $2p \times 2p$. The residual term $e_i^{(p)}(t)$ is assumed to be mutually independent, and $e_i^{(p)}(t) \sim N(0, \sigma_e^{(p)})$.

For notational convenience, we let $\boldsymbol{a} = (\boldsymbol{a}_1', \ldots, \boldsymbol{a}_k', \ldots, \boldsymbol{a}_K')'$, and $\boldsymbol{a}_k = (a_{k,1}, \ldots, a_{k,n_k-1})'$ for the kth ordinal outcome with n_k categories. We let $\boldsymbol{b} = \boldsymbol{(b_1, \ldots, b_K)'}$, a K by P matrix, where $\boldsymbol{b_k} = (b_k^{(1)}, \ldots, b_k^{(p)})'$. Because the model is over-parameterized, additional constraints are required to make it identifiable. The indeterminacy between the latent variable loadings $\boldsymbol{b}_k$ and the scales of the latent variables $\boldsymbol{\theta}_i(t)$ can be fixed by either setting one element in each column of $\boldsymbol{b}$ to be 1, or letting $\sigma_e^{(p)} = 1$ for $p = 1, \ldots, P$ with at least one of the loadings constrained to be positive for each factor [20]. Finally, to identify parameters $\boldsymbol{a}$ and intercepts β in regression coefficients, we set the constraints on one selected item in each domain, we let $a_{p,1} = 0$ (or other constant) for $p = 1, \ldots, P$ ordinal outcomes and the order constraint $a_{k,1} < \ldots < a_{k,l} < \ldots < a_{k,n_{k-1}}$ must be satisfied. Besides, we set identifiability constraints on p vectors $\boldsymbol{b}$, for example, when $P = 3$ and the constraints are put on the first three items, we let $b_1^{(1)} = b_2^{(2)} = b_3^{(3)} = 1$, all other elements are 0. In real data analysis, in order to achieve the better domain calibration and locate the three presumptive optimal bases, we have to carefully select the item to put constraints for each domain.

4.3.2 Model for monotone missing data

For the monotone missing data, we use proportional hazard model to incorporate this missing data pattern. Here, we use y_{ikj} to denote the jth scheduled outcome k from subject i. When y_{ikj} belongs to monotone missing data, y_{ikq} is missing for all $q \geq j$, or it is the first missing observation for following consecutive missed visits. We use δ_i to code this missing data pattern, let $C_i = (T_i, \delta_i)$ be the endpoint observation for ith subject, where $\delta_i = 0$ indicating the latest visit and the following visits are still on going, $\delta_i = 1$ for event (either dropout or ST) and no following observation. To quantify the effect of $\theta_i(t_j)$ on the risks for events, we build the Cox model as,

$$\begin{aligned} \lambda_i(t) &= \lim_{h \to 0} \frac{P[t \leq T_i < t+h | T_i \geq t, \theta_i, \nu, \boldsymbol{X_i(t)}, \boldsymbol{\gamma}]}{h}, \\ &= \lambda_0(t) exp\{\boldsymbol{\gamma' W_i} + \boldsymbol{\nu' \theta_i(t)}\}, \end{aligned} \tag{4.5}$$

where $\lambda_0(t)$ is the baseline hazard, $\boldsymbol{W_i}$ is a vector of fixed effect covariates, $\boldsymbol{\gamma}$ is the vector of regression coefficients. The regression coefficient $\boldsymbol{\nu'} = (\nu^{(1)}, \ldots, \nu^{(P)})$ links event times and the domain-specific disease status. In this model, $\lambda_i(t)$ is the instantaneous failure rate at time t given the covariates $\boldsymbol{W_i}$ and latent traits $\boldsymbol{\theta_i}$. More precisely, if $\boldsymbol{\nu} = \mathbf{0}$, this monotone missingness is not disease related, categorized as MAR.

4.3.3 Model for intermittent missing data

To incorporate intermittent missingness, we use a mixed effect logistic regression model to model the conditional probability of missed visit. Intermittent missing observations for y_{ikj} are those y_{ikj} being missing, while y_{ikq} is observed for at least one $q > j$. We use intermittent missing indicator $r_{ij} = 1$ to code this missed visit. We now obtain the augmented data (y_{ij}, r_{ij}), where $r_{ij} = 1$ or 0, denoting whether ith subject's jth visit is missing or fully recorded respec-

tively. For example, the outcome y_{ikj} ($j < J_i$, J_i is the last visit) is missing for $k = 1, \ldots, K$, when $r_{ij} = 1$, in this setting $r_{iJ_i} = 0$ or y_{ikJ_i} (the last) must be observed.

$$logit(P(r_{ij} = 1|\theta_{ij})) = \boldsymbol{\alpha'} \boldsymbol{X_i} + \boldsymbol{\eta'} \boldsymbol{\theta_i}(\boldsymbol{t_j}), \tag{4.6}$$

where $\boldsymbol{X_i}$ is the vector of covariates and $\boldsymbol{\alpha}$ is the corresponding vector of regression coefficients. The parameter $\boldsymbol{\eta'} = (\eta^{(1)}, \ldots, \eta^{(P)})$ governs the association between the intermittent missing data process and the domain-specific disease severity process modeled by latent variable. Both $y_i(t_j)$ and r_{ij} are censored by censoring time T_i or monotone missingness. Moreover, the parameter $\boldsymbol{\eta}$ plays the role of sensitivity parameters to test MAR assumption for intermittent missing. When $\boldsymbol{\eta} = 0$ the missing data mechanism is MAR, otherwise it is MNAR, this is because the parameter can be used to test whether missingness (modeled missed visit probability) depends on unobserved data or not.

4.3.4 Likelihood

By combining all the sub-models, the joint missing data models incorporate two missingness patterns into frameworks without excluding nonignorable mechanism assumption. We provide statistical test for MAR. If both ν and η are zero (based on statistical tests) we can not reject MAR, or misingness will be independent of unobserved disease status, in this scenario the missing data are ignorable, both intermittent sub-model and cox sub-model (for monotone misingness) can be removed from the full likelihood. If just one of them is zero, the corresponding sub-model can be removed from the full likelihood.

For intermittent missing, the conditional likelihood function for the ith subject given parameters, covariates and random effects $\boldsymbol{X}, \boldsymbol{U}$ is $L_i = f(\boldsymbol{y_i}|\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{\Theta}) f(\boldsymbol{r_i}|\boldsymbol{X}, \boldsymbol{U}, \boldsymbol{\Theta})$. The validity of likelihood is based on the assumption of independence of $\boldsymbol{y_i}$ and $\boldsymbol{r_i}$ given the latent

variable (or random effect). The combined likelihood for ith subject by incorporating both monotone missing data and non-monotone missing data is,

$$L_i \ = f(\boldsymbol{y_i}|\boldsymbol{X},\boldsymbol{U},\boldsymbol{\Theta})f(\boldsymbol{r_i}|\boldsymbol{X},\boldsymbol{U},\boldsymbol{\Theta})f(T_i,\delta_i|\boldsymbol{X},\boldsymbol{U},\boldsymbol{\Theta})f(\boldsymbol{U}). \tag{4.7}$$

4.3.5 Bayesian inference

To make inference on the parameter vector $\boldsymbol{\Theta}$, we use Bayesian methods based on Markov chain Monte Carlo (MCMC) posterior simulations. We use vague priors on all elements in $\boldsymbol{\Theta}$, except for the aforementioned constrained parameters, i.e., $a_{1,1} = 0$ (or other constant) , and $b_1^{(p)} = 1$, for the selected item to ensure the item response model identifiable. To obtain the prior distributions for the threshold parameters of ordinal outcome k, we let $a_{1,1} \sim N(0, 20)$, and $a_{k,l} = a_{k,l-1} + \Delta_l$ for $l = 2, \ldots, n_k - 1$, with $\Delta_l \sim N(0, 10000)I(> 0)$ (SD=100), i.e., normal distribution left truncated at 0. The setting of high SD for the following difficulty parameters is considering the scenarios that there are rare responses to some top level of items (some items have few responses for the level 5). Prior distributions for unconstrained elements in $\boldsymbol{b}$ and $\boldsymbol{\beta}$ are from $N(0, 20)$. We use the Cholesky factorization to estimate the correlation coefficients, the random effects covariance matrix is expressed as $\boldsymbol{\Sigma} = \boldsymbol{\sigma_u'}\boldsymbol{\Sigma_U}\boldsymbol{\sigma_u}$, where Σ_U is the correlation matrix. All variances are from Inverse-Gamma$(0.01, 0.01)$. We have investigated other selections of vague prior distributions with various hyper-parameters and obtained very similar results.

The posterior samples are obtained from the full conditional likelihood of each unknown parameter using Hamiltonian Monte Carlo (HMC) [19] and No-U-Turn Sampler (NUTS) [42]. Both HMC and NUTS samplers are implemented in `Stan` (version 2.17.0) [87], which is a probabilistic programming language implementing statistical inference. For large datasets, `Stan` may be more efficient than `BUGS` language [59] in achieving faster convergence and requiring

smaller number of samples [42]. To monitor Markov chain convergence, we use the trace plots and view the absence of apparent trends in the plot as evidence of convergence. In addition, we use the Gelman-Rubin diagnostic to ensure the scale reduction $\hat{R}$ of all parameters are smaller than 1.1 as well as a suite of convergence diagnosis criteria to ensure convergence [32].

4.4 Simulation

We conduct simulation studies to evaluate the proposed method and compare the method with the naive method (ignoring intermittent missing). We generate two continuous outcomes and a series of ordinal responses with 10 items, each item has five levels. To mimic the characteristics of PD study data, we let the outcome responses to be predicted by the latent traits from two domains. These twelve outcomes are longitudinal outcomes. Each response is predicted by two latent variables (cross loading). Here, we use multidimensional latent trait model [98] to model longitudinal observations. Now, the updated latent trait model is $\theta_i^{(p)}(t) = \beta_0^{(p)} + \beta_1^{(p)} X_i + \beta_2^{(p)} t + u_i^{(p)} + e_i^{(p)}(t)$, Cox model is $\lambda_i(t)|\boldsymbol{\theta_i(t)} = \lambda_0(t) exp(\gamma V_i + \boldsymbol{\nu' \theta_i(t)})$, and intermittent missing data model is $logit(P(r_{ij} = 1|\boldsymbol{\theta_i(t)}) = w + \boldsymbol{\eta' \theta_i(t_j)}$, where $p = 1, 2$ denoting two disease domains, the vector $\boldsymbol{u_i'} = (u_i^{(1)}, u_i^{(2)}) \sim N(0, \boldsymbol{\Sigma})$. We simulate 1000 subjects. Each subject could have 16 sequential longitudinal observations in maximum. The latent variables are simulated with $(\beta_0^{(1)}, \beta_1^{(1)}, \beta_1^{(2)}, \beta_0^{(2)}) = (-0.2, 0.2, 0.5, -0.5)$, the time effects $(\beta_2^{(1)}, \beta_2^{(2)}) = (0.4, 0.7)$. The random effects are from $N_2(0, \boldsymbol{\Sigma})$, where diagonal elements of $\boldsymbol{\Sigma}$ (variance part) are $(1, 1.69)$, while the off-diagonal element (covariance) is 0.52, or equivalent to $(\sigma^{(1)}, \sigma^{(2)}, \rho) = (1, 1.69, 0.4)$. We simulate random errors from independent normal distribution $N(0, \epsilon^{(p)})$, while $(\epsilon^{(1)}, \epsilon^{(2)}) = (1, 0.64)$. For monotone missing data, we use Cox model to generate endpoints (dropout). We set the parameters for the covariates as $\gamma = 1$, baseline hazards as $\lambda_0 = 0.006$, association parameter $\boldsymbol{\nu'} = c(0.4, 0.2)$. The censoring time is generated from exponential distribution with mean 50 in additional to the administrative

censoring time 10. About 29% of total events (monotone missingness) are generated. For intermittent missing, we use $logit[P(r_{ij} = 1)] = W + \boldsymbol{\eta\theta_i(t_j)}$ to generate missing indicators. We let $(W, \eta^{(1)}, \eta^{(2)}) = (-4, 0.5, 0.7)$ to generate around 20% intermittent missing data (based on the summaries of 240 datasets). For comparison purpose, we generate another set of data with around 30% intermittent missing by letting $(W, \eta^{(1)}, \eta^{(2)}) = (-3, 0.5, 0.7)$. Taking into account the monotone missing, there are total about 30% missed data in the first setting, and about 40% in the second setting. The other parameter settings for continuous outcomes and ordinal outcomes are presented in Appendix. Total 240 datasets are generated.

In addition, we run naive model, treating missing as ignorable and only use the observed observations (intermittent missing data are ignored) for comparison purpose. The simulation analysis is conducted using a Bayesian approach via MCMC. Two chains are used in each setting, each has 4000 iterations with 3000 burn-in. For each estimated parameter, we compute percent bias as follows, for parameter β_j, percent bias= $100(\hat{\beta}_j - \beta_j)/\beta_j$. The biases is the average over all simulations. The simulation results are presented in the Table 4.41, more results are in Appendix Table C.01.

Table 4.41: Simulation results with different intermittent missing proportions.

	Naive (setting 1)			Joint (setting 1)			Naive (setting 2)			Joint. (setting 2)		
	BIAS %	SD	CP	BIAS %	SD	CP	BIAS %	SD	CP	BIAS %	SD	CP
$\beta_0^{(1)} = -0.2$	13.000	0.045	0.900	0.000	0.044	0.949	31.000	0.044	0.691	0.000	0.046	0.930
$\beta_0^{(2)} = 0.2$	−9.000	0.051	0.927	2.500	0.050	0.927	−24.500	0.049	0.845	2.000	0.052	0.940
$\beta_1^{(1)} = 0.5$	0.200	0.034	0.959	0.800	0.037	0.967	0.400	0.037	0.973	−0.200	0.038	0.973
$\beta_1^{(2)} = -0.5$	−2.000	0.043	0.950	−0.600	0.044	0.953	−2.600	0.044	0.955	−0.600	0.045	0.958
$\beta_2^{(1)} = 0.4$	−4.250	0.008	0.568	−0.250	0.009	0.953	−3.500	0.009	0.677	−0.250	0.009	0.963
$\beta_2^{(2)} = 0.7$	−2.714	0.011	0.677	−0.143	0.013	0.963	−3.000	0.013	0.668	−0.143	0.014	0.963
$\rho = 0.4$	−7.000	0.034	0.873	−0.250	0.033	0.940	−9.000	0.036	0.823	−0.500	0.034	0.953
$\sigma_1^{(1)} = 1$	−5.600	0.065	0.827	−1.100	0.066	0.935	−6.500	0.065	0.782	−0.800	0.069	0.953
$\sigma_2^{(2)} = 1.69$	−4.970	0.100	0.818	0.059	0.100	0.953	−7.278	0.097	0.723	−0.118	0.104	0.940
$\epsilon^{(1)} = 1$	−2.100	0.058	0.886	−0.600	0.053	0.935	−2.400	0.055	0.905	−0.400	0.057	0.930
$\epsilon^{(2)} = 0.64$	2.500	0.033	0.927	0.000	0.034	0.935	−3.750	0.034	0.882	0.000	0.036	0.963
$\eta^{(1)}$=0.5	-	-	-	1.400	0.034	0.926	-	-	-	0.400	0.032	0.940
$\eta^{(2)}$=0.7	-	-	-	0.286	0.026	0.944	-	-	-	0.285	0.024	0.953
$\gamma = 1$	−0.800	0.171	0.959	−0.400	0.185	0.963	−0.600	0.185	0.973	−0.200	0.186	0.986
$\nu^{(1)} = 0.4$	6.250	0.061	0.900	4.000	0.056	0.926	6.500	0.057	0.918	4.000	0.057	0.926
$\nu^{(2)} = 0.2$	6.000	0.038	0.923	−0.500	0.035	0.958	6.500	0.036	0.932	0.000	0.036	0.935
$\lambda_0 = 0.006$	0.000	0.001	0.945	0.000	0.001	0.949	0.000	0.001	0.935	0.000	0.001	0.935

Simulation results show that parameter estimations from joint model outperform the naive model in both setting. Generally, joint model has small bias (except one, but have better coverage probability), and coverage rate. Indeed, the simulation results show that the proposed joint model can accurately estimate the covariate coefficients with the presence of both monotone missing and intermittent missing data.

