Modeling developmental trajectories with nonrandomly missing data: Investigating trajectories of frailty using data from the Manitoba Follow-up Study

By

Chendong Li

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
In partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Community Health Sciences
University of Manitoba
Winnipeg

Abstract

Frailty is a recognized syndrome characterized by age-associated declines in physiological reserve and function in multiple organ systems. Baseline frailty and longitudinal change in frailty were independently associated with all-cause mortality. Previous studies have shown large within- and between-individual heterogeneity in the frailty trajectories. It is important to untangle these diverse patterns of frailty and to identify factors associated with these patterns. Group-based trajectory analysis (GBTA) is often used for longitudinal trajectory analysis, which allows missing values. However, the conventional GBTA assumes the missing mechanism is either missing completely at random or missing at random. When the missingness is not at random, such as when missing is due to death or extreme frailty, the conventional GBTA can result in biased estimates, and the extended GBTA is recommended. The extended GBTA can additionally estimate the attrition probability while considering the nonignorable missingness. The aim of this thesis is to compare the performances of the conventional and extended GBTA in different missing mechanisms and investigate their impacts on the identification of frailty trajectory patterns.

Objectives

The objectives of the thesis are: (1) to compare the extended GBTA with conventional GBTA in estimating trajectory classes under different conditions and missing data mechanisms; (2) to investigate the trajectories of frailty among older men and factors associated with these frailty trajectories using data from the Manitoba Follow-up Study (MFUS).

Methods

A series of simulation studies was conducted to compare the conventional and extended GBTA models. Data were generated from two trajectory classes under two scenarios: (1) initially well-separated classes and (2) initially not well-separated classes. Under each scenario, different missing data mechanisms were considered, including missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). Performance criteria such as the absolute error, relative bias, and mean squared error (MSE) were used for evaluating the precision of growth parameters (e.g., intercept and slope) and trajectory class proportions generated from the conventional and extended GBTA models.

Data from the MFUS, which is the longest-running cardiovascular and ageing longitudinal cohort study in Canada, were used for the investigation of frailty trajectory among older populations. A yearly frailty index incorporating 29 activities of daily living and 39 comorbidities was constructed. The frailty index measures variation in health status as the proportion of deficits present of the 68 deficits considered (range 0-1). Both the conventional and extended GBTA methods were applied to depict the trajectories of the frailty index. Multinomial logistic regression was used to examine how living arrangements and marital status were associated with frailty trajectories.

Results

Our simulation studies indicated that when the latent trajectory classes were not well-separated, the extended GBTA better recovered trajectory estimates than the conventional GBTA across a range of missing data mechanisms and attrition rates. When the trajectories were well-separated, the extended GBTA performs as effective as the conventional GBTA in recovering trajectory classes.

The extended GBTA identified four distinctive frailty trajectories: stable with a low level (18.90%), stable with moderate level (39.43%), increase with moderate level initially (35.44%), and increase with high level initially (6.22%). The conventional GBTA model revealed similar patterns of frailty trajectories but with varying shape or sizes for each trajectory class. The extended model is preferred evidenced by the findings of the simulation study. Participants' age was found to be associated with their frailty trajectory membership, while marital status and living arrangement were not.

Conclusion and Significance

The simulation study evaluates the performance of the extension of GBTA model in accurately recovering trajectories in the presence of non-random missing data. It outlines the advantages of the extended GBTA over the conventional GBTA, particularly when the latent classes are not initially well-separated. This study makes a significant contribution to applied researchers who are working with latent classes in longitudinal studies with non-random missing data.

By applying both the conventional and extended GBTA methods to the MFUS dataset and comparing their differences, researchers can gain a better understanding of the heterogeneity of

frailty growth within the population over time, especially considering attrition due to death. Our finding on factors associated with frailty trajectories suggests that further research is needed to identify the characteristics of individuals and communities that drive these different frailty trajectories. This knowledge would be important for tailoring and targeting future prevention initiatives to have the most impact.

Keywords

Group-Based Trajectory Analysis, Missing Data, Frailty Index, Frailty Trajectory, Manitoba Follow-up Study

Acknowledge

During the course of my master's thesis, I had the privilege of receiving support and assistance from numerous individuals who generously shared their knowledge, time, and expertise. Without their contributions, this work would not have been possible. I would like to express my heartful gratitude and appreciation to all of them.

First and foremost, I extend my heartfelt gratitude to my advisor, Dr. Depeng Jiang, for his unwavering guidance, support, and mentorship. His insights, patience, and commitment to academic excellence have played a pivotal role in shaping this research.

I am also immensely grateful to my other committee members, Dr. Robert B Tate and Dr. Philip St. John. Their invaluable feedback, constructive criticism, and encouragement have been crucial throughout this endeavor. I truly appreciate their wisdom and dedication ensuring that this thesis met the highest academic standards.

I am deeply grateful to the Visual and Automated Disease Analytics (VADA) Trainee Scholarships funded by the Natural Science and Engineering Research Council of Canada (NSERC), the Canadian Institute of Health Research (CIHR) Project Grant, and the Rady Faculty of Community Health Sciences Graduate Studentship for their generous financial support. These fundings have been instrumental in providing the necessary resources to successfully complete the research.

I would like to acknowledge the Department of Community Health Sciences, the George & Fay Yee Centre for Healthcare Innovation, and the Manitoba Follow-Up Study Team for providing an enriching and supportive research environment. The knowledge, tools, and resources provided by these institutions have been invaluable in the pursuit of this study.

I would like to give a special mention to Dr. Jiang's Lab for their unrelenting support, camaraderie, and teamwork. Working among such a dedicated group of individuals has been an inspiring experience.

Lastly, but certainly not least, I want to express my profound gratitude to my family back in China. Their unwavering belief in me, constant encouragement, and boundless love have been my pillars of strength throughout this academic journey.

In culmination, while the responsibility for any errors or omissions in this thesis is solely my own, the credit for its achievements is shared with all those who have walked this journey alongside me. Thank you.

Contents

1. Introduction

Canada, similar to many other developed nations, is experiencing a notable demographic shift characterized by an aging population. Statistics Canada (2017) projected that by the year 2030, nearly one in four Canadians could be 65 years of age or older. The rise in the elderly population often corresponds with an increased prevalence of chronic diseases and multi-morbidity, leading to increased healthcare usage and costs. These trends call for research into how consequences related to aging affect the health status of seniors.

The study of change in health status over time is important as people age, which can help us to understand the effect of these changes on adverse outcomes in the population. Using a longitudinal study design, in which data are collected at two or more points in time, can provide greater insight into these changes. It has become a staple of gerontological research methods, providing myriad insights into the general processes and heterogeneity of ageing.

In terms of the statistical methods for analyzing longitudinal data of change, multilevel modeling and latent curve modeling are commonly used (Bollen & Curran, 2006). Both methods assume one population trajectory with individual variations. Starting from the 1990s, more advanced statistical techniques emerged, as researchers were eager to investigate distinctive latent trajectories among the population. Group-based trajectory analysis (GBTA), proposed by Nagin (1999), was one of the first scholarly papers to answer the call. It used unobserved latent class variables to identify distinctive trajectories among the population. This method allows for estimating not only the trajectory shape and size for each latent class but also the risk factors associated with the trajectory group membership.

Among the challenges in trajectory analysis, attrition or missingness is the one of the most prominent. The conventional GBTA can account for attrition when its mechanism is missing completely at random (MCAR) or missing at random (MAR). However, when the mechanism of attrition is missing not at random (MNAR), using the conventional GBTA method can result in biased estimates of trajectory classes. Haviland et al. (2011) proposed an extended version of GBTA to encompass the nonignorable missingness or attrition. The first aim of this thesis research is to compare the extended GBTA with conventional GBTA in terms of trajectory class recovery when true classes exist.

Frailty has attracted increasing attention in the medical literature. It is a distinct status represented by a slowly diminishing physiological reserve across different organ systems, along with increased vulnerability to physical and/or mental stressors during the process of ageing. The general scientific consensus is that frailty is a dynamic state and there is a significant heterogeneity in frailty trajectories, both within and between individuals. Thus, the frailty change over time should be investigated from longitudinal perspectives. The second aim of this thesis research is to identify distinct frailty trajectories in an elderly population and to examine factors associated with these trajectories.

The MFUS is the longest longitudinal study within Canada investigating cardiovascular disease and ageing (Tate et al., 2015). Initiated on July 1948, the MFUS recruited 3,983 men from the Royal Canadian Air Force at the end of World War II. The initial research objective focused on the role of abnormalities in routine electrocardiograms in the prediction of subsequent diagnoses of cardiovascular disease. The research focus of MFUS was expanded in 1996 to examine the roles of physical, mental, and social functioning in support of healthy and successful ageing. By 2019, 107 out of the original cohort members were still alive. With the longitudinal data over 20 years from MFUS, we will be able to examine the distinctive trajectory classes of frailty, as well as factors associated with frailty trajectories.