4.5 Application to PPMI study

In this section, we use our proposed models to handle PPMI data which carry both monotone and nonmonotone missing data. The dataset used in this study was downloaded on Nov. 9, 2017. In PPMI study, all subjects were grouped into several cohorts, Parkinson Disease (PD), Scans Without Evidence of Dopaminergic Degeneration (SWEDD) and Healthy

Control (HC) etc. We use the PD cohort, which includes 423 subjects. Excluding those having only one visit, there are 415 subjects in our study. Among these subjects, total 40 dropped out early for different reasons, and 197 individuals underwent ST. There are 3151 observations with complete recorded responses, and 151 missed records. We combine the events of dropout and ST, treating them as the initiation of monotone missing observation. There are total 237 individuals underwent events (dropouts or having ST) which are treated as having monotone missing data, and 158 missed visits which are recorded as the intermittent pattern of missing data. All the item responses in MDS-UPDRS part I, II and III are used as outcome responses. The structured questionnaire design (part I, II, and III) confines the information manifested by those ordinal responses in each part to the corresponding domains. To fit to the data structure, and incorporate the impairment of disease status across domains, we adept our models based on the assumption that item responses in three parts manifest the unobserved status of corresponding do- mains. We add dependency across domains by incorporating correlated random effects. Specifically, the full models for the two missingness patterns are $\text{logit}\{p(r_{ij} = 1|\boldsymbol{\theta_i(t_j)})\} = W_k + \boldsymbol{\eta_k'\theta_i(t_j)}$, and $\lambda_i(t) = \lambda_0(t)exp(\gamma V_i + \boldsymbol{\nu'\theta_i(t)})$, where vector $\boldsymbol{\eta_k}' = (\eta_1, \eta_2, \eta_3)$, and vector $\boldsymbol{\nu}' = (\nu_1, \nu_2, \nu_3)$, corresponding to domain-specific coefficients for testing different missing data patterns. This domain-specific setting enables us to incorporate the impairment across domains and regress the effects of heterogeneity of disease development across domains. To avoid over fitting, we first run full model incorporating all three domains into our missing data sub-models. We check the domain-specific missing data association, and run the reduced model while keeping the domains which significantly impact the missingness in missing data sub-models.

Table 4.51: Parameter estimates for PPMI with two missing patterns.

	nM-EDL				M-EDL				Motor Examination			
	MEAN	SD	95 % CI		MEAN	SD	95% CI		MEAN	SD	95% CI	
Disease Status (Latent Variable)												
Int.	0.013	0.057	−0.092	0.109	0.154	0.063	0.0483	0.259	0.222	0.095	0.064	0.388
Age (yr)	0.163	0.034	0.106	0.236	0.154	0.033	0.083	0.209	0.199	0.058	0.094	0.335
Time (yr)	0.290	0.016	0.259	0.318	0.452	0.015	0.425	0.483	0.678	0.041	0.602	0.767
Intermittent pattern of missing data												
η	0.206	0.074	0.014	0.314	-	-	-	-	-	-	-	-
Monotone pattern of missing data												
ν	-	-	-	-	0.248	0.041	0.161	0.337	-	-	-	-

*: random intercept

**: random slope

Table 4.51 shows the domain-specific parameter estimates and the 95% confidence intervals based on reduced model. We find significant age effects in all domains. Several studies addressed the role of age in PD severity with mixed results [54]. In this study, our analysis which is based on a small cohort of early stage PD patients discloses that age is an important factor in PD, and it significantly impacts both motor and non-motor progression. The time effects are significant in all three domains. Specifically, the disease is getting worse at average rate of 0.290 units per year (CI: [0.259, 0.318]) in nM-EDL domain, the disease progresses to worse status at average rate of 0.452 units per year (CI: [0.425, 0.483]) in M-EDL domain, while in Motor Examination domain, the disease is getting worse at average rate of 0.678 units per year (CI: [0.602, 0.767]). The intermittent missingness pattern is significant associated with disease status and development in nM-EDL with log odd increased 0.206 (CI=[0.014, 0.314]) for ever unit worsening of disease in nM-EDL while controlling other two disease status unchanged. The occurrence of monotone missingness is significantly associated with the disease

status and progression in M-EDL, the log hazard ratio increases 0.248 units (CI=[0.161, 0.337]) for every unit worsening of disease in M-EDL while controlling other two disease status unchanged. Overall, we can not exclude the MNAR for both missing data patterns. This new finding signifies the unique advantage of our model.

The correlation matrix (Appendix Table C.04) provides additional information of interdependency of diseases and correlated (in high dimension) subject-specific characteristics between and within domains. Every individual has 3 pairs of random effect terms sampled from a 6×6 covariance matrix, denoting the random effects in 3 dimensions or domains, each domain has its domain-specific random intercept and random slope. For each random effects vector, the 1st and 2nd elements are for nM-EDL domain, the 3rd and 4th are for domain in M-EDL, while the last two are for motor examination domain. The within-domain correlations are significant (between random intercept and random slope) in two domains, for nM-EDL we have ($\rho_{01} = -0.179,\ 95\%$ CI: $[-0.324, -0.027]$), while for motor examination we have ($\rho_{01} = -0.387,\ 95\%$ CI: $[-0.516, -0.258]$), explained as: if a participant's nM-EDL disease status is worse at the start of the trial, his or her nM-EDL's disease progression is slow during follow-up. Same interpretation is for disease in Motor Examination domain. The intercepts between-domain correlations are all positive and significant (random intercept vs random intercept), these associations reveal that in the study if any initial disease status is worse, the other two are also worse, for example, the initial disease status in nM-EDL is worse, the disease severities in M-EDL and motor examination are also worse. Same finding is discovered in the correlations of within-domain random slope (random slope vs random slope), if disease progresses fast in any domain, it also develops fast in other domains, for example, disease develops fast in M-EDL, it also deteriorates fast in nM-EDL and motor examination. Overall, our high dimensional random effect matrix (6X6) reveals the internal, hidden, between and within domain disease correlation.

4.6 Discussion

This study provides a joint model approach to the longitudinal data in presence of two missing data patterns without excluding MNAR assumption. We provide a framework on modeling two missing data patterns (monotone and intermittent) simultaneously. We use a logistic model to describe the intermittent missing data pattern, while modeling the monotone missing data pattern using Cox model. We use latent trait to link outcome responses and missing data measures (indices and events). In addition, our model provides a statistical test for missing data mechanism in sub-models for the two missing data patterns, we simplify this process to covariate effect hypothesis test. Moreover, in analyzing multiple mixed types data, our model incorporates domain-specific variability for disease. To incorporate the impairment across domains, we refine our model to obtain the heterogeneity of disease from different domains. We quantify the impact on missing data mechanisms from domain-specific disease status, and provide a direct interpretation for the impacts. This is the strength of our proposed model. In addition, our model can be extended to incorporate the high order impact, such as those from time-dependent disease progression. This approach can also be extended to competing risk model for multiple cause of endpoints.

We apply our approach to PPMI study, our model discloses that the intermittent missingness pattern is significantly associated with the disease development in nM-EDL, while the monotone missing pattern are mainly affected by the disease status in M-EDL. Specifically, both missingness patterns can not exclude MNAR. However, these two missingness patterns are not related to the disease progression in motor examination. One possible reason may be from the design of study, and cohort used in this study, as the goal of PPMI is to identify, test and verify markers of progression for early-stage Parkinson's disease, at which stage, motor impairment is less severe compared with non-motor impairment.

Our model has some limitations. Our joint analysis is based on latent traits to provide inference for nonignorable missingness with fully correlated random effects cross domain. We do not provide a method to address the partial missing data. To assess the underlying association when patients intentionally avoid certain medical tests or just answer part of questions, we need build a more general approach to the partial missing data, and test the missing data mechanism. In this study we provide a statistical test to test missing data mechanism, there are other models with different frameworks, we do not provide comparison of test results to other applicable models, such as selection model and mixture model. In general, local sensitivity analysis can be use to evaluate the robustness of inference of departure from assumption [61, 69, 95].

Appendix A

Appendix for article 1

Table A.01: Unified Parkinson's Disease Rating Scale (UPDRS).

Items	0	1	2	3	4
Part I: Mentation, Behavior and Mood (4 items)					
Intellectual Impairment	None	Mild	Moderate memory loss	Severe memory loss with disorientation for time and often to place	Severe memory loss with orientation preserved to person only
Thought Disorder	None	vivid dreaming	"Benign" hallucinations with insight retained	Occasional to frequent hallucinations or delusions	Persistent Hallucinations, delusions, or florid psychosis.
Depression	Not present	Periods of sadness or guilt greater than normal, never sustained for days or weeks.	sustained depression	Sustained depression with vegetative symptoms	Sustained depression with vegetative symptoms and suicide thoughts or intent
Motivation/ Initiative	Normal	Less assertive than usual, more passive.	Loss of initiative or disinterest in elective activities	Loss of initiative or disinterest in day to day activities	Withdraw, complete loss of motivation.
Part II: Activities of Daily Living (13 items)					
Speech	Normal	Mildly affected	Moderately affected	Severely affected	Unintelligible most of the time
Salivation	Normal	Slight but definitely excess of saliva in mouth	Moderately excessive saliva	Marked excess of saliva with some drooling	Marked drooling, requires constant tissue or handkerchief.
Swallowing	Normal	Rare choking	Occasional choking	Requires soft food	Requires NG tube or gastrotomy feeding
Handwriting	Normal	Slightly slow or small	Moderately slow or small; all words are legible.	Severely affected; not all words are legible.	The majority of words are not legible
Cutting Food & Handling Utensils	Normal	Somewhat slow and clumsy, but no help needed.	Can cut most foods, although clumsy and slow; some help needed.	Food must be cut by someone, but can still feed slowly.	Needs to be fed
Dressing	Normal	Somewhat slow, but no help needed.	Occasional assistance with buttoning, getting arms in sleeves.	Considerable help required, but can do some things alone.	Helpless

continued

Table A.01: Unified Parkinson's Disease Rating Scale (UPDRS).

Items	0	1	2	3	4
Hygiene	Normal	Somewhat slow, but no help needed.	Needs help to slower or bathe; or very slow in hygienic care.	Requires assistance for washing, brushing teeth, combing hair, going to bathroom.	Foley catheter or other mechanical aids
Turning in bed and Adjusting Bed clothes	Normal	Somewhat slow and clumsy, but no help needed	Can turn alone or adjust sheets, but with great difficulty.	Can initiate, but not turn or adjust sheets alone.	Helpless
Falling	None	Rare falling	Occasionally falls, less than once per day.	Falls an average of once daily.	Falls more than once daily
Freezing when Walking	None	Rare freezing	Occasional freezing	Frequently freezing	Frequently falls from freezing
Walking	Normal	Mild difficult	Moderate difficulty, but requires little or no assistance.	Severe disturbance of walking, requiring assistance.	Can not walk at all, even with assistance.
Tremor	Absent	Slight and infrequently present	Moderate; bothersome to patient.	Severe; interferes with many activities.	Marked; interferes with most activities.
Sensory Complaints Related to Parkinson-ism	None	Occasionally has numbness, tingling, or mild aching.	Frequently has numbness, tingling, or aching; not distressing.	Frequent painful sensations	Excruciating pain
Part III: Motor Examination (27 items)					
Speech	Normal	Slight loss of expression, diction and /or volume.	Monotone, slurred but understandable; moderately impaired.	Marked impairment, difficult to understand	Unintelligible
Facial Expression	Normal	Minimal hypomimia, could be normal "Poker Face".	Slight but definitely abnormal diminution of facial expression	Moderate hypomimia; lips parted some of the time.	Masked or fixed facies with severe or complete loss of facial expression; lips parted 1/4 inch or more.
Tremor at Rest (head, upper, and lower ex-tremities, 5 items)	Absent	Slight and infrequently present	Mild in amplitude and persistent. Or moderate in amplitude, but only intermittently present.	Moderate in amplitude and present most of the time	Marked in amplitude and present most of the time

continued

Table A.01: Unified Parkinson's Disease Rating Scale (UPDRS).

Items	0	1	2	3	4
Action or Postural Tremor of Hands (right & left, 2 items)	Absent	Slight; present with action	Moderate in amplitude, present with action.	Moderate in amplitude with posture holding as well as action	Marked in amplitude; interferes with feeding
Rigidity (major joints, 5 items)	Absent	Slight or detectable only when activated by mirror or other movements	Mild to moderate	Marked, but full range of motion easily achieved.	Severe, range of motion achieved with difficulty.
Finger Taps (right & left, 2 items)	Normal	Mild slowing and/or reduction in amplitude	Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.	Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement	Can barely perform the task
Hand Movements (right & left, 2 items)	Normal	Mild slowing and/or reduction in amplitude	Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.	Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement	Can barely perform the task
Rapid Alternating Movements of Hands (right & left, 2 items)	Normal	Mild slowing and/or reduction in amplitude	Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.	Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement	Can barely perform the task
Leg Agility (right & left, 2 items)	Normal	Mild slowing and/or reduction in amplitude	Moderately impaired. Definite and early fatiguing. May have occasional arrests in movement.	Severely impaired. Frequent hesitation in initiating movements or arrests in ongoing movement.	Can barely perform the task
Arising from Chair	Normal	Slow; or may need more than one attempt.	Pushes self up from arms of seat	Tends to fall back and may have to try more than one time, but can get up without help.	Unable to arise without help

continued

Table A.01: Unified Parkinson's Disease Rating Scale (UPDRS).

Items	0	1	2	3	4
Posture	Normal erect	Not quite erect, slightly stooped posture; could be normal for older person.	Moderately stooped posture, definitely abnormal; can be slightly leaning to one side.	Severely stooped posture with kyphosis; can be moderately leaning to one side.	Marked flexion with extreme abnormality of posture
Gait	Normal	Walks slowly, may shuffle with short steps, but no festination (hastening steps) or propulsion.	Walks with difficulty, but requires little or no assistance; may have some festination, short steps, or propulsion.	Severe disturbance of gait, requiring assistance.	Cannot walk at all, even with assistance.
Postural Stability	Normal	Retropulsion, but recovers unaided.	Absence of postural response; would fall if not caught by examiner.	Very unstable, tends to lose balance spontaneously.	Unable to stand without assistance
Body Bradykinesia and Hypokinesia	None	Minimal slowness, giving movement a deliberate character; could be normal for some persons. Possibly reduced amplitude.	Mild degree of slowness and poverty of movement which is definitely abnormal. Alternatively, some reduced amplitude.	Moderate slowness, poverty or small amplitude of movement.	Marked slowness, poverty or small amplitude of movement

Table A.02: Results of simulation for ordinal parameters.