In Section 2, literature on methods used in trajectory analysis, missing data patterns and strategies, and frailty analysis is reviewed. The objectives of the thesis are also presented in this section. Section 3 introduces the simulation study, where model performance under a range of conditions is evaluated and compared. Section 4 offers the model estimation of the conventional and the extended GBTA for analyzing the frailty trajectory based on the MFUS. In Section 5, we discuss ethical considerations, the significance of the thesis work, its limitations, and future directions.

2. Literature Review

In this section, we review the literature on developmental trajectory models for ageing research, frailty measurement and frailty trajectory analysis, and the Manitoba Follow-Up Study (MFUS) in ageing research. In Section 2.1, we review statistical methods for trajectory analysis in ageing research. We discuss different missing mechanisms and their impacts on estimations of trajectories. In Section 2.2, we review the frailty definitions and commonly used measurement models of frailty. We review frailty prevalence, the consequences of frailty, and the factors associated with frailty. In Section 2.3, we introduce the MFUS and review some trajectory analyses and research on frailty using data from the MFUS. In Section 2.4, we summarize the research gaps in the literature and the thesis objectives.

2.1 Developmental trajectory models for ageing research

2.1.1 Longitudinal studies for ageing research

Ageing is a complicated process. At the personal level, ageing is accompanied by multifaceted changes throughout an individual's lifetime. The effects of complex interactions among changing biological, psychological, and social factors can take years to unfold (Raina et al., 2009). At the population level, the variance in the ageing process across individuals further impedes understanding. If future interventions and policies are to fulfill the general welfare aim of improving both the quality and length of life, researchers in this field will need to achieve an integrative understanding of both between-person and within-person sources of information. This understanding must be based on results derived from appropriate designs and tenable assumptions, requiring different levels of analysis – a longitudinal study design (Hofer & Sliwinski, 2006). Longitudinal studies are efficient tools for understanding the process of ageing and changes in quality of life. These studies appeal to researchers in ageing analysis by not only depicting the overall trajectory of the population but also including the heterogeneous information across subjects. Their follow-up nature contributes to understanding ageing and geriatric syndrome. There has been an increasing number of longitudinal studies over the past several decades focusing on older adults, assessing biological pathways that lead to adverse health outcomes, social and behavioral outcomes, and disease endpoints (Newman, 2010). The data sets collected from these longitudinal studies have provided a wide array of measures and large representative samples. Although researchers initiating the longitudinal study may have limited and specific aims in mind, the research hypotheses can reach beyond the original study

goals once the data set exists. The accessibility of such rich data sources for students and junior faculty, who might otherwise lack the resources needed for the initiation of large-scale studies, provides an unmatched opportunity for scientific contributions to ageing research.

The longitudinal design also presents researchers with a number of challenges. These studies capture information at both individual and population levels, necessitating sophisticated statistical methods to illustrate the heterogeneity among populations and correlated errors of repeated measurements. Attrition or missing data is a common issue in longitudinal studies, particularly when the focus is on older adults, who are more likely to encounter health and functional problems that limit data collection (Hardy et al., 2009). In addition, some individuals may be lost to follow-up until the end of the study, while others might miss one measurement but return to the study for the next, resulting in a non-monotonic attrition pattern (Ibrahim & Molenberghs, 2009). If missingness is not handled properly, it could result in biased estimates and misleading results (Hardy et al., 2009).

2.1.2 Growth curve modeling for longitudinal data analysis

The past decade has witnessed a rapid rise to a bunch of new statistical methods for analyzing longitudinal data. The growth curve models (Curran et al., 2010) are a popular one. They refer to a set of statistical methods that can estimate the between-person differences in within-person change (Bollen & Curran, 2006; Goldstein et al., 1993). The within-person patterns of change indicate the time trends or latent trajectories, and these trajectories may vary from person to person with different characteristics. This is achieved by including both the fixed and random effects in the growth curve models. The fixed effect part of the growth curve model captures the mean trajectory for the population, while the random effect part represents the individual deviations from the population mean trajectory.

Growth models can be estimated by using all available information, including data from individuals with missing assessments. There are two commonly used procedures to handle missing data in estimating growth curve models. The first is the use of full information maximum likelihood (FIML) to estimate the parameters. In the FIML, the model is estimated by grouping the individual contributions of each case, which results in the observations with many data points weighted more heavily than those with a small number of data points (Curran et al., 2010). Another approach is Multiple Imputation, which is a two-stage process (Rubin, 2008).

This method first constructs a data set by replacing missing data with simulated values multiple times and then fits the postulated growth model to the constructed data sets. Although both the FIML estimation and the two-stage MI handle the missing data effectively, resulting in unbiased estimations, the mechanism of the missing data that these two methods could account for must be characterized as missing completely at random or missing at random. Rubin (1976) initially identified three types of attrition or missingness: (1) missing completely at random (MCAR, where attrition is unrelated to the data), (2) missing at random (MAR, where attrition depends on the observed data but not on the unobserved), and (3) missing not at random (MNAR, where attrition depends on unobserved data). When the data is missing not at random, standard growth models cannot be used, and more advanced methods must by employed (Rubin, 2008).

There are two common approaches to growth curve modeling: (1) the multilevel modeling framework and (2) the structural equation modeling (SEM) framework (Bollen & Curran, 2006). Multilevel models are also referred to as hierarchical models, mixed-effects models, and random-effects models. Pioneers in the field, Laird & Ware (1982), assumed that subjects from a single population would follow the same trajectory as the population means, and allowed for subject-specific deviations from the overall mean response pattern by adding random effects to the parameters. A milestone in growth curve modeling was reached by Meredith & Tisak (1990), who modeled growth essentially following the same premise as mixed-effects models, but formulated it within a general SEM framework. This approach falls into the same region as multilevel models, where latent curve approaches depict the overall mean trajectory for the sample, but each subject is assigned random effects to capture individual variation. The random effects, in contrast to multilevel models, are specified as latent variables in the context of SEM. It's worth noting that the latent curve models in the SEM context usually yield the same outputs in trajectory estimates as multilevel models do (McNeish & Matta, 2018).

2.1.3 Conventional group-based trajectory analysis

In the 1990s, researchers in the field were not satisfied with using a single continuous distribution function to represent the trajectory among the population. Instead, they assumed that multiple classes existed in the developmental profiles. Previously, both multilevel models and latent curve models could account for different growth patterns across different sub-groups (e.g., gender, race, and intervention/control) by treating the grouping variables as time-invariant

covariates in both frameworks. However, researchers could only identify the sub-population trajectory when grouping variables were measured. It remained a mystery in trajectory analysis, 'what if there was a potentially important grouping variable that had not been directly observed?' (Curran et al., 2010).

Nagin (1999) answered the call by proposing a semi-parametric, group-based statistical method analyzing longitudinal outcomes. This method is now referred to as Group-based Trajectory Analysis (GBTA), or, as Muthén (2001) referred to it, the Latent Class Growth Model (LCGA) in the SEM context. Conceptually, instead of assuming one trajectory of growth among populations, they suggested that the population is composed of a mixture of distinct groups (latent classes) defined by their developmental trajectories (Nagin, 1999). The latent classes are subgroups of individuals who are taxonomically similar in their response to the outcome variable (Nagin & Odgers, 2010a). Thus, subjects within different latent classes follow different developmental pathways, but subjects from the same latent class follow the same growth trajectory. The optimal number of latent classes is determined by the Bayesian information criterion (BIC; (Raftery, 1995)), the Lo-Mendell-Rubin likelihood ratio test (LMR-LRT; (Lo, 2001), and entropy (Nagin & Odgers, 2010b). The posterior probability of each individual's membership in the various classes that make up the model can be calculated, and individuals are assigned to the development trajectory classes based on the class with the highest posterior probability of latent class membership. The model is shown below.

$$P(Y_i|Time_i) = \sum_{j=1}^{J} \pi^j \, P\left(Y_j \middle| Time_i, j; \beta^j\right). \quad (1)$$

Conditional on time, the distribution of outcome trajectories denoted by $P(Y_i|Time_i)$, where Y_i represents subject i's longitudinal outcomes and $Time_i$ represents the time when Y_i is measured. The GBTA assumes that the population distribution of trajectories arises from a finite mixture of the total number of latent classes J, and the probability of membership in class j is indicated by π^j, which is a time-invariant quantity. The shape of the class-specific trajectory is determined by β^j. For an individual i from the given latent class j, conditional independence is assumed for the outcome y_{it}at each measurement t. Thus, we have

$$P\left(Y_i \middle| Time_i, j; \beta^j\right) = \prod_{t=1}^{T} p\left(y_{it} \middle| time_{it}, j; \beta^j\right), \quad (2)$$

where $p(.)$ is the distribution of y_{it} conditional on membership in class j and the time when individual i is measured at time t.

2.1.4 Missing data in group-based trajectory analysis

The conventional GBTA that has been reviewed can handle missing data as long as the missing mechanism is either MAR or MCAR. The MAR assumption posits that attrition is independent of any unobserved values, implying independence between attrition and latent class group membership. As the conventional GBTA treats group membership as time-invariant, the probability of a trajectory can be interpreted as the proportion of the population following a specific growth curve, unaffected by attrition or missing data.