	EST	Bias	SD	CP	RMSE
Difficult parameters					
$a_{1,2} = 2.000$	2.001	0.001	0.032	0.950	0.032
$a_{1,3} = 2.500$	2.503	0.003	0.040	0.942	0.040
$a_{1,4} = 3.500$	3.503	0.003	0.048	0.958	0.047
$a_{2,2} = 2.000$	2.001	0.001	0.029	0.933	0.029
$a_{2,3} = 2.500$	2.502	0.002	0.035	0.917	0.035
$a_{2,4} = 3.500$	3.500	−0.000	0.047	0.933	0.047
$a_{3,2} = 2.000$	2.000	0.000	0.027	0.942	0.027
$a_{3,3} = 2.500$	2.499	−0.001	0.029	0.967	0.029
$a_{3,4} = 3.500$	3.503	0.003	0.039	0.967	0.039
$a_{4,1} = 2.000$	1.997	−0.003	0.074	0.983	0.074
$a_{4,2} = 3.000$	2.995	−0.005	0.081	0.992	0.081
$a_{4,3} = 3.500$	3.494	−0.006	0.086	0.967	0.086
$a_{4,4} = 4.500$	4.494	−0.006	0.089	0.992	0.088
$a_{5,1} = 0.800$	0.800	−0.000	0.082	0.942	0.082
$a_{5,2} = 1.800$	1.804	0.004	0.089	0.942	0.089
$a_{5,3} = 2.300$	2.307	0.007	0.092	0.925	0.092
$a_{5,4} = 3.300$	3.307	0.007	0.093	0.950	0.093
$a_{6,1} = 0.500$	0.506	0.006	0.069	0.925	0.069
$a_{6,2} = 1.500$	1.508	0.008	0.077	0.933	0.077
$a_{6,3} = 2.000$	2.009	0.009	0.081	0.933	0.081
$a_{6,4} = 3.000$	3.016	0.016	0.105	0.917	0.106
$a_{7,1} = 0.300$	0.292	−0.008	0.058	0.967	0.058
$a_{7,2} = 1.300$	1.294	−0.006	0.059	0.975	0.059
$a_{7,3} = 1.800$	1.793	−0.007	0.058	0.983	0.058
$a_{7,4} = 2.800$	2.794	−0.006	0.062	0.983	0.062

Table A.02 – continued from previous page

	EST	Bias	SD	CP	RMSE
$a_{8,1} = 0.600$	0.605	0.005	0.071	0.917	0.071
$a_{8,2} = 1.600$	1.606	0.006	0.072	0.933	0.072
$a_{8,3} = 2.100$	2.108	0.008	0.073	0.925	0.073
$a_{8,4} = 3.100$	3.112	0.012	0.079	0.900	0.080
$a_{9,1} = 1.200$	1.210	0.010	0.076	0.925	0.077
$a_{9,2} = 2.200$	2.214	0.014	0.077	0.967	0.078
$a_{9,3} = 2.700$	2.717	0.017	0.085	0.950	0.086
$a_{9,4} = 3.700$	3.712	0.012	0.116	0.967	0.116
$a_{10,1} = 1.500$	1.507	0.007	0.068	0.975	0.068
$a_{10,2} = 2.500$	2.511	0.011	0.072	0.975	0.072
$a_{10,3} = 3.000$	3.015	0.015	0.073	0.967	0.074
$a_{10,4} = 4.000$	4.023	0.023	0.079	0.942	0.082
Discrimination parameters					
$b_4^{(1)} = 0.600$	0.607	0.007	0.043	0.942	0.043
$b_5^{(1)} = 0.700$	0.705	0.005	0.048	0.900	0.048
$b_6^{(1)} = 0.200$	0.202	0.002	0.040	0.958	0.040
$b_7^{(1)} = 0.500$	0.501	0.001	0.033	0.925	0.033
$b_8^{(1)} = 0.100$	0.105	0.005	0.036	0.925	0.036
$b_9^{(1)} = -0.300$	−0.304	−0.004	0.052	0.925	0.052
$b_{10}^{(1)} = 0.300$	0.311	0.011	0.033	0.967	0.035
$b_4^{(2)} = -0.100$	−0.102	−0.002	0.045	0.975	0.045
$b_5^{(2)} = 0.600$	0.600	−0.000	0.060	0.925	0.060
$b_6^{(2)} = 0.200$	0.197	−0.003	0.049	0.975	0.049
$b_7^{(2)} = -0.200$	−0.199	0.001	0.040	0.958	0.040
$b_8^{(2)} = 0.200$	0.201	0.001	0.043	0.942	0.043

Table A.02 – continued from previous page

	EST	Bias	SD	CP	RMSE
$b_9^{(2)} = 0.300$	0.301	0.001	0.065	0.933	0.065
$b_{10}^{(2)} = -0.400$	−0.409	−0.009	0.047	0.925	0.048
$b_4^{(3)} = 0.200$	0.199	−0.001	0.026	0.992	0.026
$b_5^{(3)} = -0.400$	−0.401	−0.001	0.034	0.950	0.034
$b_6^{(3)} = -0.300$	−0.298	0.002	0.029	0.942	0.029
$b_7^{(3)} = 0.100$	0.098	−0.002	0.023	0.925	0.023
$b_8^{(3)} = 0.100$	0.099	−0.001	0.025	0.925	0.025
$b_9^{(3)} = -0.200$	−0.199	0.001	0.035	0.875	0.035
$b_{10}^{(3)} = 0.400$	0.404	0.004	0.026	0.950	0.026

Table A.03: Estimated random effects.

	Mean	SD	95% CI	
$\sigma_0^{(1)}$	1.646	0.060	1.537	1.773
$\sigma_1^{(1)}$	0.337	0.019	0.302	0.373
$\sigma_0^{(2)}$	1.844	0.048	1.749	1.943
$\sigma_1^{(2)}$	0.474	0.017	0.439	0.507
$\sigma_0^{(3)}$	1.250	0.032	1.190	1.315
$\sigma_1^{(3)}$	0.327	0.011	0.307	0.349

σ_0 : random intercept.
σ_1 : random slope.

Table A.04: Correlation coefficient estimates for random effects correlation matrix.

MBM		ADL		Motor		
1.000	$-0.069^{*}_{0.054}$	$0.449_{0.027}$	$0.065_{0.035}$	$0.204_{0.030}$	$0.096_{0.035}$	MBM_{sd}
	$[-0.170, 0.036]^{**}$	$[0.398, 0.501]$	$[0.000, 0.138]$	$[0.147, 0.263]$	$[0.032, 0.172]$	$95\%CI$
	1.000	$-0.087_{0.044}$	$0.734_{0.034}$	$0.006_{0.043}$	$0.497_{0.042}$	MBM_{sd}
		$[-0.176, 0.003]$	$[0.664, 0.799]$	$[-0.074, 0.094]$	$[0.409, 0.575]$	$95\%CI$
		1.000	$-0.093_{0.035}$	$0.628_{0.018,}$	$-0.034_{0.035,}$	ADL_{sd}
			$[-0.158, -0.021]$	$[0.591, 0.663]$	$[-0.099, 0.032]$	$95\%CI$
			1.000	$-0.025_{0.033}$	$0.861_{0.016}$	ADL_{sd}
				$[-0.091, 0.038]$	$[0.828, 0.892]$	$95\%CI$
				1.000	$-0.009_{0.033}$	$Motor_{sd}$
					$[-0.070, 0.056]$	$95\%CI$
					1.000	$Motor$

* $-0.069_{0.054}$: correlation coefficient mean estimate=-0.069, SD=0.054.

** $[-0.170, 0.036]$: 95% CI=$[-0.171, 0.040]$ for the above element ($-0.069_{0.054}$).

Table A.05: Estimated random errors in three UPDRS parts.

	Mean	SD	95% CI	
$\sigma_e^{(1)}$	0.006	0.002	0.003	0.010
$\sigma_e^{(2)}$	0.280	0.019	0.247	0.317
$\sigma_e^{(3)}$	0.310	0.016	0.282	0.342

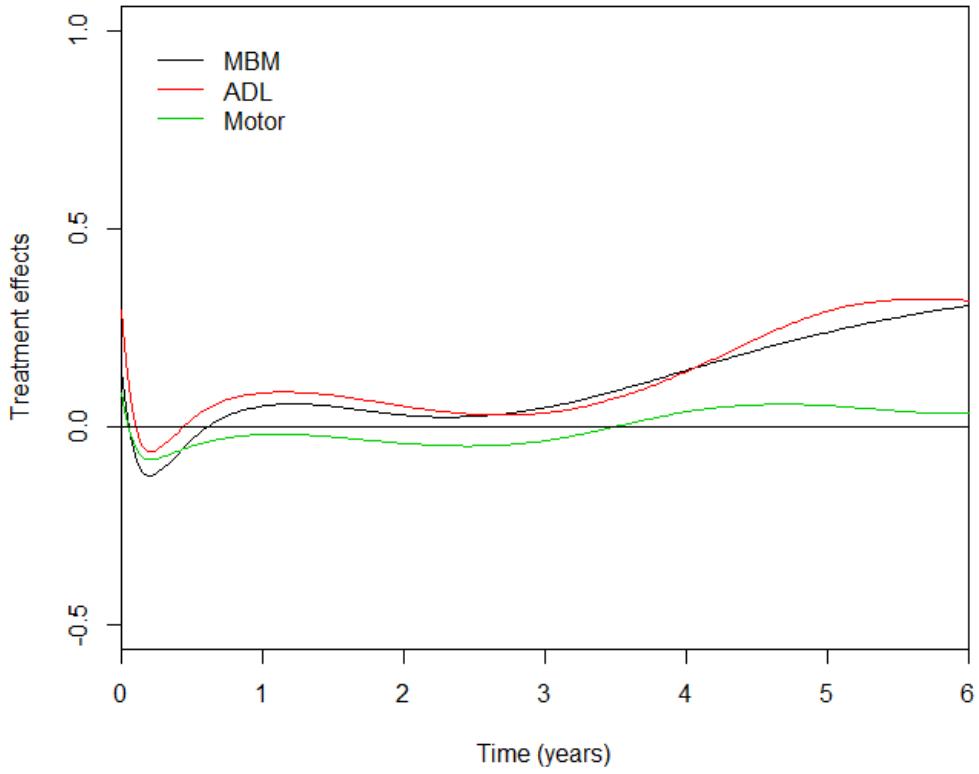

Figure A.01: Impaired domain-specific and heterogeneous treatment effects in time horizon for UPDRS I, II and III. MBM: mentation, behavior and mood. ADL: activities of daily living.

Table A.06: Estimated difficulty parameters in UPDRS Part I: MBM.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Mentation*	0.400	3.937	6.669	8.675	1.000	0.000	1.000	1.000
Thought Disorder	0.787	3.127	5.402	7.717	0.496	0.025	0.445	0.546
Depression	1.010	3.651	5.511	8.717	0.705	0.035	0.636	0.773
Motivation/Initiative	0.827	3.026	4.644	8.110	0.831	0.039	0.753	0.905

$a_{k,l}$: item k's level l.
*: item to put constraints.

Table A.07: Estimated difficulty parameters in UPDRS Part II: ADL.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Speech	−0.578	1.321	3.764	7.654	0.605	0.017	0.573	0.639
Salivation	−0.017	1.938	4.115	6.378	0.462	0.015	0.434	0.492
Swallowing	1.571	2.964	5.966	7.653	0.348	0.015	0.318	0.379
Handwriting	−1.570	0.335	2.007	4.340	0.576	0.016	0.546	0.606
Cutting food	−0.021	3.405	6.341	10.102	0.896	0.023	0.852	0.943
Dressing	−0.913	2.655	6.508	10.169	1.036	0.026	0.987	1.089
Hygiene*	0.396	5.093	7.503	10.728	1.000	0.000	1.000	1.000
Turing in bed	0.027	3.064	6.272	9.143	0.850	0.023	0.806	0.895
Falling	2.455	4.272	6.223	7.275	0.536	0.020	0.497	0.575
Freezing	1.643	3.193	5.032	7.745	0.628	0.020	0.590	0.667
Walking	−1.563	2.595	4.766	8.101	0.705	0.020	0.668	0.743
Tremor	−1.492	0.825	4.022	6.112	0.045	0.009	0.028	0.064
Sensory symptoms	0.531	2.002	3.499	6.604	0.195	0.011	0.173	0.216

$a_{k,l}$: item k's level l.

*: item to put constraints.

Table A.08: Estimated parameters in UPDRS Part III: motor.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Speech	−0.595	2.450	5.633	9.603	0.773	0.021	0.735	0.815
Facial Expression	−1.867	0.775	3.662	7.253	0.744	0.021	0.705	0.785
Tremor at rest: face, lips,chin	2.557	4.393	6.615	89.269	0.273	0.024	0.229	0.319
Tremor at rest: right hands	0.640	1.586	3.417	6.645	0.077	0.013	0.052	0.103
Tremor at rest: left hands	0.851	1.843	3.840	7.247	0.076	0.014	0.048	0.104
Tremor at rest: right feet	2.197	3.584	5.670	8.395	0.098	0.021	0.057	0.138
Tremor at rest: left feet	2.404	3.659	5.702	8.893	0.145	0.021	0.106	0.186
Action tremor: right	0.773	2.996	4.505	8.945	0.237	0.015	0.207	0.265
Action tremor: left	0.562	2.821	4.732	7.867	0.222	0.014	0.195	0.248
Rigidity: neck	−0.170	1.740	4.420	6.348	0.827	0.022	0.785	0.870
Rigidity: right upper extremity	−1.034	0.978	4.476	7.835	0.605	0.019	0.566	0.644
Rigidity: left upper extremity	−0.706	1.298	4.722	8.354	0.757	0.020	0.717	0.799
Rigidity: right lower extremity	0.163	1.968	5.196	8.746	0.775	0.023	0.729	0.821
Rigidity: left lower extremity	0.166	2.005	5.342	8.137	0.869	0.024	0.824	0.912
Finger taps: right	−1.093	1.335	3.940	6.974	0.770	0.023	0.726	0.815
Finger taps: left	−1.169	1.180	3.906	7.147	0.942	0.023	0.899	0.987
Hand grips: right	−0.582	2.089	5.115	8.425	0.892	0.026	0.844	0.940
Hand grips: left*	−0.643	1.829	4.698	7.847	1.000	0.000	1.000	1.000
Hand pronate: right	−0.400	2.074	4.569	7.707	0.858	0.024	0.814	0.907
Hand pronate: left	−0.516	1.800	4.244	6.775	0.967	0.024	0.920	1.014
Leg agility: right	0.244	2.760	5.369	7.800	0.872	0.025	0.824	0.920
Leg agility: left	0.054	2.348	5.041	7.455	0.949	0.023	0.903	0.992
Arise from chair	2.120	4.262	5.899	6.645	0.900	0.028	0.847	0.954
Posture	−0.445	2.364	5.311	7.502	0.811	0.022	0.771	0.854

STAN code

```
data {
int<lower=0> num_subject;
int<lower=0> num_obs;
int<lower=1> num_rho;


int subject[num_obs];

int<lower=1> num_basis;
int<lower=1> num_pred;
int<lower=0> K;
int<lower=1> Y[num_obs, K];
int<lower=0> n_ordi;
int<lower=0> n_theta;
vector [n_theta] a0;

real<lower=0> time_obs[num_obs];
int<lower=0> treat_obs[num_obs];


vector<lower=0> [num_basis] time_obs_Bt [num_obs];
vector [num_basis] pred_t_Bt [num_pred];
}
parameters {
vector <lower=-10, upper=10> [2] beta [n_theta];
vector[n_theta] U[num_subject];
vector <lower=0> [n_theta] Var_U;
corr_matrix[n_theta] Omega;
vector[n_theta] e[num_obs];
vector <lower=0> [n_theta] Var_e;

vector [K-n_theta] a_random;

vector[n_theta] b_random[K-n_theta];
vector<lower=0>[n_ordi-2] delta[K];
vector [num_basis] cc [n_theta];
```

```
vector <lower=0> [n_theta] Var_spline ;
}
transformed parameters {
cov_matrix [n_theta] Sigma_U;

ordered[n_ordi-1] a[K];
vector[n_theta] b[K];

vector [n_theta] fbt [num_obs];

vector[n_theta] theta[num_obs];

vector <lower=0> [n_theta] sig;
vector  <lower=0> [n_theta]  sd_spline;
vector <lower=0> [n_theta] sd_e;
vector [K] y_ordi_hat [num_obs];

for (i in 1:num_obs) {
for(p in 1:n_theta) {
fbt[i,p] <- cc[p]'*time_obs_Bt[i] ;
theta[i,p]<- beta[p,1] + beta[p,2]*time_obs[i] + fbt[i,p]*treat_obs[i] + U[subject[i],p]+e[i,p];
}
}

for (k in 1:n_theta) {
a[k,1] <- a0[k];
for (l in 2:(n_ordi-1)) {
a[k, l] <- a[k, l-1] + delta[k, l-1] ;
}
}

for (k in (n_theta+1):K) {
a[k, 1] <- a_random[k-n_theta];
for (l in 2:(n_ordi-1)) a[k, l] <- a[k, l-1] + delta[k, l-1];
}


b[1, 1] <- 1;
b[1, 2] <- 0;
```

```
b[1, 3] <- 0;

b[2, 1] <- 0;
b[2, 2] <- 1;
b[2, 3] <- 0;

b[3, 1] <- 0;
b[3, 2] <- 0;
b[3, 3] <- 1;


for (k in (n_theta+1):K)
b[k] <- b_random[k-n_theta]; // b[k] is a vector


for (i in 1: num_obs){
for (k in 1:K) y_ordi_hat[i,k]<- b[k]'* theta[i];

}



sig       = sqrt(Var_U);
sd_spline = sqrt(Var_spline);
Sigma_U   = quad_form_diag(Omega, sig);
sd_e      = sqrt(Var_e);

}

model {
vector [n_theta] zero=[0,0,0]';
U ~ multi_normal(zero, Sigma_U);
for (i in 1:num_obs)    e[i] ~ normal(0, sd_e);

for(p in 1:n_theta) {
cc[p,1] ~ normal(0, 10);   // intercept and first c initialization

for (j in 2: num_basis) cc[p,j]~ normal(cc[p, j-1], sd_spline[p]);
```

```
}

for(i in 1:num_obs){
for (k in 1:K) Y[i,k] ~ ordered_logistic(y_ordi_hat[i,k], a[k]);
}

for (k in 1:K) {
for (l in 1:(n_ordi-2)) {
delta[k, l] ~ normal(0, 100) T[0,] ;
}
}
for (k in 1:(K-n_theta)) {
b_random[k] ~ normal(0, 20);
}
;
a_random ~ normal(0, 20);
beta ~ normal(0,20);

for (p in 1:n_theta) beta[p] ~ normal(0,20);


Var_U ~ inv_gamma(0.01, 0.01);
Omega ~ lkj_corr(2.0);

Var_spline ~ inv_gamma(0.01,0.01);
Var_e ~ inv_gamma(0.01, 0.01);
```

Appendix B

Appendix for article 2

Table B.01: Results of simulation for ordinal parameters.

	EST	Bias	SD	CP
Difficult parameters				
$a_{1,2} = 1.700$	1.699	−0.001	0.025	0.948
$a_{1,3} = 2.500$	2.496	−0.004	0.031	0.943
$a_{1,4} = 3.500$	3.496	−0.004	0.042	0.919
$a_{2,2} = 1.700$	1.699	−0.001	0.024	0.929
$a_{2,3} = 2.500$	2.499	−0.001	0.030	0.914
$a_{2,4} = 3.500$	3.498	−0.002	0.036	0.943
$a_{3,2} = 1.800$	1.800	0.000	0.040	0.929
$a_{3,3} = 2.600$	2.597	−0.003	0.042	0.957
$a_{3,4} = 3.600$	3.599	−0.001	0.050	0.967
$a_{4,1} = 0.100$	0.102	0.002	0.035	0.971
$a_{4,2} = 1.800$	1.801	0.001	0.041	0.952
$a_{4,3} = 2.600$	2.602	0.002	0.046	0.938
$a_{4,4} = 3.600$	3.605	0.005	0.053	0.957

Table B.01 – continued from previous page

	EST	Bias	SD	CP
$a_{5,1} = -0.100$	-0.098	0.002	0.037	0.952
$a_{5,2} = 1.600$	1.601	0.001	0.041	0.952
$a_{5,3} = 2.400$	2.400	0.000	0.044	0.948
$a_{5,4} = 3.400$	3.400	0.000	0.049	0.962
$a_{6,1} = 0.200$	0.200	0.000	0.034	0.952
$a_{6,2} = 1.900$	1.900	0.000	0.037	0.962
$a_{6,3} = 2.700$	2.702	0.002	0.043	0.943
$a_{6,4} = 3.700$	3.701	0.001	0.053	0.952
$a_{7,1} = -0.200$	-0.199	0.001	0.035	0.938
$a_{7,2} = 1.500$	1.503	0.003	0.038	0.952
$a_{7,3} = 2.300$	2.303	0.003	0.041	0.948
$a_{7,4} = 3.300$	3.305	0.005	0.052	0.943
$a_{8,1} = 0.300$	0.298	-0.002	0.038	0.943
$a_{8,2} = 2.000$	1.998	-0.002	0.041	0.933
$a_{8,3} = 2.800$	2.801	0.001	0.046	0.943
$a_{8,4} = 3.800$	3.799	-0.001	0.062	0.924
Discrimination parameters				
$b_4^{(1)} = 0.200$	0.202	0.002	0.014	0.938
$b_5^{(1)} = 0.500$	0.502	0.002	0.018	0.933
$b_6^{(1)} = 0.100$	0.100	0.000	0.013	0.933
$b_7^{(1)} = -0.200$	-0.202	-0.002	0.014	0.924
$b_8^{(1)} = 0.200$	0.199	-0.001	0.013	0.962
$b_4^{(2)} = -0.300$	-0.302	-0.002	0.018	0.933

Table B.01 – continued from previous page

	EST	Bias	SD	CP
$b_5^{(2)} = -0.200$	-0.200	0.000	0.018	0.924
$b_6^{(2)} = 0.200$	0.200	0.000	0.015	0.962
$b_7^{(2)} = 0.300$	0.303	0.003	0.015	0.952
$b_8^{(2)} = -0.100$	-0.100	0.000	0.017	0.938

Table B.02: Results of simulation for continuous parameters.

	EST	BIAS	SD	CP
Outcome specific parameters				
a_1	-4.020	-0.020	0.065	0.962
a_2	-2.017	-0.017	0.080	0.948
a_3	1.980	-0.020	0.103	0.967
$b_1^{(1)}$	1.006	0.006	0.030	0.952
$b_1^{(2)}$	0.504	0.004	0.029	0.938
$b_2^{(1)}$	1.511	0.011	0.046	0.929
$b_2^{(2)}$	1.506	0.006	0.039	0.943
$b_3^{(1)}$	1.006	0.006	0.040	0.957
$b_3^{(2)}$	1.003	0.003	0.049	0.938
Random errors				
$\epsilon^{(1)}$	0.635	-0.005	0.040	0.910
$\epsilon^{(2)}$	0.358	-0.002	0.025	0.914

Table B.03: Model selection criteria for PPMI study.

	Model 1	Model 2	Model 3
DIC_3	323,306	323,321	323,319
WAIC	323,668	323,702	323,687

Table B.04: Posterior estimations with three domains.

	nM-EDL				M-EDL				Motor examination			
	Mean	SD	95% CI		Mean	SD	95% CI		Mean	SD	95% CI	
Disease Status												
Int.	0.055	0.086	−0.110	0.216	−0.315	0.107	−0.502	−0.102	−0.265	0.182	−0.621	0.094
Age (yr)	0.105	0.060	−0.018	0.226	0.131	0.076	−0.044	0.273	0.271	0.157	−0.005	0.588
Competing risks' Linking Parameters												
Dropout	0.083	0.183	−0.275	0.441	0.309	0.158	0.004	0.630	0.010	0.061	−0.105	0.131
ST	0.076	0.101	−0.127	0.277	0.194	0.083	0.029	0.362	0.013	0.027	−0.038	0.062

Table B.05: Estimated coefficients for continuous outcome (SDM).

	Mean	SD	95% CI	
a	41.221	0.310	40.611	41.853
$b^{(1)}$	−1.248	0.350	−1.936	−0.575
$b^{(2)}$	−1.040	0.284	−1.600	−0.475
$b^{(3)}$	−0.190	0.098	−0.374	0.013
σ	10.404	0.196	10.009	10.815

Table B.06: Estimated random effects.

	Mean	SD	95% CI	
$\sigma_0^{(1)}$	1.174	0.072	1.045	1.317
$\sigma_1^{(1)}$	0.243	0.021	0.204	0.284
$\sigma_0^{(2)}$	1.528	0.082	1.386	1.706
$\sigma_1^{(2)}$	0.358	0.029	0.308	0.415
$\sigma_0^{(3)}$	3.485	0.171	3.154	3.837
$\sigma_1^{(3)}$	0.545	0.039	0.475	0.624

σ_0 : random intercept.
σ_1 : random slope.
superscript (k): kth domain.

Table B.07: Estimated correlation coefficients for random effects.

nM-EDL	M-EDL		Motor		
1.000	$-0.204^{*}_{0.078}$	$0.625_{0.037}$	$-0.105_{0.074}$	$0.194_{0.051}$	$0.014_{0.074}$	$nM-EDL_{sd}$
	$[-0.341, -0.043]^{**}$	$[0.547, 0.694]$	$[-0.254, 0.041]$	$[0.092, 0.296]$	$[-0.130, 0.152]$	$95\%CI$
	1.000	$-0.043_{0.082}$	$0.627_{0.062}$	$-0.105_{0.077}$	$0.200_{0.089}$	$nM-EDL_{sd}$
		$[-0.203, 0.119]$	$[0.499, 0.734]$	$[-0.246, 0.057]$	$[0.030, 0.374]$	$95\%CI$
		1.000	$-0.168_{0.073}$	$0.375_{0.045}$	$-0.080_{0.072}$	$M-EDL_{sd}$
			$[-0.311, -0.028]$	$[0.284, 0.460]$	$[-0.213, 0.067]$	$95\%CI$
			1.000	$-0.129_{0.073}$	$0.639_{0.065}$	$M-EDL_{sd}$
				$[-0.278, 0.021]$	$[0.507, 0.755]$	$95\%CI$
				1.000	$-0.426_{0.064}$	$motor_{sd}$
					$[-0.544, -0.295]$	$95\%CI$
					1.000	$Motor$

* $-0.204_{0.078}$: correlation coefficient mean estimate=-0.204, SD=0.078.

** $[-0.341, -0.040]$: 95% CI=$[-0.341, -0.040]$ for the above element ($-0.204_{0.078}$).

Table B.08: Estimated random errors in latent traits.

	Mean	SD	95% CI	
$\sigma_e^{(1)}$	0.276	0.022	0.238	0.319
$\sigma_e^{(2)}$	0.527	0.028	0.478	0.588
$\sigma_e^{(3)}$	1.250	0.050	1.157	1.345

Table B.09: Estimated parameters in MDS-UPDRS Part I: nM-EDL.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Cognitive Impairment	1.319	3.682	5.358	7.435	0.952	0.061	0.836	1.070
Hallucinations	3.398	5.696	6.625	8.570	0.818	0.079	0.665	0.969
Depressed Mood*	1.500	3.524	5.171	7.343	1.000	0.000	1.000	1.000
Anxious Mood	1.039	3.134	4.827	7.372	0.824	0.056	0.717	0.935
Apathy	2.093	3.909	6.033	8.380	1.183	0.077	1.035	1.331
Dopamine Dysregulation	3.932	5.689	8.415	90.597	0.761	0.093	0.579	0.953
Sleep Problem	−0.049	1.333	2.651	4.611	0.783	0.056	0.675	0.895
Daytime Sleepiness	−0.214	1.488	5.053	7.480	1.009	0.071	0.876	1.151
Pain & other sensations	−0.153	2.024	3.351	5.244	0.772	0.054	0.666	0.882
Urinary	−0.019	1.980	3.560	5.185	0.716	0.053	0.613	0.822
Constipation	0.567	2.825	4.462	9.388	0.758	0.057	0.650	0.870
Light Headedness	1.324	3.370	5.126	7.856	0.937	0.066	0.809	1.071
Fatigue	0.223	3.108	5.199	7.167	1.623	0.102	1.427	1.831

$a_{k,l}$: item k's level l.

*: Constrained item.

Table B.010: Estimated difficulty parameters in MDS-UPDRS Part II: M-EDL.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Speech	0.708	2.277	4.640	8.771	0.774	0.044	0.691	0.864
Saliva & Drooling	0.722	1.608	2.970	4.864	0.661	0.040	0.588	0.742
Chewing & Swallowing	1.968	4.537	5.429	9.624	0.664	0.043	0.583	0.748
Eating Tasks*	0.800	3.748	7.361	89.451	1.000	0.000	1.000	1.000
Dressing	0.396	4.243	8.742	11.359	1.454	0.073	1.310	1.604
Hygiene	1.342	6.471	9.243	11.172	1.098	0.059	0.984	1.220
Handwriting	−0.618	1.388	3.173	5.374	0.701	0.040	0.625	0.786
Doing Hobbies	0.563	3.283	5.657	7.297	1.142	0.057	1.037	1.256
Turning in Bed	0.961	4.899	7.265	10.462	0.940	0.053	0.838	1.049
Tremor	−1.681	1.083	3.382	6.185	0.199	0.024	0.154	0.247
Getting out of bed	0.344	3.785	6.325	10.470	1.238	0.068	1.108	1.376
Walking & Balance	0.509	4.053	5.063	7.937	0.917	0.053	0.818	1.031
Freezing	3.725	5.811	7.743	10.101	1.134	0.076	0.991	1.288

$a_{k,l}$: item k's level l.

*: Constrained item.