However, the conventional GBTA cannot handle missing data that is not at random. Haviland et al. (2011) extended the conventional GBTA to accommodate nonrandom attrition. This extended model relaxed the assumption of independence between latent class membership and attrition, and introduced additional quantities relevant when attrition is present. The extension allows the attrition probability to follow a functional form of observed outcomes before dropout, as well as other covariates. This makes it possible to estimate latent-class-specific attrition rates and the trajectory group probability for the remaining population after dropout. This generalization of the conventional GBTA is well-suited to analyzing nonrandom attrition due to death or other events causing the outcome of interest to become undefined (i.e., truncation).

The extended GBTA model is described below. We suppose that participants are measured over a total of T occasions. An attrition indicator is introduced as ω_{it}. If $\omega_{it} = 1$, it indicates that individual i drops out by $t \leq T$ and 0 otherwise. The τ_i is the measurement time at which individual i drops out (i.e., $\omega_{\tau_i} = 1$). If an individual stays at the study until the end, the τ_i will be equal to $T + 1$. Thus, the $\omega_{it} = 0$ implies $\omega_{it} = 0$ in any prior measurements and the $\omega_{it} = 1$ implies $\omega_{it} = 1$ in any subsequent measurements. The θ_t^j is the probability of attrition for each measurement t given latent class j. For each individual at each measurement up to τ_i (the observations exist), the probability of the observed data given membership in group j is $p\left(y_{it} \middle| \omega_{it} = 0, time_{it}, j; \beta^j\right)(1 - \theta_t^j)$. Multiplying across the period before τ_i, the probability

becomes $\prod_{t=1}^{\tau_i-1} p\left(y_{it}\middle|\omega_{it}=0, time_{it}, j; \beta^j\right)(1-\theta_t^j)$. Finally, by including the attrition that happens at the measurement τ_i at the probability equaling to $\theta_{\tau i}^j$, equation (2) has a more general form.

$$P\left(Y_i\middle|Time_i, j; \beta^j, \theta_t^j\right) = \prod_{t=1}^{\tau_i-1} p\left(y_{it}\middle|\omega_{it}=0, time_{it}, j; \beta^j\right)\left(1-\theta_t^j\right)\theta_{\tau_i}^j. \quad (3)$$

The above-described model extension allows the probability of attrition θ_t^j to vary by trajectory and across time. Let H_t denote a vector composed of realizations of y_{it} before individual i drops out (i.e., $y_{i1}, y_{i2}, \ldots, y_{i\tau_i-1}$). Let x_i indicate the i^{th}individual-specific covariates. Then θ_{it}^j can be expressed as $\theta_t^j\left(x_i, H_t, t; \gamma^j\right)$, where the γ^j is the vector of group-specific parameters of the dropout process (i.e., the intercept and the linear slope). The functional form of the θ_t^j is assumed to be the binary logit function.

$$\theta_t^j\left(x_i, H_t, t; \gamma^j\right) = \frac{1}{1+e^{f^j\left(x_i, H_t, t; \gamma^j\right)}}, \quad (4)$$

Where $f^j\left(x_i, H_t, t; \gamma^j\right)$ is specified to be linear in its parameter γ^j.

Moreover, the interpretation of the probability of latent class membership π^j in the extended model is different from the conventional GBTA model. The conventional GBTA assumes that the group membership probability is immune to time. However, in the extended model, with attrition existing in each follow-up occasion, the estimated probability π^j should be properly interpreted as the latent group proportion at the initial stage that the attrition has not occurred, and it will change over time. With the help of additional estimates of dropout probability θ^j (assuming constant dropout rates across time within groups for simplicity), a time-varying group membership probability for the nonattracting members could be estimated using the equation below.

$$\tilde{\pi}_t^j = \frac{\pi^j\left(1-\theta^j\right)^{t-1}}{\sum_j \pi^j(1-\theta^j)^{t-1}}, \quad (5)$$

where the initial group proportion π^j is weighted by group j's probability of survival through period t, $\left(1-\theta^j\right)^{t-1}$. Thus, if attrition is due to truncation (e.g., death), then $\tilde{\pi}_t^j$ should be interpreted as the expected proportion of participants in group j who are alive at time t.

2.2 Trajectory of frailty index and associated factors

2.2.1 Frailty and frailty index

Frailty is a state of increased vulnerability to poor resolution of homeostasis following a stressor event, which increases the risk of adverse outcomes, including falls, incident disability, hospitalization, and mortality (Xue, 2011). These stressor events can be seemingly minor insults, such as a new drug, minor infection, or minor surgery Unlike fit elderly individuals who, after a minor stressor event, experience a small deterioration in function and then return to their previous level of homeostasis, frail elderly individuals may experience a larger deterioration and never return to their baseline homeostasis. This condition has been described as a transition phase between successful ageing and disability (Clegg et al., 2013).

Frailty does not have a gold standard definition. The general premise is that frailty can be a geriatric syndrome reflecting multisystem dysfunction (Dent et al., 2016), and it is assumed to be caused by the interaction of the progressive age-related decline in physiologic systems with chronic diseases and conditions, consequently leading to decreased functional reserve capacities (Cesari et al., 2016). These physiological systems that are inherently correlated include the brain, endocrine, immune, skeletal muscle, cardiovascular, respiratory, and renal (Clegg et al., 2013). Among these, the brain is one of the most studied organ systems in the development of frailty (Walston et al., 2006).

Although consensus on defining frailty has not been reached in academia, there is a growing body of operational models for measuring frailty from the perspective of clinical practice. The first breakthrough was witnessed in the mid-1990s when Sager et al. (1996) demonstrated that when frailty presentations (e.g., slow walking speed and weight loss) were combined to generate a summary score, their ability to predict adverse health events was better than when the components were considered alone. This inspired researchers in the field to generate a comprehensive, multi-perspective score to measure frailty. The frailty phenotype model and frailty index model stood out as the most commonly used ones in this field. There are also a growing number of instruments, including the Clinical Frailty Scale (CFS) (Kenneth Rockwood

et al., 2005) and Multidimensional Prognostic Instrument (MPI) (Pilotto et al., 2008), but they are less commonly used compared with the frailty phenotype and frailty index models.

Frailty Phenotype Model

Fried et al. (2001) viewed frailty as a pre-disability syndrome and proposed the landmark frailty phenotype model, which uses five physical components to assess frailty. These components are unintentional weight loss, low energy, slow gait speed, reduced handgrip strength, and reduced physical activity. Frailty is considered when three out of five of these components exist in a participant. Pre-frailty is considered when one or two of the components are present. Otherwise, the participant is considered to be in the robust stage.

Frailty Index Model

Another group of researchers in Canada developed the frailty index model in the same year (Arnold et al., 2001). This measurement model was based on a belief that the more deficits that individuals have, the higher the likelihood that they will be frail. The deficits can include symptoms, signs, disabilities, diseases, and laboratory test results. And a frailty index could be generated by calculating the ratio of deficits present to the total number of deficits considered. For example, if a total of 45 deficits are considered and a subject exhibits nine of these deficits, the frailty index for the subject is nine out of 45, which is .2. However, not all the deficits available can be selected into the frailty index model, for which process we should follow five criteria (Searle et al., 2008):

1. Selected deficits must be those related to health status. Variables such as graying hair, while they are age-related, should not be included.
2. Selected deficits should have an increasing prevalence rate with ageing (although some age-related adverse conditions can decrease in prevalence at senior ages due to survival effects).
3. Selected deficits should not be saturated too early. Attributes that are almost universal by age 55 (e.g., presbyopia) saturate too early to be considered a deficit.
4. Selected deficits should cover multiple systems. If all variables selected are measuring or related to cognition, then the index is a cognition index instead of a frailty index.
5. If the frailty index is used serially on the same group of people, the items across the population over the studying period should be the same.

After the selection of items for the deficits, binary indicators are created, with '0' representing the absence of the deficit and '1' representing its presence. Ordinal variables are accommodated by grading the rank into a score between 0 (where there is no presence of a deficit) and 1 (where the maximumly expressed presence of a deficit). For continuous variables, a recognized cutoff point can be used. Quantiles or percentiles among the population could be a substitute if there are no recognized cutoff values from a clinical perspective.

When constructing the frailty index, it was suggested that at least 30-40 items need to be included in the measurement (Rockwood & Mitnitski, 2007). Although including more items in the frailty index model may yield more precise estimates, and there is no scientific reason not to bring more, it has been confirmed that the model remains stable with around 40 deficits added Additional items could be selected randomly without affecting its performance (Ferrucci et al., 2004; Mitnitski et al., 2005). Multiple studies have reported an age-invariant submaximal for the frailty index (Kulminski et al., 2007; Mitnitski et al., 2005; Kenneth Rockwood & Mitnitski, 2007), and it has been accepted as a standard that an individual's frailty index should not exceed 0.67. Although the frailty index is a continuous variable, researchers have categorized it into a three-level ordinal variable (robustness, pre-frailty, and frailty), similar to the practice in the frailty phenotype model, but with different cutoff values across various papers (Romero-Ortuno, 2013).