Table B.011: Estimated parameters in MDS-UPDRS Part III: Motor Examination.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Speech	−0.101	2.646	6.271	89.377	0.212	0.013	0.187	0.241
Facial Expression	−2.197	0.868	3.326	6.835	0.303	0.015	0.273	0.332
Rigidity-Neck	−0.089	1.471	4.218	7.014	0.305	0.016	0.274	0.339
Rigidity-RUE	−1.204	0.292	2.750	7.141	0.012	0.008	0.001	0.028
Rigidity-LUE	−0.501	1.721	5.156	11.291	0.624	0.027	0.575	0.677
Rigidity-RLE	−0.073	1.284	3.284	6.303	0.069	0.010	0.049	0.091
Rigidity-LLE	0.639	2.220	4.578	8.417	0.464	0.021	0.423	0.509
Finger Tapping-Right	−1.148	0.527	2.204	4.981	0.034	0.009	0.017	0.052
Finger Tapping-Left	−1.108	2.259	5.362	9.806	0.987	0.040	0.912	1.066
Hand Movement-Right	−0.713	0.992	2.727	5.982	0.049	0.009	0.029	0.067
Hand Movement-Left*	−0.400	2.976	6.212	10.896	1.000	0.000	1.000	1.000
Pronation-Right	−0.672	0.987	2.857	6.351	0.009	0.006	0.000	0.024
Pronation-Left	−0.121	2.674	5.404	9.380	0.846	0.035	0.784	0.921
Toe Tapping-Right	−0.664	1.100	2.955	5.608	0.060	0.010	0.042	0.079
Toe Tapping-Left	−0.660	1.946	4.481	8.348	0.680	0.028	0.626	0.739
Leg Agility-Right	0.171	2.153	4.194	6.525	0.099	0.011	0.078	0.120
Leg Agility-Left	0.731	3.178	5.688	8.685	0.649	0.030	0.593	0.710
Arising from chair	1.701	3.748	5.429	7.087	0.221	0.018	0.186	0.257
Gait	−0.717	2.461	4.723	6.944	0.186	0.013	0.158	0.211
Freezing of Gait	4.589	6.339	6.942	7.323	0.340	0.045	0.256	0.429
Postural Stability	2.232	3.269	3.843	6.160	0.175	0.019	0.138	0.212
Posture	−0.474	1.853	4.027	6.047	0.201	0.013	0.175	0.227
Body Bradykinesia	−2.558	0.373	2.855	7.860	0.349	0.017	0.317	0.383
Postural Tremor-Right	0.611	2.460	4.606	15.133	0.002	0.002	0.000	0.006

STAN code

```
functions {
real expit(real x){
real ps;

ps=exp(x)/(1+exp(x));
return  ps;
}

vector convert_p(real[] psi){
int N= size(psi);
vector [N+1] pr;

pr[1]= psi[1];
for (k in 2:N) pr[k]=psi[k]-psi[k-1];
pr[N+1]=1-psi[N];
return pr;
}

real Sum_const_nupart (vector nu_vect, vector u0_vect, vector beta_vect, real cov_x_v) {

real sum_const_nupart;

sum_const_nupart=nu_vect' *(u0_vect + cov_x_v * beta_vect); // vector operation scalar * vector

return sum_const_nupart;

}

real pointV( vector nu_vect, vector u1_vect, real tee_X15_v, vector ft_vect){
real sub_fk;
sub_fk=nu_vect' * (tee_X15_v * u1_vect+ft_vect);
return sub_fk;
}

}

data {
```

```
int<lower=1> num_subject;
int<lower=1> num_obs;
int<lower=1> num_conti;
int<lower=1> KK;//KK=15 Gauss_Kronrod points #
int<lower=1> num_long;
int<lower=1> num_basis;
int<lower=1> num_pw;
int<lower=1> num_pred;

int<lower=1> num_part1;
int<lower=1> num_part2;
int<lower=1> num_part3;
int<lower=1> num_ordi;  // level of each question
int<lower=0> num_theta; // num of domain

int subj_long[num_obs]; // subject ID in long fmt

int<lower=1> updrs1[num_obs, num_part1];
int<lower=1> updrs2[num_obs, num_part2];
int<lower=1> updrs3[num_obs, num_part3];

vector [num_conti] SDM ;
int conti_match[num_conti] ;
vector[num_theta] a0;

real<lower=0> time_obs[num_obs];

real age_norm[num_subject];  // normalized data

int <lower=0, upper=1> gender_subj[num_subject];

real <lower=0> tee [num_subject];
int <lower=0> event[num_subject];   // possible value 0, 1, 2
int <lower=1, upper=num_pw> tee_id [num_subject];   // use to indicate with piece h to use

vector [KK] c15;
vector <lower=0> [KK] tee_X15 [num_subject];  //tee_x15 for each subject, 15 pts matrix
vector <lower=0> [num_basis] teeX15_long_Bt [num_long]; // num_long= num_subj * 15, long X15 bs format Matrix
vector <lower=0> [num_basis] tee_Bt [num_subject];  // use to construct fb_tee nonpar basis for tee
```

```
int <lower=1, upper=num_pw> pwht_X15_ind [num_subject, KK];//indicator for piecewise h, indicator for which pw for each x15 point

vector<lower=0> [num_basis] time_obs_Bt [num_obs];
vector <lower=0> [num_basis] pred_t_Bt [num_pred];
}


parameters {

vector [num_theta*2] U[num_subject];
corr_matrix[2*num_theta] Omega; // cholesky correlation matrix only intercept
vector<lower=0> [2*num_theta] Var_U; //  random scale
real<lower=0> Var_conti;
vector[num_theta] ee [num_obs];
vector<lower=0>[num_theta] Var_e;
vector [num_theta] beta0;
real a_conti;
vector [num_theta] b_conti;

vector [num_part1] a_random1;
vector [num_part2] a_random2;
vector [num_part3] a_random3;

vector <lower=0>  [num_part1] b_random1;
vector <lower=0>  [num_part2] b_random2;
vector <lower=0>  [num_part3] b_random3;


vector<lower=0> [num_ordi-2] delta1[num_part1];
vector<lower=0> [num_ordi-2] delta2[num_part2];
vector<lower=0> [num_ordi-2] delta3[num_part3];
vector [num_basis] cc [num_theta];  // B-spline

vector <lower=0> [num_theta] Var_spline ;


vector [2] gam ;
vector <lower=0> [num_pw] h_pw [2];
```

```
vector [num_theta] nu[2];

}

transformed parameters {
cov_matrix [2*num_theta] Sigma_U;
vector<lower=0> [2*num_theta] sd_U;
real<lower=0> sd_conti;

vector [num_conti] mu_SDM;
vector<lower=0> [num_theta] sd_e;
ordered[num_ordi-1] a_ordi_part1[num_part1];
ordered[num_ordi-1] a_ordi_part2[num_part2];
ordered[num_ordi-1] a_ordi_part3[num_part3];


vector   [num_part1] b_ordi_part1 ;
vector   [num_part2] b_ordi_part2 ;
vector   [num_part3] b_ordi_part3 ;


vector [num_theta] ft [num_obs];

vector[num_theta] theta[num_obs];

vector  <lower=0> [num_theta]  sd_spline;
vector [num_part1] updrs1_ordi_hat [num_obs];
vector [num_part2] updrs2_ordi_hat [num_obs];
vector [num_part3] updrs3_ordi_hat [num_obs];


vector [num_subject] exp_const[2];

vector [num_theta] U0  [num_subject];
vector [num_theta] U1  [num_subject];



vector [num_theta] fb_tee [num_subject];
```

```
vector [num_basis] fb1_X15_long [num_subject, KK];
vector [num_basis] fb2_X15_long [num_subject, KK];
vector [num_basis] fb3_X15_long [num_subject, KK];

vector [3] temp_ft_tee;



vector [KK] Y1_15 [num_subject];
vector [KK] Y2_15 [num_subject];
vector [num_subject] K1_15 ;
vector [num_subject] K2_15 ;

real h[num_subject];
vector  [num_pw] integral_pw1 [num_subject];
vector  [num_pw] integral_pw2 [num_subject];

real log_S1 [num_subject];
real log_S2 [num_subject];
real LL [num_subject];

for (i in 1:num_obs) {
for(p in 1:num_theta) {
ft[i,p] = cc[p]'*time_obs_Bt[i] ;
theta[i,p]= beta0[p]*age_norm[subj_long[i]] + ft[i,p] + U[subj_long[i],(2*p-1)] + U[subj_long[i],(2*p)]*time_obs[i]+ee[i,p]; // in
}
}



for(i in 1: num_conti) mu_SDM[i]= a_conti+ b_conti'* theta[conti_match[i]];

for (k in 1:num_part1) {
a_ordi_part1[k, 1]   = a_random1[k];
for (lev in 2:(num_ordi-1)) a_ordi_part1[k, lev] = a_ordi_part1[k, lev-1] + delta1[k, lev-1];
}



for (k in 1:num_part2) {
```

```
a_ordi_part2[k, 1]   = a_random2[k];
for (lev in 2:(num_ordi-1)) a_ordi_part2[k, lev] = a_ordi_part2[k, lev-1] + delta2[k, lev-1];
}


for (k in 1:num_part3) {
a_ordi_part3[k, 1]   = a_random3[k];
for (lev in 2:(num_ordi-1)) a_ordi_part3[k, lev] = a_ordi_part3[k, lev-1] + delta3[k, lev-1];
}


a_ordi_part1[3,1]=a0[1];
for(lev in 2:(num_ordi-1)) a_ordi_part1[3, lev] = a_ordi_part1[3, lev-1] + delta1[3, lev-1];

a_ordi_part2[4,1]=a0[2];
for(lev in 2:(num_ordi-1)) a_ordi_part2[4, lev] = a_ordi_part2[4, lev-1] + delta2[4, lev-1];

a_ordi_part3[11,1]=a0[3];
for(lev in 2:(num_ordi-1)) a_ordi_part3[11, lev] = a_ordi_part3[11, lev-1] + delta3[11, lev-1];

b_ordi_part1 = b_random1;
b_ordi_part2 = b_random2;
b_ordi_part3 = b_random3;


b_ordi_part1[3]  = 1;

b_ordi_part2[4]  = 1;

b_ordi_part3[11] = 1;

for (i in 1: num_obs){
for (k in 1: num_part1) updrs1_ordi_hat[i, k]= b_ordi_part1[k]* theta[i,1];
for (k in 1: num_part2) updrs2_ordi_hat[i, k]= b_ordi_part2[k]* theta[i,2];
for (k in 1: num_part3) updrs3_ordi_hat[i, k]= b_ordi_part3[k]* theta[i,3];

}
```

```
for (i in 1:num_subject){

for(p in 1: num_theta)   {
U0[i,p]= U[i, (2*p-1)];
U1[i,p]= U[i,    2*p ];
fb_tee[i, p]= cc[p]' * tee_Bt[i]; //for event 1, 2
}

exp_const[1,i]<- exp(gam[1]*gender_subj[i]+ Sum_const_nupart(nu[1], U0[i], beta0, age_norm[i]) );
exp_const[2,i]<- exp(gam[2]*gender_subj[i]+ Sum_const_nupart(nu[2], U0[i], beta0, age_norm[i]) );


for(k in 1: KK){    //KK=15 for x15

fb1_X15_long[i,k] <- cc[1] .* teeX15_long_Bt[(i-1)*15  + k] ;
fb2_X15_long[i,k] <- cc[2] .* teeX15_long_Bt[(i-1)*15  + k] ;
fb3_X15_long[i,k] <- cc[3] .* teeX15_long_Bt[(i-1)*15  + k] ;



temp_ft_tee= [sum(fb1_X15_long[i,k]), sum(fb2_X15_long[i,k]), sum(fb3_X15_long[i,k])]' ; // build temp vector for function pointV

Y1_15[i,k] = h_pw[1, pwht_X15_ind[i,k]]* exp( pointV(nu[1], U1[i], tee_X15[i,k], temp_ft_tee) );  // f1(x15) for ith subj kth poi
Y2_15[i,k] = h_pw[2, pwht_X15_ind[i,k]]* exp( pointV(nu[2], U1[i], tee_X15[i,k], temp_ft_tee) );  // f2(x15)

}

K1_15[i] <- Y1_15[i]' * c15;
K2_15[i] <- Y2_15[i]' * c15;


sd_spline = sqrt(Var_spline);
sd_e    = sqrt(Var_e);
sd_U    = sqrt(Var_U);
Sigma_U = quad_form_diag(Omega, sd_U);
sd_conti= sqrt(Var_conti);

}
```

```
for(i in 1: num_subject){


log_S2[i] =  -  K2_15[i] *tee[i] /2 * exp_const[2,i];   //Gauss_krnorod integration

log_S1[i] =  -  K1_15[i] *tee[i] /2 * exp_const[1,i];

if (event[i]==2)      h[i] = h_pw[2,tee_id[i]] * exp_const[2,i]*exp(pointV(nu[2], U1[i], tee[i], fb_tee[i]));

if (event[i]==1)      h[i] = h_pw[1,tee_id[i]] * exp_const[1,i]*exp(pointV(nu[1], U1[i], tee[i], fb_tee[i]));

if (event[i]==0)      h[i] = 1;

LL[i] = log(h[i]) + log_S1[i] + log_S2[i];

}

}

model {

vector [2*num_theta] Zero=[0,0,0,0,0,0]';


U ~  multi_normal(Zero, Sigma_U);

for(i in 1:num_obs)  ee[i] ~ normal(0, sd_e); // this constrain exclude other onstrains

for(p in 1:num_theta) {
cc[p,1] ~ normal(0, 10);   //no intercept

for (j in 2: num_basis) cc[p,j]~ normal(cc[p, j-1], sd_spline[p]);

}

SDM ~ normal(mu_SDM, sd_conti);
```

```
for(i in 1:num_obs){
for (k in 1: num_part3) updrs3[i, k] ~ ordered_logistic(updrs3_ordi_hat[i, k], a_ordi_part3[k]) ;
for (k in 1: num_part2) updrs2[i, k] ~ ordered_logistic(updrs2_ordi_hat[i, k], a_ordi_part2[k]) ;
for (k in 1: num_part1) updrs1[i, k] ~ ordered_logistic(updrs1_ordi_hat[i, k], a_ordi_part1[k]) ;

}


target +=LL;

beta0 ~ normal(0,20);
beta1 ~ normal(0,20);


a_conti ~ normal(0,100);
b_conti ~ normal(0,100);

for (l in 1:(num_ordi-2)) {
for (k in 1: num_part1)   delta1[k, l] ~ normal(0, 100) T[0,] ;
for (k in 1: num_part2)   delta2[k, l] ~ normal(0, 100) T[0,] ;
for (k in 1: num_part3)   delta3[k, l] ~ normal(0, 100) T[0,] ;
}



for (k in 1:(num_part1))   b_random1[k] ~ normal(0, 20);

for (k in 1:(num_part2))   b_random2[k] ~ normal(0, 20);

for (k in 1:(num_part3))   b_random3[k] ~ normal(0, 20);

a_random1 ~ normal(0, 20);
a_random2 ~ normal(0, 20);
a_random3 ~ normal(0, 20);

Var_U     ~ inv_gamma(0.01, 0.01);
Var_conti ~ inv_gamma(0.01, 0.01);

Omega ~ lkj_corr(2.0); //Omega=L_Omega*L_Omega'
```

```
Var_spline ~ inv_gamma(0.01,0.01);
Var_e      ~ inv_gamma(0.01, 0.01);

for(i in 1: 2){
h_pw[i] ~ gamma(0.01, 0.01);   // piecewise base havard
nu[i]  ~ normal(0,20);
gam[i] ~ normal(0,20);
}


}
```

Appendix C

Appendix for article 3

Table C.01: Simulation results with different intermittent missing proportions.

	Naive (setting 1)			Joint (setting 1)			Naive (setting 2)			Joint (setting 2)		
	BIAS %	SD	CP	BIAS %	SD	CP	BIAS %	SD	CP	BIAS %	SD	CP
Parameters for continuous outcomes												
$a_1^{(1)} = -4$	0.475	0.053	0.959	0.125	0.051	0.967	0.525	0.055	0.950	0.175	0.055	0.967
$a_1^{(2)} = -2$	0.950	0.070	0.914	0.300	0.066	0.949	0.800	0.070	0.927	0.350	0.070	0.935
$b_1^{(1)} = 1$	0.800	0.038	0.923	0.800	0.038	0.930	1.200	0.040	0.909	0.800	0.039	0.921
$b_1^{(2)} = 1.5$	0.800	0.050	0.950	0.667	0.049	0.953	1.000	0.052	0.923	0.600	0.051	0.949
$b_2^{(1)} = 1.5$	0.867	0.031	0.923	0.067	0.031	0.921	1.467	0.034	0.905	0.133	0.033	0.949
$b_2^{(2)} = 1$	0.800	0.032	0.936	0.000	0.031	0.953	1.400	0.034	0.905	0.100	0.034	0.949
Parameters for ordinal outcomes**												
$a_{3,1} = 0.1$	−3.000	0.029	0.936	−1.000	0.029	0.935	−5.000	0.031	0.932	−2.000	0.030	0.940
$a_{4,1} = -0.1$	1.000	0.027	0.936	−1.000	0.027	0.930	1.000	0.029	0.914	−0.000	0.029	0.921
$a_{5,1} = 0.2$	3.500	0.032	0.955	1.500	0.032	0.967	4.500	0.032	0.964	1.500	0.032	0.963
$a_{6,1} = -0.2$	−1.000	0.023	0.959	0.500	0.021	0.977	−1.500	00.023	0.950	−1.000	0.023	0.953
$a_{7,1} = 0.3$	0.333	0.031	0.964	−1.000	0.031	0.963	1.333	0.033	0.955	−0.333	0.032	0.944
$a_{8,1} = -0.3$	−1.000	0.023	0.950	−0.667	0.023	0.953	−0.667	0.024	0.964	−0.333	0.024	0.958
$a_{9,1} = 0.4$	0.500	0.026	0.945	0.500	0.025	0.949	1.000	0.025	0.945	0.500	0.025	0.949
$a_{10,1} = -0.4$	1.250	0.034	0.968	0.750	0.033	0.977	1.750	0.035	0.959	0.500	0.034	0.972
$a_{1,2} = 1.5$	−0.067	0.029	0.950	−0.067	0.027	0.972	−0.067	0.029	0.950	−0.133	0.029	0.963
$a_{2,2} = 1.5$	−0.200	0.030	0.959	−0.133	0.030	0.963	−0.200	0.032	0.959	−0.067	0.032	0.963
$a_{3,2} = 1.6$	−0.188	0.033	0.955	0.125	0.032	0.967	−0.188	0.036	0.950	0.062	0.035	0.958
$a_{4,2} = 1.4$	−0.071	0.031	0.927	0.000	0.032	0.926	−0.071	0.032	0.923	0.071	0.033	0.926
$a_{5,2} = 1.7$	0.471	0.038	0.936	0.294	0.037	0.949	0.706	0.040	0.936	0.353	0.040	0.940
$a_{6,2} = 1.3$	0.231	0.026	0.968	0.000	0.026	0.972	0.385	0.027	0.964	0.231	0.027	0.963
$a_{7,2} = 1.8$	0.222	0.038	0.959	0.000	0.037	0.963	0.333	0.038	0.968	0.000	0.038	0.967
$a_{8,2} = 1.2$	0.250	0.027	0.941	0.167	0.026	0.944	0.167	0.028	0.923	0.083	0.028	0.949
$a_{9,2} = 1.9$	0.053	0.034	0.936	0.000	0.034	0.940	0.158	0.036	0.941	0.053	0.036	0.935
$a_{10,2} = 1.1$	−0.273	0.033	0.973	−0.091	0.034	0.967	−0.636	0.035	0.968	−0.182	0.035	0.953
$a_{1,3} = 2$	−0.100	0.037	0.941	−0.100	0.035	0.944	−0.100	0.036	0.945	−0.100	0.035	0.953
$a_{2,3} = 2$	−0.050	0.034	0.941	0.000	0.033	0.949	−0.150	0.036	0.945	0.000	0.036	0.940
$a_{3,3} = 2.1$	−0.095	0.037	0.950	0.095	0.037	0.949	−0.095	0.040	0.945	0.095	0.039	0.944
$a_{4,3} = 1.9$	−0.105	0.033	0.955	−0.053	0.034	0.949	−0.105	0.037	0.941	−0.053	0.038	0.935
$a_{5,3} = 2.2$	0.364	0.040	0.945	0.227	0.039	0.949	0.545	0.042	0.959	0.273	0.043	0.967
$a_{6,3} = 1.8$	0.278	0.029	0.959	0.111	0.029	0.963	0.389	0.030	0.950	0.278	0.030	0.958
$a_{7,3} = 2.3$	0.261	0.041	0.950	0.087	0.040	0.958	0.391	0.042	0.964	0.130	0.041	0.967
$a_{8,3} = 1.7$	0.176	0.029	0.941	0.176	0.029	0.935	0.118	0.030	0.936	0.059	0.030	0.940

Table C.02: Parameter estimates for PPMI's three domains.

	nM-EDL				M-EDL				Motor Examination			
	MEAN	SD	95 % CI		MEAN	SD	95% CI		MEAN	SD	95% CI	
Disease Status (Latent Variable)												
Int.	0.011	0.079	−0.149	0.152	−0.045	0.089	−0.213	0.124	0.046	0.189	−0.333	0.392
Age (yr)	0.093	0.058	−0.024	0.200	0.113	0.070	−0.021	0.245	0.254	0.148	−0.031	0.531
Time (yr)	0.289	0.023	0.246	0.333	0.416	0.033	0.355	0.479	0.656	0.044	0.571	0.745
Random effects												
ρ	−0.179	0.077	-0.324	−0.027	−0.120	0.074	−0.259	0.028	−0.387	0.067	−0.516	-0.258
σ_i *	1.160	0.065	1.037	1.286	1.484	0.083	1.327	1.655	3.506	0.177	3.172	3.860
σ_s **	0.246	0.021	0.206	0.287	0.360	0.031	0.302	0.422	0.576	0.042	0.500	0.662
σ_e	0.264	0.025	0.215	0.309	0.532	0.030	0.477	0.596	1.280	0.051	1.179	1.393
Intermittent pattern of missing data												
η	0.237	0.107	0.030	0.452	−0.056	0.093	−0.234	0.125	−0.013	0.030	−0.074	0.046
Monotone pattern of missing data												
ν	0.105	0.086	−0.061	0.273	0.183	0.072	0.044	0.326	0.009	0.025	−0.038	0.058

*: random intercept
**: random slope

Table C.03: Estimated random effects.

	Mean	SD	95% CI	
$\sigma_0^{(1)}$	1.160	0.065	1.037	1.286
$\sigma_1^{(1)}$	0.246	0.021	0.206	0.287
$\sigma_0^{(2)}$	1.484	0.083	1.327	1.655
$\sigma_1^{(2)}$	0.360	0.031	0.302	0.422
$\sigma_0^{(3)}$	3.506	0.177	3.172	3.860
$\sigma_1^{(3)}$	0.576	0.042	0.500	0.662

σ_0 : random intercept.
σ_1 : random slope.
superscript (k): kth domain.

Table C.04: Estimated correlation coefficients for random effects.

nM-EDL		M-EDL		Motor		
1.000	$-0.179^{*}_{0.077}$	$0.626_{0.037}$	$-0.079_{0.077}$	$0.187_{0.053}$	$0.039_{0.076}$	$nM-EDL_{sd}$
	$[-0.324, -0.027]^{**}$	$[0.553, 0.695]$	$[-0.236, 0.064]$	$[0.082, 0.288]$	$[-0.111, 0.180]$	$95\%CI$
	1.000	$-0.017_{0.082}$	$0.665_{0.060}$	$-0.090_{0.073}$	$0.246_{0.087}$	$nM-EDL_{sd}$
		$[-0.212, 0.110]$	$[0.509, 0.751]$	$[-0.263, 0.033]$	$[0.025, 0.382]$	$95\%CI$
		1.000	$-0.120_{0.074}$	$0.370_{0.046}$	$-0.045_{0.074}$	$M-EDL_{sd}$
			$[-0.259, 0.028]$	$[0.279, 0.460]$	$[-0.192, 0.100]$	$95\%CI$
			1.000	$-0.105_{0.075}$	$0.669_{0.057}$	$M-EDL_{sd}$
				$[-0.245, 0.042]$	$[0.549, 0.772]$	$95\%CI$
				1.000	$-0.387_{0.067}$	$motor_{sd}$
					$[-0.516, -0.258]$	$95\%CI$
					1.000	$Motor$

* $-0.179_{0.077}$: correlation coefficient mean estimate=-0.179, SD=0.077.

** $[-0.324, -0.027]$: 95% CI=$[-0.324, -0.027]$ for the above element ($-0.179_{0.077}$).

Table C.05: Estimated random errors in three MDS-UPDRS parts.

	Mean	SD	95% CI	
$\sigma_e^{(1)}$	0.264	0.025	0.215	0.309
$\sigma_e^{(2)}$	0.532	0.030	0.477	0.596
$\sigma_e^{(3)}$	1.280	0.051	1.179	1.393

Table C.06: Estimated parameters in Cox model and Logistic model.

	Mean	SD	95% CI	
w	-3.081	0.095	-3.299	-2.902
γ	0.036	0.139	-0.239	0.304

Table C.07: Estimated parameters in MDS-UPDRS Part I: nM-EDL.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Cognitive Impairment	1.326	3.686	5.357	7.418	0.957	0.061	0.843	1.079
Hallucinations	3.400	5.705	6.636	8.583	0.824	0.081	0.673	0.993
Depressed Mood*	1.500	3.516	5.164	7.306	1.000	0.000	1.000	1.000
Anxious Mood	1.047	3.140	4.832	7.372	0.838	0.051	0.743	0.947
Apathy	2.103	3.912	6.038	8.393	1.201	0.073	1.065	1.345
Dopamine Dysregulation	3.934	5.692	8.394	17.407	0.778	0.091	0.605	0.958
Sleep Problem	−0.038	1.348	2.669	4.633	0.804	0.050	0.713	0.907
Daytime Sleepiness	−0.201	1.500	5.071	7.500	1.030	0.059	0.922	1.156
Pain & other sensations	−0.142	2.034	3.360	5.254	0.790	0.049	0.701	0.886
Urinary	−0.010	1.987	3.561	5.183	0.730	0.048	0.642	0.823
Constipation	0.576	2.823	4.459	9.246	0.772	0.049	0.683	0.872
Light Headedness	1.341	3.388	5.150	7.885	0.961	0.061	0.850	1.082
Fatigue	0.247	3.135	5.228	7.204	1.658	0.091	1.487	1.845

$a_{k,l}$: item k's level l.

*: item to put constrains.

Table C.08: Estimated difficulty parameters in MDS-UPDRS Part II: M-EDL.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Speech	0.721	2.291	4.657	8.769	0.789	0.046	0.703	0.884
Saliva & Drooling	0.734	1.619	2.978	4.867	0.673	0.041	0.597	0.754
Chewing & Swallowing	1.978	4.548	5.436	9.460	0.679	0.047	0.593	0.778
Eating Tasks*	0.800	3.722	7.312	17.501	1.000	0.000	1.000	1.000
Dressing	0.422	4.279	8.786	11.422	1.491	0.078	1.348	1.650
Hygiene	1.363	6.501	9.286	11.097	1.124	0.063	1.001	1.250
Handwriting	−0.602	1.402	3.186	5.395	0.717	0.040	0.641	0.799
Doing Hobbies	0.588	3.318	5.701	7.350	1.177	0.062	1.065	1.304
Turning in Bed	0.980	4.924	7.291	10.441	0.963	0.054	0.863	1.077
Tremor	−1.675	1.086	3.380	6.175	0.202	0.024	0.158	0.252
Getting out of bed	0.369	3.831	6.381	10.508	1.278	0.069	1.154	1.424
Walking & Balance	0.529	4.079	5.096	7.973	0.941	0.054	0.835	1.048
Freezing	3.743	5.827	7.765	10.113	1.159	0.076	1.010	1.311

$a_{k,l}$: item k's level l.

*: item to put constrains.

Table C.09: Estimated parameters in MDS-UPDRS Part III: Motor Examination.

	Difficulty Parameters				Discrimination Parameters			
	$a_{k,1}$	$a_{k,2}$	$a_{k,3}$	$a_{k,4}$	Mean	SD	95% CI	
Speech	−0.123	2.626	6.233	15.871	0.210	0.014	0.183	0.237
Facial Expression	−2.227	0.837	3.294	6.819	0.301	0.016	0.271	0.331
Rigidity-Neck	−0.120	1.437	4.186	6.981	0.302	0.016	0.270	0.333
Rigidity-RUE	−1.208	0.289	2.745	7.108	0.012	0.008	0.001	0.028
Rigidity-LUE	−0.567	1.656	5.093	11.130	0.618	0.026	0.568	0.669
Rigidity-RLE	−0.079	1.276	3.282	6.279	0.069	0.010	0.049	0.091
Rigidity-LLE	0.592	2.171	4.529	8.329	0.460	0.022	0.417	0.506
Finger Tapping-Right	−1.153	0.524	2.198	4.971	0.034	0.009	0.017	0.052
Finger Tapping-Left	−1.216	2.163	5.278	9.738	0.982	0.041	0.908	1.063
Hand Movement-Right	−0.718	0.984	2.716	5.964	0.048	0.009	0.029	0.067
Hand Movement-Left*	−0.500	2.894	6.149	10.839	1.000	0.000	1.000	1.000
Pronation-Right	−0.671	0.987	2.857	6.351	0.009	0.006	0.000	0.024
Pronation-Left	−0.213	2.579	5.300	9.249	0.837	0.032	0.773	0.901
Toe Tapping-Right	−0.670	1.097	2.954	5.595	0.060	0.010	0.042	0.079
Toe Tapping-Left	−0.735	1.879	4.413	8.294	0.676	0.029	0.622	0.734
Leg Agility-Right	0.160	2.141	4.183	6.517	0.098	0.011	0.078	0.120
Leg Agility-Left	0.661	3.107	5.612	8.586	0.642	0.030	0.585	0.702
Arising from chair	1.676	3.721	5.395	7.055	0.219	0.017	0.187	0.252
Gait	−0.736	2.443	4.704	6.936	0.185	0.013	0.158	0.211
Freezing of Gait	4.554	6.320	6.934	7.321	0.339	0.045	0.256	0.429
Postural Stability	2.215	3.252	3.825	6.147	0.173	0.019	0.136	0.214
Posture	−0.494	1.832	4.010	6.015	0.199	0.013	0.176	0.224
Body Bradykinesia	−2.594	0.337	2.815	7.802	0.346	0.017	0.314	0.381
Postural Tremor-Right	0.611	2.460	4.606	15.133	0.002	0.002	0.000	0.006