Comparative Evaluation for the Two Measurements

Although the frailty phenotype and the frailty index are tools designed to measure frailty for operational purposes in clinical settings, and while they dominate the field, their uniqueness outweighs their overlap. Each of these models has strengths and weaknesses in different study settings. The main characteristics of the two measurements are shown in Table 1.

Table 1. Main characteristics of the frailty phenotype and the frailty index

Frailty Phenotype	Frailty Index
Frailty as a pre-disability syndrome	Frailty as an accumulation of deficits
Signs, symptoms	Disease, activities of daily living, clinical evaluation
Designed for prospective studies but could be modified to fit retrospective studies	Designed for retrospective studies
Ordinal variable	Continuous variable
Pre-defined criteria	Relatively vague criteria
Targeted at non-disable older persons	Targeted at the general population

In summary, as Cesari et al. (2014) mentioned in their paper, the two models should not be treated as alternatives or substitutes for each other but are complementary to each other's weaknesses. The frailty phenotype is suitable for individuals in the pre-disability stage in clinical practice and can be applied at the first contact with the subject. The frailty index tends to be popular in retrospective longitudinal studies that have comprehensive health assessments at hand, focusing on group-level results, which is the aim of this thesis.

2.2.2 Prevalence and consequence of frailty

The prevalence of frailty significantly varies across different countries and study settings. Collard et al. (2012) conducted a systematic review focused on community-dwelling adults aged 65 and older, aiming to compare and pool the prevalence of frailty. The studies incorporated into the review were conducted in developed countries, including the United States, Canada, and European countries. The reported prevalence of frailty within the community ranged from 4.0% to 59.1%, measured by either the frailty index or frailty phenotype. The overall weighted prevalence was 10.7% (95% CI = 10.5 - 10.9%). Choi et al. (2015) published a more recent review article focusing on community-dwellers aged 65 or older. They reported a smaller range of prevalence, from 4.9% to 27.3% in the six studies conducted in North America, Europe, and Japan, using only the frailty phenotype. However, when pre-frailty was included in the definition, a similar rate of frailty prevalence was observed (4.9 – 50.9%). In developing countries, frailty prevalence was reported to be higher than in developed nations, with considerable heterogeneity. Siriwardhana et al. (2018) conducted a systematic review focusing on the prevalence rate in China, Cuba, Brazil, and Tanzania. The prevalence of frailty varied from 3.9% (China) to 51.4% (Cuba), using the frailty phenotype model for community dwellers

aged 60 years and older. The pooled prevalence in developing countries was 17.4% (95% CI = 14.4% - 20.7%), higher than the reported prevalence in developed countries.

The prevalence of frailty among individuals in nursing homes differs from that in communities. Kojima (2015) conducted a systematic review of the prevalence of frailty in nursing homes in the United States. This review included studies using the frailty phenotype model to measure frailty. The studies, both cross-sectional and longitudinal designs with baseline data, included patients aged 60 years or older. The average prevalence of frailty ranged from 19.0 to 75.6%, and the pooled estimate was 52.3% (95% CI = 37.9 - 66.5%), both higher than their community-dwelling peers in the United States.

The prevalence of frailty increases with age. A study conducted by Lee et al. (2016) targeted older adults aged 90 and older. It was a longitudinal study including 824 participants living in California, United States, at the baseline. The overall prevalence rate at the beginning of the study was 24% in those aged 90 to 94 and 39.5% in those aged 95 and older. Pérez-Zepeda et al. (2020) used the Canadian Longitudinal Study on Aging (CLSA) baseline data to focus on a relatively younger population, reporting a lower prevalence rate for those aged 45 and older.

Frailty was found to be more prevalent in women than in men. It was verified by studies conducted in Britain, Canada, Korea, and China (Baek & Min, 2022; Gale et al., 2015; Pérez-Zepeda et al., 2020; K Rockwood & Mitnitski, 2012). A systematic review using an extended frailty phenotype model showed that the prevalence rate for frailty among community-dwelling women was 9.6% (95% CI = 9.2 – 10.0%), compared to 5.2% in men from the same study cohort (95% CI = 4.9 – 5.5%; Collard et al., 2012).

There is a growing body of evidence suggesting that frailty predicts adverse health outcomes, including disability, hospitalization, long-term care admission, and death. Vermeulen et al. (2011) demonstrated that frailty is a predictor of future disability in Activities of Daily Living (ADL). A Canadian prospective cohort study using the frailty index reported that frailty was significantly associated with new disabilities post-surgery (McIsaac et al., 2018). Kojima (2016) further showed that both frailty and pre-frailty measured by the phenotype model were significantly associated with a higher hospitalization risk. Irrespective of the measurement models used, frailty is associated with a higher mortality rate. McIsaac et al. (2018) found that frailty, measured by the frailty index model and clinical frailty scale, was associated with a higher risk

of death. Another review demonstrated that individuals deemed frail, measured by either the frailty phenotype model or the frailty index model, have a higher risk of death than robust elderly persons (Pereira et al., 2017).

2.2.3 Factors associated with frailty

Risk factors for frailty in older adults have been established from sociodemographic, physical, biological, and clinical perspectives. The frailty index model and multisystem frailty risk index both use physical, clinical and biological factors as components or deficits to measure frailty. Therefore, in this thesis, only sociodemographic factors (e.g., neighborhood, education, and access to health care) are considered as risk factors as they are not used in the outcomes measured by either the frailty index or frailty phenotype model.

Education is a recognized risk factor for frailty, as indicated in multiple analyses. For instance, a study conducted in the U.S. used the frailty phenotype to measure frailty and reported that those identified as frail at the baseline have lower educational attainment (i.e., without a college degree) (Woods et al., 2005). After a 3-year follow-up, education emerged as a risk factor for incident frailty, even after controlling for age, income, and diseases. However, this study focused only on women. Another study conducted in Greece included 1,867 participants aged 65 years and above, and it used five different instruments to measure frailty, including the frailty phenotype and frailty index models (Ntanasi et al., 2020). Among the sociodemographic factors, only education was consistently associated with frailty, irrespective of the measurement models for frailty. Additionally, Brigola et al. (2019) conducted a secondary data analysis on a large study in Brazil. They used Fried's phenotype model to measure frailty and reported that having no formal education doubles the likelihood of being categorized as frail (95% CI = 1.0 – 3.9). They also found a similar magnitude of correlation between having 12-24 months of education and frailty (OR = 2.0, 95% CI = 1.0 – 4.1). A longitudinal study on aging in Amsterdam, the Netherlands, showed that a low educational level (i.e., general intermediate education) corresponds to a 2.94 times higher risk of becoming frail compared to individuals with a secondary or higher school completion (Hoogendijk et al., 2014) (95% CI = 1.84 – 4.71). These differences persisted throughout 13 years of follow-up, underscoring the role of education as a risk factor for frailty. The results call for interventions aimed at reducing frailty, particularly among less educated groups.

Income or financial status is another well-studied risk factor for frailty. Individuals with low income (i.e., below the national average) were found to have more than twice the risk of becoming frail compared with those with high income (i.e., above the national average) (OR = 2.29, 95% CI = 1.41 – 3.73) (Myers et al., 2014). However, this study only targeted the post-myocardial infarction population. Similar findings were reported by Woods et al. (2005), but the study design was flawed in that it included only women. Meanwhile, some studies dispute the existence of an association between income and frailty in older adults. Aranda et al. (2011), who used the ability to pay bills as a measure of financial strain, found that the association between financial stress and frailty was not significant (OR = 0.99, 95% CI = 0.7 – 1.27). Given the contradictions in the literature, the contribution of income to frailty growth warrants further investigation.

Other factors of concern include living arrangement (i.e., whether the person is living alone or not) and marital status. Theoretically, they both serve as components of the variables used to define social frailty (Bessa et al., 2018), and may predict frailty that is defined from physical and clinical perspectives.

The meta-analysis conducted by Kojima, Taniguchi, et al. (2020) suggested that living alone might increase the risk of frailty. Among the 46 cohorts selected in cross-sectional studies, only men living alone were at an increased risk, while women were not. When assessing the living arrangement as a time-varying covariate in longitudinal studies, the association between status change in living arrangement and frailty was not observed. However, the number of longitudinal studies was relatively small in this systematic review compared to cross-sectional ones (6 versus 44), and the selected longitudinal studies also had their limitations (e.g., targeting only women (Woods et al., 2005)). As for marital status, limited evidence was found. Although a systematic review including 35 cross-sectional studies suggested that unmarried individuals were more likely to be frail in later life than married couples (Kojima, Walters, et al., 2020), only three studies that treated marital status as a time-varying covariate and analyzed its impact on the development of frailty growth using longitudinal study designs were included in the review, and a high degree of heterogeneity was witnessed. This could be explained by different reasons for being unmarried (i.e., widowed, divorced, or separated) and the cultural context in which the

study was conducted. Therefore, further analysis of the impact of marital status and living arrangements on frailty growth over time in a longitudinal setting is required.

2.2.4 Trajectories of the frailty index

There are significant numbers of scholarly papers using trajectory analysis to study the frailty growth over time in the elderly population. This literature review examines the topics from the standpoint of frailty measurement, statistical models for depicting the trajectory, and strategies for handling missing data.

A systematic review was carried out by Welstead et al. (2020) that encompassed 25 longitudinal observational studies, quantifying frailty trajectory. The frailty index was the prevailing measurement model for frailty in trajectory analysis, as evidenced by its use in 19 studies to derive the trajectory outcomes. The remaining studies employed frailty phenotype measurement. Nevertheless, within the range of publications reviewed, there were variations in the statistical methods utilized. These included mixed-effects models/latent curve analysis (10 publications), mixture models (2 publications), and other methods (13 publications).

Stow et al. (2018) were the pioneering researchers to apply GBTA to identify differing longitudinal patterns of frailty change over time and the consequent effects on mortality. Electronic health records were extracted for 13,149 seniors aged 75 or older who passed away within a year. Frailty was measured by the frailty index model, and a monthly frailty index was calculated for one year prior to death for these individuals. The use of GBTA led to the identification of three latent classes with linear growth: (1) rapidly rising frailty (2.2%; baseline: 0.21, the monthly rate of change: 0.022), (2) moderately increasing frailty (21.2%; baseline: 0.26, the monthly rate of change: 0.007), and (3) stable frailty (76.6%; baseline: 0.26, the monthly rate of change: 0.001). In another study, Hwang et al. (2021) sought to investigate changes in frailty status as defined by frailty phenotype and frailty index concurrently, using GBTA to present their findings. The study was based on the Taiwan Longitudinal Study of Ageing, which included 2,807 subjects aged 50 or older at the inception of the study. When frailty was gauged by the frailty index model, three latent classes with linear growth were identified: (1) stable (69.1%; baseline: 0.10, the yearly rate of change: 0.0025), (2) moderate increase (21.5%; baseline: 0.20, the yearly rate of change: 0.0014), and (3) rapid increase (9.3%; baseline: 0.28, the yearly rate of change: 0.0039).

The largest research gap lies in the missing data strategies in analyzing developmental pathways of frailty. When conducting frailty trajectory analyses using mixture models, the majority of publications either failed to report a method for dealing with missing data (Liu et al., 2018; Peek et al., 2012; Stow et al., 2018) or employed a complete-cases analysis excluding participants who had a certain level of missing throughout the follow-up period (Hwang et al., 2021). Other papers addressed missing data through two-stage MI or assuming the missing mechanism to be MAR. As mentioned above, missing data could be handled well using either the FIML estimation or the two-stage MI if and only if the missing mechanism is MCAR or MAR. However, this is not always the case in frailty analysis. A missing subject or a missing survey is likely due to the high level of frailty that should have been observed (e.g., the subject is too frail to answer the survey to report his frailty level). Thus, the missingness or attrition is dependent on an unobserved variable (i.e., the missing frailty level of the subject), and the resulting missing mechanism is MNAR. It violates the assumption of FIML and two-stage MI. If researchers overlook the assumption or ignore the missing data, the estimators will be biased, and the frailty level will be underestimated (i.e., assuming the missing is at random while the value should be high). Moreover, if the assumption of missingness is violated when using GBTA, researchers are at risk of underestimating the number of latent classes in the population. Assuming the non-ignorable attrition MAR, GBTA will fail to find the high-level frailty sub-group(s) consisting of seniors with a high risk of losing follow-ups. Thus, this thesis aims to examine the impact of non-ignorable missing data on frailty trajectory, particular through the use of an extended group-based trajectory analysis to find latent classes among the population.

2.3 Manitoba Follow-up Study

The Manitoba Follow-up Study (MFUS) represents the most enduring longitudinal study in Canada focusing on cardiovascular disease and ageing (Tate et al., 2015). Launched on July 1, 1948, the study originally consisted of 3,983 male participants recruited from the Royal Canadian Air Force at the conclusion of World War II. At the start of the study, their mean age was 31 years, with 90% falling within the age range of 20 to 39 years. By the spring of 2019, 107 original cohort members remained alive.

Initially, the primary research goal was to examine whether abnormalities in electrocardiograms could predict future cardiovascular disease. As such, the members of the cohort were free of

ischemic heart disease, and the MFUS committed to routine medical examinations at regular intervals. In 1996, the research focus expanded to assess the roles of physical, mental, and social functioning in the process of ageing. Subsequently, the Successful Ageing Questionnaire (SAQ) was disseminated in 1996, 2000, 2002, and annually thereafter. The SAQ includes Activities of Daily Living (ADL), Instrumental Activities of Daily Living (IADL), and a 36-item short-form health survey (SF-36; Ware & Sherbourne, 1992). The existing survey data, along with illnesses and disease diagnoses in medical history record, enabled the potential application of the Frailty Index model to measure frailty.

Two scholarly articles have been published that employed the MFUS data for trajectory analyses. Lengyel et al. (2017) leveraged the Group-Based Trajectory Analysis (GBTA) to assess nutritional risk. Over a four-year period, they identified five subgroups of nutritional risk SCREEN II scores among older men in the MFUS cohort. Those with poor nutritional risk trajectories demonstrated a higher risk of mortality over the subsequent two and a half years.

A different trajectory analysis using MFUS data sought to characterize participants' quality of life over a decade in both mental and physical health domains, and to discern if these trajectories predicted mortality in the ensuing nine years (John et al., 2018). A dual trajectory method was used to concurrently evaluate mental and physical health trajectories and their correlation. Three trajectory classes were discerned for both mental and physical function scores, with 37% of the participants maintaining a high level in both health domains. The study found that deteriorating mental and physical health statuses were linked to lower survival rates in subsequent years.

Two scholarly papers focusing on frailty that utilized the MFUS data were identified. Both shifted the perspective on how frailty was defined, advocating for its definition from the standpoint of seniors, rather than as a tool generated by researchers (John et al., 2019; Sachs et al., 2021). Despite inconsistent definitions and diverse experiences of frailty among participants (John et al., 2019), Sachs et al. (2021) found that self-rated frailty correlates with other health measures and predicts mortality over a three-year period (HR = 3.3, 95% CI = 1.5 - 7.1).

2.4 Objectives

Most statistical techniques utilized in longitudinal data analysis, including the conventional GBTA, adeptly handle missingness or attrition provided the mechanism is either MCAR or MAC. However, when the mechanism becomes MNAR, the conventional GBTA yields biased

estimates of trajectory classes and misleading results concerning factors correlated with trajectory classes. The extended GBTA, as discussed above, shows potential for trajectory analysis under conditions of MNAR attrition or missingness. Nevertheless, the circumstances under which the extended GBTA outperforms the conventional GBTA remain uncertain. Existing literature on frailty trajectory analysis assumes missingness as either MCAR or MAC, a presumption potentially inappropriate when participants are too frail to continue the study. To our knowledge, the extended GBTA has not been employed to examine the frailty trajectory classes and their associated risk factors. Therefore, this thesis research is delineated by two objectives.

The first objective is to illuminate the benefits of the extended GBTA model, which accounts for nonrandom participant attrition as opposed to the conventional model, with respect to trajectory group proportions, shapes, and sizes, using a simulation study in instances of non-random missing data. The hypothesis is that the GBTA modeling extension yields more precise latent classes when these classes are not well-separated in the presence of non-ignorable missing data.

The second objective is to explore the trajectories of frailty among older men and the factors correlated with frailty trajectories using data from the MFUS by contrasting the results derived from the conventional GBTA model and its extended variant. The MFUS data, offering more than 20 years of observations on medical, functional, and psychosocial age-related information, is advantageous in developing a reliable frailty index and scrutinizing frailty trajectory classes while accounting for attrition due to death. The hypothesis is that the extended GBTA model facilitates the identification of four latent classes of frailty index trajectories with class-specific attrition rates. This approach is expected to be more accurate and informative than the conventional GBTA model, a conclusion supported by simulation results. Additionally, it is hypothesized that factors such as age and living arrangement influence latent class membership for the participants.

In this thesis project, we illustrate the importance of considering the non-ignorable missing when modeling developmental trajectories for older populations. The extended GBTA can reveal not only the distinctive frailty trajectory patterns, but also the mortality rates for the latent classes, which may be overlooked by conventional methods. Accurately recovering the frailty trajectories

can be beneficial in primary care settings, as it can indicate an increased risk of mortality and the potential need for palliative care.

3. Simulation Study

Simulation is an effective way to compare model performance, such as the precision of parameter estimation under various conditions with known model settings. The simulation for the thesis is designed to compare the performance of the extended GBTA (accounting for nonrandom attrition) and the conventional GBTA (treating the attrition as random) in estimating the shapes of latent trajectories and the proportion of each trajectory class.