STAN code

```
functions {
real expit(real x){
real ps;
ps=exp(x)/(1+exp(x));
return  ps;
}
vector convert_p(real[] psi){
int N= size(psi); // use rows(psi) if define psi as vector
vector [N+1] pr;

pr[1]= psi[1];
for (k in 2:N) pr[k]=psi[k]-psi[k-1];
pr[N+1]=1-psi[N];
return pr;
}

real Sum_const_nupart (real nu_v,  real int_v, real beta_v, real cov_x_v, real u0_v) {
real sum_const_nupart;

sum_const_nupart=nu_v *( int_v+ cov_x_v * beta_v + u0_v); // vector operation scalar * vector
return sum_const_nupart;

}

real Theta_Const_nupart (vector nu_vect, vector int_vect , vector beta1_vect, real cov_x1_v,  vector u0_vect) {
real sum_const;
sum_const=nu_vect'*( int_vect+ cov_x1_v * beta1_vect  + u0_vect ); // vector operation scalar * vector

return sum_const;

}

real pointV( real nu_v,  real ft_v, real tee_v, real ut_v){
real sub_fk;
sub_fk=nu_v * (ft_v+tee_v*ut_v);
return sub_fk;
}
```

```
real cox_time_nupart( vector nu_vect, vector alpha_vect, vector u1_vect){
real time_eff;
time_eff= nu_vect'*(alpha_vect +u1_vect);
return time_eff;

}

real exp_intgl(real start_t_v, real end_t_v, real time_effect_v){
real pw_integal;
pw_integal= (exp(time_effect_v*end_t_v)-exp(time_effect_v*start_t_v))/time_effect_v;
return pw_integal;

}

real sum_intgl(int k_v, vector h0_vect, vector tee_sect_vect, real time_effect_v){
vector [k_v] piece_intgl;
vector [k_v] piece_h0;
real sum_pw;
for(i in 1: k_v) {
piece_intgl[i]=exp_intgl(tee_sect_vect[i], tee_sect_vect[i+1], time_effect_v);
piece_h0[i]=h0_vect[i];
}
sum_pw=piece_h0'*piece_intgl;
return  sum_pw;

}

vector h0_vct( int k_v, vector h0_vect){
vector [k_v] piece_h0;
for(i in 1: k_v) piece_h0[i]=h0_vect[i];

return piece_h0;
}

}
data {
int<lower=1> num_subject;
int<lower=1> NN;
```

```
int<lower=1> num_obs;
int<lower=1> num_mis;

int<lower=1> num_ordi;

int<lower=1> num_part1;
int<lower=1> num_part2;
int<lower=1> num_part3;

int <lower=1, upper=num_subject>subj_long[NN];

int <lower=-1, upper=num_ordi> Y_ordi_part1[NN, num_part1];
int <lower=-1, upper=num_ordi> Y_ordi_part2[NN, num_part2];

int <lower=-1, upper=num_ordi> Y_ordi_part3[NN, num_part3];

vector [3] a0;

real<lower=0> time[NN];

real age_norm[num_subject];
real  gender_subj[num_subject];
real<lower=0> HY_subj[num_subject];
real  duration_subj[num_subject];

real <lower=0> tee [num_subject];
int <lower=0> event[num_subject];
//int <lower=1, upper=num_pw> tee_id [num_subject];
int <lower=1, upper=NN> obs_ind [num_obs];
int <lower=1, upper=NN> mis_ind [num_mis];

int rr [NN];

int num_pw;
int <lower=1> subj_pw_ind [num_subject];
vector <lower=0> [num_pw+1] tee_pw [num_subject];
}

parameters {
```

```
vector [6] U [num_subject];
corr_matrix [6] Omega;
vector<lower=0> [6] Var_U;

vector [3] beta0;
vector [3] beta1;
vector [3]  alpha;

vector [num_part1] a_random1;
vector [num_part2] a_random2;
vector [num_part3] a_random3;

vector <lower=0>  [num_part1] b_random1;
vector <lower=0>  [num_part2] b_random2;
vector <lower=0>  [num_part3] b_random3;

vector<lower=0> [num_ordi-2] delta1[num_part1];
vector<lower=0> [num_ordi-2] delta2[num_part2];
vector<lower=0> [num_ordi-2] delta3[num_part3];

real w;
vector [3] eta;
real gam ; // survival gender num_risk
vector <lower=0> [num_pw]  h0 ;
vector [3] nu ;

}


transformed parameters {
cov_matrix [6] Sigma_U;
vector<lower=0> [6] sd_U;

ordered[num_ordi-1] a_ordi_part1[num_part1];   // this is required to use ordered_logistic
ordered[num_ordi-1] a_ordi_part2[num_part2];
ordered[num_ordi-1] a_ordi_part3[num_part3];

vector   [num_part1] b_ordi_part1 ;
```

```
vector   [num_part2] b_ordi_part2 ;
vector   [num_part3] b_ordi_part3 ;

vector [3]  theta [NN];



vector [3] U0  [num_subject];
vector [3] U1  [num_subject];


real<lower=0, upper=1>             psi1[NN, num_part1, num_ordi];
vector<lower=0, upper=1>[num_ordi] prob_y1[NN, num_part1];

real<lower=0, upper=1>             psi2[NN, num_part2, num_ordi];
vector<lower=0, upper=1>[num_ordi]  prob_y2[NN, num_part2];

real<lower=0, upper=1>           psi3[NN, num_part3, num_ordi];
vector<lower=0, upper=1>[num_ordi]  prob_y3[NN, num_part3];

vector [num_subject] cox_const;
real cox_time [num_subject]  ;

vector [num_subject] sum_pw_intgl;
real h[num_subject];

real log_S [num_subject];

real LL [num_subject];

for (i in 1:NN) {
for(p in 1:3) {

U0[subj_long[i],p]= U[subj_long[i], (2*p-1)];
U1[subj_long[i],p]= U[subj_long[i],  (2*p) ];
theta[i,p]= beta0[p] + beta1[p]*age_norm[subj_long[i]] + alpha[p]*time[i] + U0[subj_long[i],p] + U1[subj_long[i],p]*time[i]; // in
}

}
```

```
for (k in 1:num_part1) {
a_ordi_part1[k, 1]   = a_random1[k];
for (lev in 2:(num_ordi-1)) a_ordi_part1[k, lev] = a_ordi_part1[k, lev-1] + delta1[k, lev-1];
}


for (k in 1:num_part2) {
a_ordi_part2[k, 1]   = a_random2[k];
for (lev in 2:(num_ordi-1)) a_ordi_part2[k, lev] = a_ordi_part2[k, lev-1] + delta2[k, lev-1];
}


for (k in 1:num_part3) {
a_ordi_part3[k, 1]   = a_random3[k];
for (lev in 2:(num_ordi-1)) a_ordi_part3[k, lev] = a_ordi_part3[k, lev-1] + delta3[k, lev-1];
}

a_ordi_part1[3,1]=a0[1];
for(lev in 2:(num_ordi-1)) a_ordi_part1[3, lev] = a_ordi_part1[3, lev-1] + delta1[3, lev-1];

a_ordi_part2[4,1]=a0[2];
for(lev in 2:(num_ordi-1)) a_ordi_part2[4, lev] = a_ordi_part2[4, lev-1] + delta2[4, lev-1];

a_ordi_part3[11,1]=a0[3];
for(lev in 2:(num_ordi-1)) a_ordi_part3[11, lev] = a_ordi_part3[11, lev-1] + delta3[11, lev-1];

b_ordi_part1 = b_random1;
b_ordi_part2 = b_random2;
b_ordi_part3 = b_random3;

b_ordi_part1[3]  = 1;

b_ordi_part2[4]  = 1;

b_ordi_part3[11] = 1;

for (i in 1:NN) {
```

```
for (k in 1:num_part1) {
for (l in 1:(num_ordi-1)) psi1[i, k, l] = inv_logit(a_ordi_part1[k, l] - b_ordi_part1[k]*theta[i,1]);
psi1[i, k, num_ordi] = 1;
prob_y1[i, k, 1] = psi1[i, k, 1];
for (l in 2:num_ordi)    prob_y1[i, k, l] = psi1[i, k, l] - psi1[i, k, l-1];

}


for(k in 1: num_part2){
for (l in 1:(num_ordi-1)) psi2[i, k, l] = inv_logit(a_ordi_part2[k, l] - b_ordi_part2[k]*theta[i,2]);
psi2[i, k, num_ordi] = 1;
prob_y2[i, k, 1] = psi2[i, k, 1];
for (l in 2:num_ordi)  prob_y2[i, k, l] = psi2[i, k, l] - psi2[i, k, l-1];

}


for (k in 1: num_part3){
for (l in 1:(num_ordi-1)) psi3[i, k, l] = inv_logit(a_ordi_part3[k, l] - b_ordi_part3[k]*theta[i,3]);
psi3[i, k, num_ordi] = 1;
prob_y3[i, k, 1] = psi3[i, k, 1];
for (l in 2:num_ordi)  prob_y3[i, k, l] = psi3[i, k, l] - psi3[i, k, l-1];

}

}


for (i in 1:num_subject){

cox_const[i]= gam*HY_subj[i] + Theta_Const_nupart(nu, beta0, beta1, age_norm[i] , U0[i]);
cox_time[i] = cox_time_nupart(nu, alpha, U1[i]);

sum_pw_intgl[i]=sum_intgl(subj_pw_ind[i], h0, tee_pw[i], cox_time[i]) ;   // integral part

log_S[i] =  -   exp(cox_const[i])*sum_pw_intgl[i];   //piecewise summision
```

```
if (event[i]==1)      h[i] = h0[subj_pw_ind[i]] * exp(cox_const[i]+ cox_time[i]*tee[i]);

if (event[i]==0)      h[i] = 1;

LL[i] = log(h[i]) + log_S[i] ;

}

sd_U      = sqrt(Var_U);
Sigma_U   = quad_form_diag(Omega, sd_U);



}

model {

vector [num_ordi] ll_ordi1_mis_temp;
vector [num_ordi] ll_ordi2_mis_temp;
vector [num_ordi] ll_ordi3_mis_temp;

vector [6] zero=[0,0,0,0,0,0]';

U ~  multi_normal(zero, Sigma_U);

for(i in 1: NN)  {
rr[i] ~ bernoulli_logit(w + eta' * theta[i]);

if(rr[i]==0){

for (n in 1: num_part1) target += categorical_lpmf(Y_ordi_part1[i, n]| prob_y1[i,n]) ;
for (n in 1: num_part2) target += categorical_lpmf(Y_ordi_part2[i, n]| prob_y2[i,n]) ;
for (n in 1: num_part3) target += categorical_lpmf(Y_ordi_part3[i, n]| prob_y3[i,n]) ;

}



if(rr[i]==1){
```

```
for(n in 1: num_part1){
for (k in 1: num_ordi) ll_ordi1_mis_temp[k]= categorical_lpmf(k| prob_y1[i,n] );
target +=2*log_sum_exp(ll_ordi1_mis_temp) ;
}

for (n in 1: num_part2){
for (k in 1: num_ordi) ll_ordi2_mis_temp[k]= categorical_lpmf(k| prob_y2[i,n] );
target +=2*log_sum_exp(ll_ordi2_mis_temp) ;
}

for (n in 1: num_part3){
for (k in 1: num_ordi) ll_ordi3_mis_temp[k]= categorical_lpmf(k| prob_y3[i,n] );
target +=2*log_sum_exp(ll_ordi3_mis_temp) ;
}

}



}

target +=LL;

beta0    ~ normal(0,20);
beta1    ~ normal(0,20);

alpha   ~ normal(0,20);


w ~ normal(0,20);
eta ~ normal(0,20);

for (l in 1:(num_ordi-2)) {

for (k in 1: num_part1)   delta1[k, l] ~ normal(0, 100) T[0,] ;
for (k in 1: num_part2)   delta2[k, l] ~ normal(0, 100) T[0,] ;
for (k in 1: num_part3)   delta3[k, l] ~ normal(0, 100) T[0,] ;
}
```

```
b_random1 ~ normal(0, 20);
b_random2 ~ normal(0, 20);
b_random3 ~ normal(0, 20);

a_random1 ~ normal(0, 20);
a_random2 ~ normal(0, 20);
a_random3 ~ normal(0, 20);


Var_U     ~ inv_gamma(0.01, 0.01);

Omega ~ lkj_corr(2.0); //Omega=L_Omega*L_Omega'


h0  ~ gamma(0.1, 0.1);
nu  ~ normal(0,20);
gam ~ normal(0,20);
}
```

References

[1] Alamri, Y., M. Pascoe, J. Dalrymple-Alford, T. Anderson, and M. MacAskill (2017). Through new eyes: gaze analysis of parkinson disease patients performing the symbol digit modalities test (p6. 006). *Neurology 88*(16 Supplement), P6–006.

[2] Andersen, P. K., S. Z. Abildstrom, and S. Rosthøj (2002). Competing risks as a multi-state model. *Statistical Methods in Medical Research 11*(2), 203–215.

[3] Auclair-Ouellet, N., P. Lieberman, and O. Monchi (2017). Contribution of language studies to the understanding of cognitive impairment and its progression over time in parkinson's disease. *Neuroscience & Biobehavioral Reviews 80*(9), 657–672.

[4] Balsis, S., A. A. Unger, J. F. Benge, L. Geraci, and R. S. Doody (2012). Gaining precision on the alzheimer's disease assessment scale-cognitive: a comparison of item response theory-based scores and total scores. *Alzheimer's & Dementia 8*(4), 288–294.

[5] Bega, D., S. Kim, Y. Zhang, J. Elm, J. Schneider, R. Hauser, A. Fraser, and T. Simuni (2015). Predictors of functional decline in early parkinson's disease: Net-pd ls1 cohort. *Journal of Parkinson's Disease 5*(4), 773–782.

[6] Bellou, V., L. Belbasis, I. Tzoulaki, E. Evangelou, and J. P. Ioannidis (2016). Environmental risk factors and parkinson's disease: an umbrella review of meta-analyses. *Parkinsonism & Related Disorders 23*, 1–9.

[7] Bjornestad, A., E. B. Forsaa, K. F. Pedersen, O.-B. Tysnes, J. P. Larsen, and G. Alves (2016). Risk and course of motor complications in a population-based incident parkinson's disease cohort. *Parkinsonism & Related Disorders 22*, 48–53.

[8] Buatois, S., S. Retout, N. Frey, and S. Ueckert (2017). Item response theory as an efficient tool to describe a heterogeneous clinical rating scale in de novo idiopathic parkinson's disease patients. *Pharmaceutical Research 34*(10), 2109–2118.

[9] Caslake, R., K. Taylor, N. Scott, J. Gordon, C. Harris, K. Wilde, A. Murray, and C. Counsell (2013). Age-, gender-, and socioeconomic status-specific incidence of parkinson's disease and parkinsonism in north east scotland: the pine study. *Parkinsonism & Related Disorders 19*(5), 515–521.

[10] Celeux, G., F. Forbes, C. P. Robert, D. M. Titterington, et al. (2006). Deviance information criteria for missing data models. *Bayesian Analysis 1*(4), 651–673.

[11] Chaudhuri, K. R., D. G. Healy, and A. H. Schapira (2006). Non-motor symptoms of parkinson's disease: diagnosis and management. *The Lancet Neurology 5*(3), 235–245.

[12] Chaudhuri, K. R. and Y. Naidu (2008). Early parkinson's disease and non-motor issues. *Journal of Neurology 255*(5), 33.

[13] Chi, Y.-Y. and J. G. Ibrahim (2006). Joint models for multivariate longitudinal and multivariate survival data. *Biometrics 62*(2), 432–445.

[14] Dantan, E., P. Joly, J.-F. Dartigues, and H. Jacqmin-Gadda (2011). Joint model with latent state for longitudinal and multistate data. *Biostatistics 12*(4), 723–736.

[15] Daskalakis, C., N. M. Laird, and J. M. Murphy (2002). Regression analysis of multiple-source longitudinal outcomes: a "stirling county" depression study. *American Journal of Epidemiology 155*(1), 88–94.

[16] De Gruttola, V. and X. M. Tu (1994). Modelling progression of cd4-lymphocyte count and its relationship to survival time. *Biometrics 50*(4), 1003–1014.

[17] Demidenko, E. (2013). *Mixed models: theory and applications with R.* John Wiley & Sons.

[18] Diggle, P. (2002). *Analysis of Longitudinal Data.* Oxford University Press.

[19] Duane, S., A. Kennedy, B. Pendleton, and D. Roweth (1987). Hybrid Monte Carlo. *Physics Letters B 195*(2), 216–222.

[20] Dunson, D. (2003). Dynamic latent trait models for multidimensional longitudinal data. *Journal of the American Statistical Association 98*(463), 555–563.

[21] Eilers, P. H. and B. D. Marx (1996). Flexible smoothing with b-splines and penalties. *Statistical Science 11*(2), 89–102.

[22] Ekholm, A. and C. Skinner (1998). The muscatine children's obesity data reanalysed using pattern mixture models. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 47*(2), 251–263.

[23] Elashoff, R. M., G. Li, and N. Li (2007). An approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in Medicine 26*(14), 2813–2835.

[24] Elashoff, R. M., G. Li, and N. Li (2008). A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics 64*(3), 762–771.