3.1 Simulation Design and Population Model

To simulate longitudinal data with latent classes, we used the following population model with a linear growth including J latent classes and N subjects with T repeated measures.

$$P(Y_i|t) = \sum_{j=1}^{J} \pi^{(j)} P\left(Y_j \middle| t, j; \beta^{(j)}\right)$$

$$\sum_{j=1}^{J} \pi^{(j)} = 1$$

$$P\left(Y_i^{(j)} \middle| t, j; \boldsymbol{\beta}^{(j)}\right) = \prod_{t=0}^{T-1} p\left(y_{it}^{(j)} \middle| t, j; \beta^{(j)}\right)$$

$$y_{it}^{(j)} = \beta_{00}^{(j)} + \beta_{01}^{(j)} t + \epsilon_{it};\ \pi^{(j)}$$

$$\epsilon_{it} \backsim N(0, \sigma^2)$$

$$i = 1, 2, \ldots, N;\ t = 0, 1, \ldots, T$$

$y_{it}^{(j)}$ is the simulated outcome for subject i in the jth class at time t; $\beta^{(j)}$ is a vector of latent class-specific growth parameters (i.e., the intercept and the linear slope in the model shown above); $\pi^{(j)}$ is the proportion of latent class j. The residuals ϵ_{it} from all latent classes are assumed to follow the same normal distribution with 0 mean and constant variance σ^2.

For simplicity, we further assumed that attrition only exists in the k^{th} latent class, named as attrition class. It is worth noting that acknowledging the computational complexities, an exhaustive examination of all possible scenarios where participant attrition might lead to potential biases in GBTA models was not performed. Instead, our simulation design was

specifically narrowed to situations where a minor subgroup suffered from attrition. The objective was to examine how this particular pattern of attrition could potentially impair the detection and estimation of the size of the latent class. An example of such a latent class is the frailest in the study of the developmental course of disability in an elderly population.

For subjects in k^{th} latent class, we used ω_{it} as the attrition indicator, with $\omega_{it} = 0$ indicating that the subject did not drop out from the study due to death before and at the measurement t. Let τ_i indicates the measurement when the attrition happens for individual i and, therefore, $\omega_{\tau_i} = 1$ and $\omega_{\tau_i} = 1$ for any subsequent measurement $t \geq \tau_i$. Thus, the observation for subject i in class k is a mixture of continuous scores following a normal distribution and a certain portion of missing responses shown below.

$$y_{it}^{(k)} \sim \begin{cases} attrition, & with\ probability\ of\ \theta_{it}^{(k)} \\ y_{it}^{(k)} | \omega_{it} = 0, & with\ probability\ of\ (1 - \theta_{it}^{(k)}) \end{cases}$$

$$y_{it}^{(k)} = \beta_{00}^{(k)} + \beta_{01}^{(k)} t + \epsilon_{it},$$

where $\theta_{it}^{(k)}$ is the probability of attrition for individual i in the k^{th} latent class at measurement t.

We considered three different attrition patterns in the latent class k.

1. MCAR

When attrition is MCAR in the latent class, it is independent of either observed or unobserved measurements, and the probability of attrition at the class level will follow a logit function below

$$logit\left(\theta_{it}^{(k)}\right) = \gamma_0$$

where γ_0 is a constant indicating that the proportion of the population lost to follow-up in the latent class k will remain a constant across different measurements.

2. MAR

When attrition is MAR in the latent class, the likelihood of attrition is dependent on the observed values and/or covariates but independent of the unobserved outcomes. The probability of attrition at the class level will follow a logit function below

$$logit\left(\theta_{it}^{(k)}\right) = \gamma_0 + \gamma_1 y_{it-1}^{(k)}$$

where γ_0 is the intercept and it could be used to change the attrition proportion in the population, and γ_1 is the coefficient measuring how $\theta_{it}^{(k)}$ relates to $y_{it-1}^{(k)}$, which is the previous observation for individual i from latent class k at measurement $t-1$. It indicates that the missing is only dependent on the observation before the dropout happens.

3. MNAR

When attrition is MNAR in the latent class, it is dependent on the would-be score of the measurements, and the probability of attrition at the class level will follow a logit function below

$$logit\left(\theta_{it}^{(k)}\right) = \gamma_0 + \gamma_1 y_{it}^{(k)}$$

where γ_0 is the intercept and we could use it to change the attrition proportion in the population, and γ_1 is the coefficient measuring how $\theta_t^{(k)}$ relates to $y_{it}^{(k)}$ (i.e., the would-be observation for individual i from latent class k at measurement t). In contrast to the constant missing rate across measurements, when attrition is MNAR, whether attrition exists or not is determined by the potential values of the outcome if it were not missed.

3.2 Data Generation and Simulation Conditions

3.2.1 Data Generation Steps

Data are generated by two steps: (1) generating a full observations data set using GBTA and (2) replacing a certain portion of observations with missing values determined by attrition patterns. We assume that there are two latent classes in the population and attrition only exists in the first latent class. Attrition is due to truncation by assumption (i.e., once attrition happens, the subject would not come back to the study). There are six repeated measurements for each individual. The data generation process is shown in Table 2.

Table 2. Data generation for simulation study

Population Models	Variables and Parameters
Step 1: Generating score with no attrition	
Class 1 (with a proportion of $\pi^{(1)}$)	$i = 1, 2, \dots, N$
$y_{it}^{(1)} = \beta_{00}^{(1)} + \beta_{01}^{(1)} t + \epsilon_{it};$	$t = 0, 1, \dots, 5$
	$\epsilon_{it} \backsim N(0, \sigma^2)$
Class 2 (with a proportion of $\pi^{(2)}$)	
$y_{it}^{(2)} = \beta_{00}^{(2)} + \beta_{01}^{(2)} t + \epsilon_{it};$	
And	
$\pi^{(1)} + \pi^{(2)} = 1$	
Step 2: Generating attrition in latent class 1	
(1) MCAR $logit\left(\theta_t^{(1)}\right) = \gamma_0$	
(2) MAR $logit\left(\theta_t^{(1)}\right) = \gamma_0 + \gamma_1 y_{t-1}^{(1)}$	
(3) MNAR $logit\left(\theta_t^{(1)}\right) = \gamma_0 + \gamma_1 y_t^{(1)}$	

3.2.2 Simulation conditions

We considered two scenarios for the proposed model and the conventional GBTA model: (1) initially well-separated classes and (2) initially overlapped classes. The design of an initially well-separated class scenario was established by assigning disparate trajectory intercepts with substantial difference to the two latent classes (i.e., 10 vs. 5). The design of an initially not well-separated class scenario was established by assigning equivalent trajectory intercepts to the two latent classes (i.e., 10) but with different growth rates (i.e, 0.5 vs 0). Both scenarios were subjected to additional analysis to explore varying missing data mechanisms (i.e., MCAR, MAR, and MNAR). For each missing data mechanism under each scenario, this analysis was conducted at low attrition rate, characterized by $\gamma_0^{(1)}$ at .05 and high attrition rate, characterized by $\gamma_0^{(1)}$ at .2, targeting the latent class experiencing the attrition. Detailed conditions are shown in Table 3.

Table 3. Simulation Conditions

Scenarios	Conditions
The latent classes are well-separated	
class 1: $y_{it}^{(1)} = 10 + .5t + \epsilon_{it}$ class 2: $y_{it}^{(2)} = 5 + \epsilon_{it}$ $\epsilon_{it} \backsim N(0,1)$	(1) Attrition at MCAR proportion $\gamma_0^{(1)}$: .05/.2 (2) Attrition at MAR proportion $\gamma_0^{(1)}$: .05/.2 (3) Attrition at MNAR proportion $\gamma_0^{(1)}$: .05/.2
The latent classes are not well-separated	
class 1: $y_{it}^{(1)} = 10 + .5t + \epsilon_{it}$ class 2: $y_{it}^{(2)} = 10 + \epsilon_{it}$ $\epsilon_{it} \backsim N(0,1)$	(1) Attrition at MCAR proportion $\gamma_0^{(1)}$: .05/.2 (2) Attrition at MAR proportion $\gamma_0^{(1)}$: .05/.2 (3) Attrition at MNAR proportion $\gamma_0^{(1)}$: .05/.2

3.3 Model comparisons and performance measures

3.3.1 Models for comparison

Two models were used to analyze the generated data: (1) the conventional GBTA and (2) the extended GBTA accounting for nonrandom attrition. The conventional GBTA assumes that the attrition is MCAR or MAR and reported estimated shape and size parameters β_{00} and β_{01} (i.e., the intercept and the linear slope) for each latent class, and the probability of latent trajectory membership $\pi^{(1)}$ but it does not have an estimation of attrition-related quantities like dropout probability $\theta_t^{(1)}$. The output of applying the extended GBTA model accounting for nonrandom attrition to the data not only includes the growth parameters and the latent class proportions but has the dropout probability estimates as well. It was achieved by using the logit function of before-period y_{t-1} to predict the probability of attrition for the trajectory experiencing the attrition. Estimation of the model for both the extended and the conventional GBTA was performed employing the SAS Proc Traj (Jones 2007). The Proc Traj only allowed for the utilization of measurements before dropout (e.g., y_{t-1}) to predict the attrition process. The inherent limitations of this analysis, in conjunction with future directions for resolution, are presented in section 5.

Table 4. Models for comparison

Models	Parameters to Report
Conventional GBTA	
$y_{it}^{(1)} = \beta_{00}^{(1)} + \beta_{01}^{(1)} t + \epsilon_{it}; \pi^{(1)}$ $y_{it}^{(2)} = \beta_{00}^{(2)} + \epsilon_{it}; \pi^{(2)}$ $\epsilon_{it} \backsim N(0, \sigma^2)$	Group membership probability: $\widehat{\pi^{(1)}}$ Growth parameters: $\hat{\beta}s$
Extended GBTA	
$y_{it}^{(1)} = \beta_{00}^{(1)} + \beta_{01}^{(1)} t + \epsilon_{it}; \pi^{(1)}$ $y_{it}^{(2)} = \beta_{00}^{(2)} + \epsilon_{it}; \pi^{(2)}$ $\theta^{(1)}(y_{t-1}; \gamma) = \frac{1}{1 + e^{f^{(1)}(y_{t-1}; \gamma^{(1)})}}$ $f^{(1)}(y_{t-1}) = \gamma_0 + \gamma_1 y_{t-1}$ $\epsilon_{it} \backsim N(0, \sigma^2)$	Group membership probability: $\widehat{\pi^{(1)}}$ Growth parameters: $\hat{\beta}$s Dropout probability: $\widehat{\theta^{(1)}}$

3.3.2 Performance Measures

To compare the performance of the basic GBTA with the extended model on the simulated data, we will use the following criteria. The simulation times (n_{sim}) are assumed to be 200 for each condition under each theme. The measures and formulas are shown in Table 5.

Table 5. Simulation Measures and Estimates

Measures	Definition	Estimates
Absolute Bias	$\left\|E\left[\hat{\theta}\right] - \theta\right\|$	$\left\|\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \hat{\theta}_i - \theta\right\|$
Relative Bias	$\left\|E\left[\hat{\theta}\right] - \theta\right\| \Big/ \theta$	$\frac{\left\|\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} \hat{\theta}_i - \theta\right\|}{\theta}$
Mean Squared Error (MSE)	$E\left[(\hat{\theta} - \theta)^2\right]$	$\frac{1}{n_{sim}} \sum_{i=1}^{n_{sim}} (\hat{\theta}_i - \theta)^2$

1. Absolute bias: the absolute difference between the expected estimation of the parameters $\hat{\theta}$ with the real values θ. The expectation of the estimation for $\hat{\theta}$ could be replaced by the average value across the number of simulations.
2. Relative bias: the absolute bias divided by the real value θ. The expectation of the estimation for $\hat{\theta}$ could be replaced by the average value across the number of simulations.

3. Mean Squared Error (MSE): the expected value of the squared difference between the expectation of estimation and true values. It is estimated by averaging the squared difference between $\hat{\theta}$ and true value θ across n_{sim} repetitions.

3.4 Simulation analysis results

In this section, we present the results of applying both the conventional GBTA and the extended GBTA to 200 repeatedly generated data sets following the above-mentioned simulation design, each with varying random seeds.

3.4.1 Scenario I: Well-separated latent classes

Missing data mechanisms

1. MCAR

Figure 1 illustrates the relative bias in the estimation of latent class proportions from the conventional and extended GBTA methods when the attrition mechanism for the latent class of interest is assumed to be MCAR. The numerical results can be found in Table A2 in Appendix.

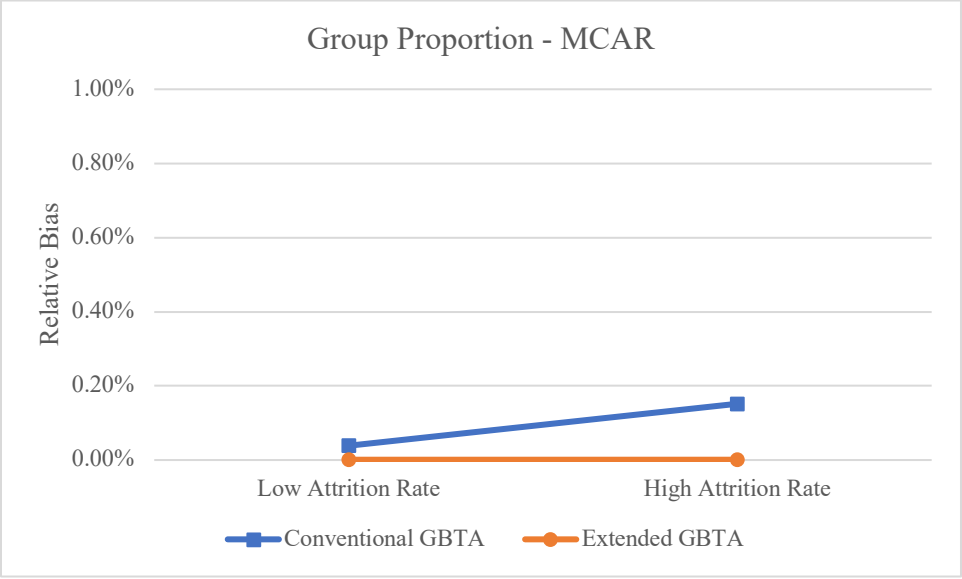

Figure 1. Relative biases of group proportion $\widehat{\pi^1}$ (π^1= 0.2) under the MCAR mechanism when groups are well separated.

The conventional and extended GBTA methods both produced estimates of the latent class proportions that were approximately unbiased, with relative biases below 0.2%. When the

attrition rate was low (5% of the population dropping out from the latent class of interest per measurement), the difference in the relative biases between the two methods was negligible. When the attrition rate was high (20% of the population dropping out from the latent class of interest per measurement), the extended GBTA had a much lower relative bias (0.04%) than the conventional GBTA (0.15%). The estimation accuracy from the extended GBTA was unaffected by the attrition rate, resulting in an unbiased estimation in both conditions.

For parameters in the linear growth model (intercept and slope), the conventional and extended GBTA methods yielded similar estimates and this similarity does not depend on attrition rates (detailed results in Table A2 in the Appendix).

Using the extended GBTA method, we can also estimate the likelihood of dropout for the class with attrition. The relative biases for this parameter are 0.26% when the attrition rate is low and 0.20% when the attrition rate was high.

2. MAR

Figure 2 depicts the comparative bias in the estimations of latent class proportions derived from both the conventional and extended GBTA methods when the dropout mechanism for the latent class is assumed to follow MAR. Detailed numerical results are provided in Table A3 in the Appendix.

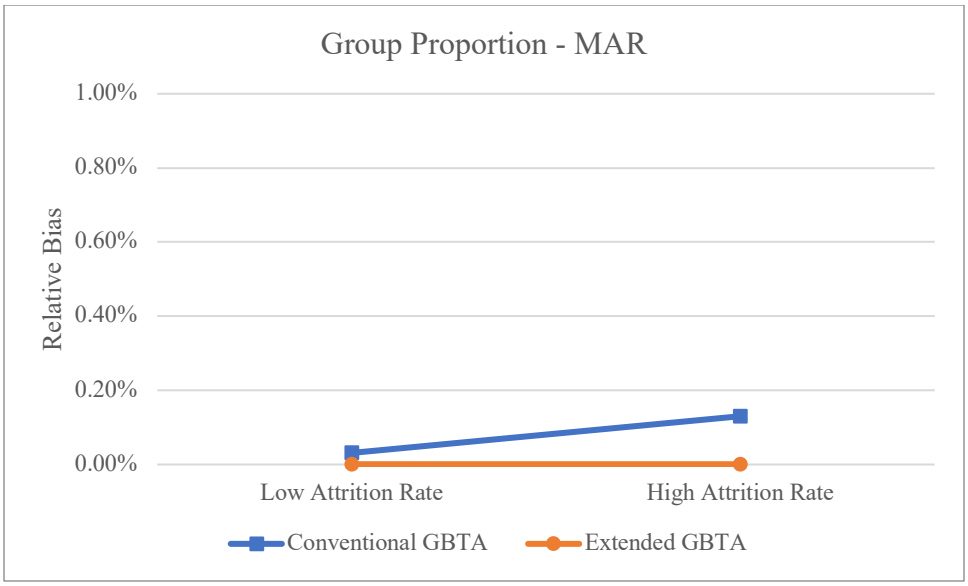

Figure 2. Relative biases of group proportion $\widehat{\pi^1}$ (π^1= 0.2) under the MAR mechanism when groups are well separated.