[25] Fang, J. Y., A. Pérez, C. W. Christine, M. Leehey, M. J. Aminoff, J. T. Boyd, J. C. Morgan, R. Dhall, A. P. Nicholas, I. Bodis-Wollner, et al. (2015). Parkinson's disease severity and use of dopaminergic medications. *Parkinsonism & Related Disorders 21*(3), 297–299.

[26] Ferguson, L. W., A. H. Rajput, and A. Rajput (2016). Early-onset vs. late-onset parkinson's disease: a clinical-pathological study. *Canadian Journal of Neurological Sciences 43*(1), 113–119.

[27] Fine, J., J. Duff, R. Chen, W. Hutchison, A. M. Lozano, and A. E. Lang (2000). Long-term follow-up of unilateral pallidotomy in advanced parkinson's disease. *New England Journal of Medicine 342*(23), 1708–1714.

[28] Fine, J. P. and R. J. Gray (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association 94*(446), 496–509.

[29] Fitzmaurice, G., M. Davidian, G. Verbeke, and G. Molenberghs (2008). *Longitudinal Data Analysis*. CRC press.

[30] Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics 19*, 1–67.

[31] Gelman, A. (2015). The connection between varying treatment effects and the crisis of unreplicable research: a bayesian perspective. *Journal of Management 41*(2), 632–643.

[32] Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2014). *Bayesian Data Analysis*. Chapman & Hall/CRC Boca Raton, FL, USA.

[33] Glas, C. A., H. Geerlings, M. A. van de Laar, and E. Taal (2009). Analysis of longitudinal randomized clinical trials using item response models. *Contemporary Clinical Trials 30*(2), 158–170.

[34] Gnaldi, M. (2017). A multidimensional irt approach for dimensionality assessment of standardised students' tests in mathematics. *Quality & Quantity 51*(3), 1167–1182.

[35] Goetz, C. G., B. C. Tilley, S. R. Shaftman, G. T. Stebbins, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, M. B. Stern, R. Dodel, et al. (2008). Movement disorder society-

sponsored revision of the unified parkinson's disease rating scale (mds-updrs): Scale presentation and clinimetric testing results. *Movement Disorders 23*(15), 2129–2170.

[36] Gorter, R., J.-P. Fox, and J. W. Twisk (2015). Why item response theory should be used for longitudinal questionnaire data analysis in medical research. *BMC Medical Research Methodology 15*(1), 15–55.

[37] Gray, R. J. (1992). Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association 87*(420), 942–951.

[38] Guimaraes, P., K. Kieburtz, C. G. Goetz, J. J. Elm, Y. Y. Palesch, P. Huang, B. Ravina, C. M. Tanner, and B. C. Tilley (2005). Non-linearity of parkinson's disease progression: implications for sample size calculations in clinical trials. *Clinical Trials 2*(6), 509–518.

[39] Hastie, T. and R. Tibshirani (1993). Varying-coefficient models. *Journal of the Royal Statistical Society. Series B (Methodological) 55*(4), 757–796.

[40] He, B. and S. Luo (2016). Joint modeling of multivariate longitudinal measurements and survival data with applications to parkinson's disease. *Statistical Methods in Medical Research 25*(4), 1346–1358.

[41] Henderson, R., P. Diggle, and A. Dobson (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics 1*(4), 465–480.

[42] Hoffman, M. D. and A. Gelman (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *Journal of Machine Learning Research 15*(1), 1593–1623.

[43] Huang, P., R. F. Woolson, and P. C. O'Brien (2008). A rank-based sample size method for multiple outcomes in clinical trials. *Statistics in Medicine 27*(16), 3084–3104.

[44] Ibrahim, J. G., M.-H. Chen, and S. R. Lipsitz (2001). Missing responses in generalised linear mixed models when the missing data mechanism is nonignorable. *Biometrika 88*(2), 551–564.

[45] Investigators, N. N.-P. et al. (2006). A randomized, double-blind, futility clinical trial of creatine and minocycline in early parkinson disease. *Neurology 66*(5), 664–671.

[46] Investigators, N. N.-P. et al. (2007). A randomized clinical trial of coenzyme q10 and gpi-1485 in early parkinson disease. *Neurology 68*(1), 20–28.

[47] Jellinger, K. A. (2014). The pathomechanisms underlying parkinson's disease. *Expert Review of Neurotherapeutics 14*(2), 199–215.

[48] Kieburtz, K., B. C. Tilley, J. J. Elm, D. Babcock, R. Hauser, G. W. Ross, A. H. Augustine, E. U. Augustine, M. J. Aminoff, I. G. Bodis-Wollner, et al. (2015). Effect of creatine monohydrate on clinical progression in patients with parkinson disease: a randomized clinical trial. *JAMA 313*(6), 584–593.

[49] Kleinman, K. P. and J. G. Ibrahim (1998). A semiparametric bayesian approach to the random effects model. *Biometrics*, 921–938.

[50] Kowal, S. L., T. M. Dall, R. Chakrabarti, M. V. Storm, and A. Jain (2013). The current and projected economic burden of parkinson's disease in the united states. *Movement Disorders 28*(3), 311–318.

[51] Krack, P., A. Batir, N. Van Blercom, S. Chabardes, V. Fraix, C. Ardouin, A. Koudsie, P. D. Limousin, A. Benazzouz, J. F. LeBas, et al. (2003). Five-year follow-up of bilateral stimulation of the subthalamic nucleus in advanced parkinson's disease. *New England Journal of Medicine 349*(20), 1925–1934.

[52] Lambert, P. and F. Vandenhende (2002). A copula-based model for multivariate non-normal longitudinal data: analysis of a dose titration safety study on a new antidepressant. *Statistics in Medicine 21*(21), 3197–3217.

[53] Lavori, P. W., R. Dawson, and T. B. Mueller (1994). Causal estimation of time-varying treatment effects in observational studies: application to depressive disorder. *Statistics in Medicine 13*(11), 1089–1100.

[54] Levy, G. (2007). The relationship of parkinson disease with aging. *Archives of Neurology 64*(9), 1242–1246.

[55] Levy, G., M.-X. Tang, L. J. Cote, E. D. Louis, B. Alfaro, H. Mejia, Y. Stern, and K. Marder (2000). Motor impairment in pd relationship to incident dementia and age. *Neurology 55*(4), 539–544.

[56] Little, R. J. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association 88*(421), 125–134.

[57] Little, R. J. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association 90*(431), 1112–1121.

[58] Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Routledge.

[59] Lunn, D., A. Thomas, N. Best, and D. Spiegelhalter (2000). WinBUGS - a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing 10*(4), 325–337.

[60] Luo, S. and J. Wang (2014). Bayesian hierarchical model for multiple repeated measures and survival data: an application to parkinson's disease. *Statistics in Medicine 33*(24), 4279–4291.

[61] Ma, G., A. B. Troxel, and D. F. Heitjan (2005). An index of local sensitivity to nonignorable drop-out in longitudinal modelling. *Statistics in Medicine 24*(14), 2129–2150.

[62] Marek, K., D. Jennings, S. Lasch, A. Siderowf, C. Tanner, T. Simuni, C. Coffey, K. Kieburtz, E. Flagg, S. Chowdhury, et al. (2011). The parkinson progression marker initiative (ppmi). *Progress in Neurobiology 95*(4), 629–635.

[63] McCulloch, C. E. and J. M. Neuhaus (2011). Misspecifying the shape of a random effects distribution: why getting it wrong may not matter. *Statistical Science 26*(3), 388–402.

[64] McDonald, R. P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement 24*(2), 99–114.

[65] McGhee, D. J., P. L. Royle, P. A. Thompson, D. E. Wright, J. P. Zajicek, and C. E. Counsell (2013). A systematic review of biomarkers for disease progression in parkinson's disease. *BMC neurology 13*(1), 35.

[66] Michell, A., S. Lewis, T. Foltynie, and R. Barker (2004). Biomarkers and parkinson's disease. *Brain 127*(8), 1693–1705.

[67] Miller, D. B. and J. P. O'Callaghan (2015). Biomarkers of parkinson's disease: present and future. *Metabolism 64*(3), 40–46.

[68] Molenberghs, G., M. G. Kenward, and E. Lesaffre (1997). The analysis of longitudinal ordinal data with nonrandom drop-out. *Biometrika 84*(1), 33–44.

[69] Moreno-Betancur, M. and M. Chavance (2016). Sensitivity analysis of incomplete longitudinal data departing from the missing at random assumption: Methodology and application in a clinical trial with drop-outs. *Statistical Methods in Medical Research 25*(4), 1471–1489.

[70] Nicoletti, A., R. Vasta, G. Mostile, G. Nicoletti, G. Arabia, G. Iliceto, P. Lamberti, R. Marconi, L. Morgante, P. Barone, et al. (2017). Gender effect on non-motor symptoms in parkinson's disease: are men more at risk? *Parkinsonism & Related Disorders 35*, 69–74.

[71] O'Brien, L. and G. Fitzmaurice (2004). Analysis of longitudinal multiple-source binary data using generalized estimating equations. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 53*(1), 177–193.

[72] Osteen, P. (2010). An introduction to using multidimensional item response theory to assess latent factor structures. *Journal of the Society for Social Work and Research 1*(2), 66–82.

[73] Pinheiro, J., C. Liu, and Y. Wu (2001). Efficient algorithms for robust estimation in linear mixed-effects models using the multivariate t distribution. *Journal of Computational and Graphical Statistics 10*(2), 249–276.

[74] Pocock, S. J. (1997). Clinical trials with multiple outcomes: a statistical perspective on their design, analysis, and interpretation. *Controlled Clinical Trials 18*(6), 530–545.

[75] Pringsheim, T., N. Jette, A. Frolkis, and T. D. Steeves (2014). The prevalence of parkinson's disease: a systematic review and meta-analysis. *Movement Disorders 29*(13), 1583–1590.

[76] Ramaker, C., J. Marinus, A. M. Stiggelbout, and B. J. Van Hilten (2002). Systematic evaluation of rating scales for impairment and disability in parkinson's disease. *Movement Disorders 17*(5), 867–876.

[77] Reeve, A., E. Simcox, and D. Turnbull (2014). Ageing and parkinson's disease: why is advancing age the biggest risk factor? *Ageing Research Reviews 14*, 19–30.

[78] Reise, S. P., T. M. Moore, and M. G. Haviland (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment 92*(6), 544–559.

[79] Ribaudo, H. J., S. G. Thompson, and T. G. Allen-Mersh (2000). A joint analysis of quality of life and survival using a random effect selection model. *Statistics in Medicine 19*(23), 3237–3250.

[80] Rubin, D. B. (1976). Inference and missing data. *Biometrika 63*(3), 581–592.

[81] Salat, D., A. J. Noyce, A. Schrag, and E. Tolosa (2016). Challenges of modifying disease progression in prediagnostic parkinson's disease. *The Lancet Neurology 15*(6), 637–648.

[82] Schmidt, S., O. Zlatkin-Troitschanskaia, and J.-P. Fox (2016). Pretest-posttest-posttest multilevel irt modeling of competence growth of students in higher education in germany. *Journal of Educational Measurement 53*(3), 332–351.

[83] Shoulson, I. (1998). Datatop: a decade of neuroprotective inquiry. *Annals of neurology 44*(S1), S160–S166.

[84] Simuni, T., J. D. Long, C. Caspell-Garcia, C. S. Coffey, S. Lasch, C. M. Tanner, D. Jennings, K. D. Kieburtz, and K. Marek (2016). Predictors of time to initiation of symptomatic therapy in early parkinson's disease. *Annals of Clinical and Translational Neurology 3*(7), 482–494.

[85] Song, X., M. Davidian, and A. A. Tsiatis (2002). A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics 58*(4), 742–753.

[86] Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. Van Der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(4), 583–639.

[87] Stan Development Team (201). *Stan Modeling Language Users Guide and Reference Manual, Version 2.17.0.*

[88] Taylor Tavares, A. L., G. S. Jefferis, M. Koop, B. C. Hill, T. Hastie, G. Heit, and H. M. Bronte-Stewart (2005). Quantitative measurements of alternating finger tapping in parkinson's disease correlate with updrs motor disability and reveal the improvement in fine motor control from medication and deep brain stimulation. *Movement Disorders 20*(10), 1286–1298.

[89] Tison, F., F. Yekhlef, V. Chrysostome, E. Balestre, N. P. Quinn, W. Poewe, and G. K. Wenning (2002). Parkinsonism in multiple system atrophy: Natural history, severity (updrs-iii), and disability assessment compared with parkinson's disease. *Movement Disorders 17*(4), 701–709.

[90] Troxel, A. B., S. R. Lipsitz, and D. P. Harrington (1998). Marginal models for the analysis of longitudinal measurements with nonignorable non-monotone missing data. *Biometrika 85*(3), 661–672.

[91] Vehtari, A., A. Gelman, and J. Gabry (2017). Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing 27*(5), 1413–1432.

[92] Verbaan, D., J. Marinus, M. Visser, S. M. van Rooden, A. M. Stiggelbout, H. A. Middelkoop, and J. J. van Hilten (2007). Cognitive impairment in parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry 78*(11), 1182–1187.

[93] Verbeke, G., S. Fieuws, G. Molenberghs, and M. Davidian (2014). The analysis of multivariate longitudinal data: A review. *Statistical Methods in Medical Research 23*(1), 42–59.

[94] Verbeke, G. and E. Lesaffre (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association 91*(433), 217–221.

[95] Verbeke, G., G. Molenberghs, H. Thijs, E. Lesaffre, and M. G. Kenward (2001). Sensitivity analysis for nonrandom dropout: a local influence approach. *Biometrics 57*(1), 7–14.

[96] Verhagen, J. and J.-P. Fox (2013). Longitudinal measurement in health-related surveys. a bayesian joint growth model for multivariate ordinal responses. *Statistics in Medicine 32*(17), 2988–3005.

[97] Vu, T. C., J. G. Nutt, and N. H. Holford (2012). Progression of motor and nonmotor features of parkinson's disease and their response to treatment. *British Journal of Clinical Pharmacology 74*(2), 267–283.

[98] Wang, J. and S. Luo (2017). Multidimensional latent trait linear mixed model: an application in clinical studies with multivariate longitudinal outcomes. *Statistics in Medicine 36*(20), 3244–3256.

[99] Wang, J., S. Luo, and L. Li (2017, October). Dynamic prediction for multiple repeated measures and event time data: An application to parkinson's disease. *The Annals of Applied Statistics 11*(3), 1787–1809.

[100] Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research 11*(Dec), 3571–3594.

[101] Weintraub, D., C. L. Comella, and S. Horn (2008). Parkinson's disease–part 1: Pathophysiology, symptoms, burden, diagnosis, and assessment. *Am J Manag Care 14*(2 Suppl), 40–48.

[102] Wu, H. and J.-T. Zhang (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-effects Modeling Approaches*. John Wiley & Sons.

[103] Wu, L., X. J. Hu, and H. Wu (2007). Joint inference for nonlinear mixed-effects models and time to event at the presence of missing data. *Biostatistics 9*(2), 308–320.

[104] Wu, M. C. and R. J. Carroll (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics 44*(1), 175–188.

[105] Yang, S. and R. L. Prentice (2011). Estimation of the 2-sample hazard ratio function using a semiparametric model. *Biostatistics 12*(2), 354–368.

[106] Yuan, Y. and G. Yin (2010). Bayesian quantile regression for longitudinal studies with nonignorable missing data. *Biometrics 66*(1), 105–114.

[107] Zeng, D., J. Cai, et al. (2005). Asymptotic results for maximum likelihood estimators in joint analysis of repeated measurements and survival time. *The Annals of Statistics 33*(5), 2132–2163.