Both the conventional and the extended GBTA produced unbiased estimates for the latent class proportion, as demonstrated by all relative biases being less than 0.2%. The estimation accuracy from the extended GBTA proved robust against attrition rate variations, thus generating unbiased estimates across different attrition rates. However, a rise in relative bias from 0.03% to 0.13% was observed when the attrition rate increased from low (5% of the population withdrawing from the latent class of interest per measurement) to high (20% of the population withdrawing from the latent class of interest per measurement).

Both models offered the same estimates for the growth parameters, including intercepts and linear slopes consistently unbiased, despite different attrition rates.

The extended GBTA method also produced the likelihood of dropout for the class with attrition. In instances of the low attrition rate, the 5-year average relative bias of the dropout probability estimation was observed to be 3.31%. When it comes to the high attrition rate condition, the 5-year average relative was lower, registering at 2.68%.

3. MNAR

Figure 3 elucidates the relative bias of the latent class proportion estimation derived from both the conventional GBTA method and its extended version, under the assumption that the attrition

mechanism for the latent class of interest is MNAR. For numerical results, refer to Table A4 in the Appendix.

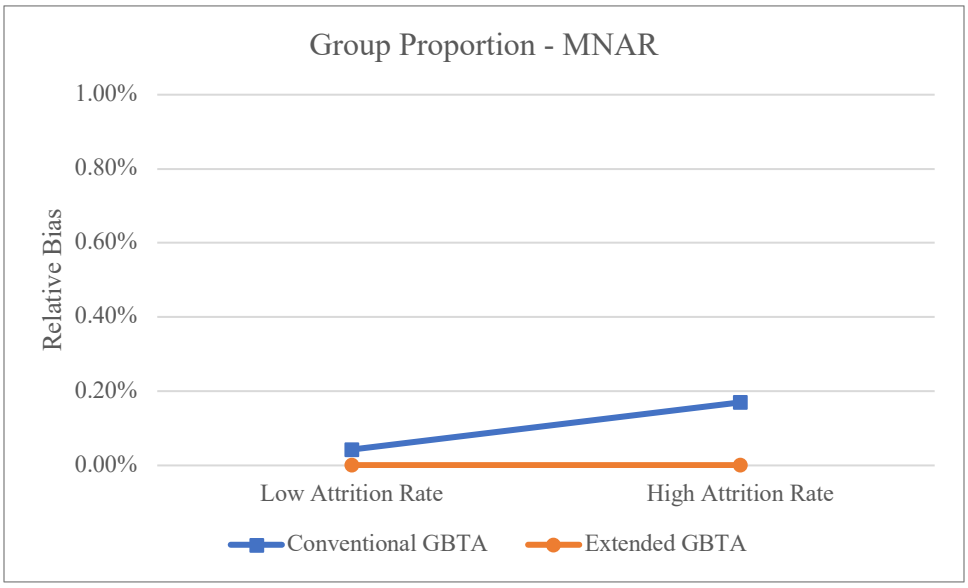

Figure 3. Relative biases of group proportion $\widehat{\pi^1}$ (π^1= 0.2) under the MNAR mechanism when groups are well separated.

It can be inferred that both the conventional and extended GBTA methods delivered near-unbiased estimates of the latent class proportion, corroborated by the relative biases all falling below a 0.2% threshold. The precision of estimations from the extended GBTA remains impervious to the attrition rate, consistently producing unbiased estimations. There was a slight surge in relative bias from 0.04% to 0.17% when the attrition rate escalated from a low (i.e., 5% of the population departing from the latent class of interest per measurement) to a high rate (i.e., 20% of the population departing from the latent class of interest per measurement) from the result of the conventional GBTA method.

The two models, under varying attrition rates, generated identical and nearly unbiased estimates for growth parameters including intercepts and linear slopes.

The extended GBTA method generated the estimation of the likelihood of dropout for the class with attrition. In situations characterized by a low attrition rate, the five-year average relative bias of the estimated dropout probability was documented at 3.60%. Similarly, under conditions

of high attrition, the relative bias was found to be slightly higher than that documented in the low attrition rate condition, recorded at 3.78%.

3.4.2 Scenario II: initially not well-separated classes

Missing data mechanisms

1. MCAR

Figure 4 provides a visualization of the relative bias in the estimation of latent class proportions from the conventional and the extended GBTA when the attrition mechanism for the latent class of interest is assumed to be MCAR. The two latent classes were posited to remain indistinguishable at the initial stage. Full results can be found in Table A5 in Appendix.

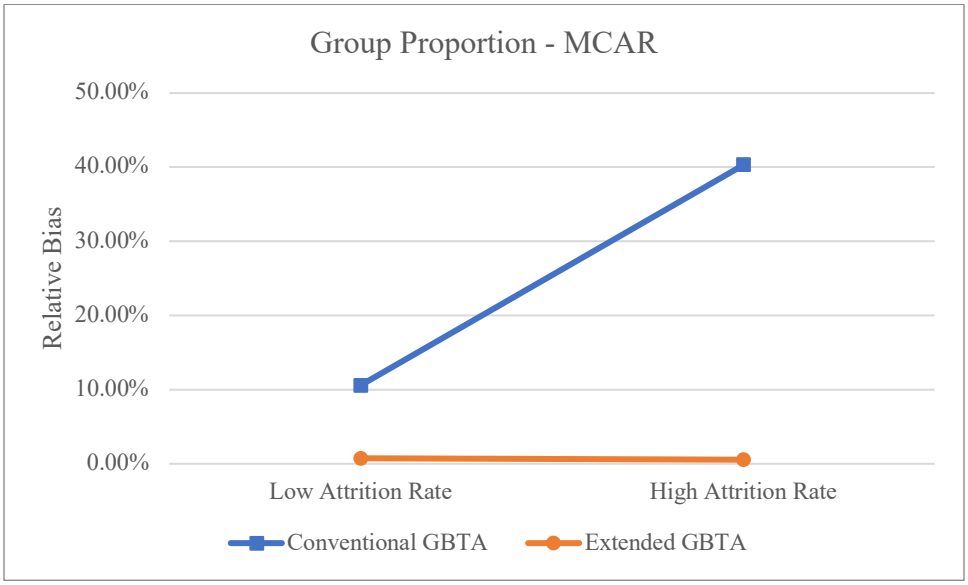

Figure 4. Relative biases of group proportion $\widehat{\pi^1}$ (π^1= 0.2) under the MCAR mechanism when groups are not well separated

A substantial relative bias was observed in the estimation of latent class proportions derived from the conventional GBTA. Specifically, a 10.59% relative bias was reported when the dropout rate was approximately 5% of the population in the latent class per measurement. This relative bias increased to 40.29% when the attrition rate was set to a higher level, with approximately 20% of participants in the latent class dropping out at each measurement. In contrast, the extended GBTA demonstrated superior performance in estimating latent class proportions compared to the

conventional model. Regardless of the attrition rate, the proposed model estimated the latent class proportion with a relative bias of less than 1%.

The two models yielded nearly identical estimates for the intercept consistently unbiased, and demonstrated resistance to changes in the attrition rate from 5% to 20%.

Figure 5 shows the relative bias of the linear slope estimation for the latent class with attrition when the missing data mechanism is MCAR.

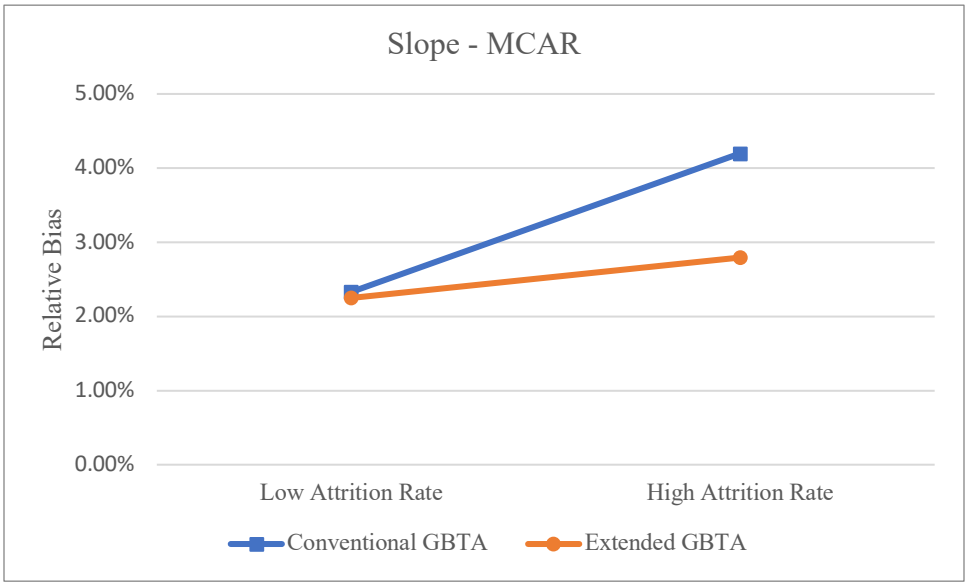

Figure 5. Relative biases of intercept $\widehat{\beta_1^1}$ (β_1^1= 0.5) under the MCAR mechanism when groups are not well separated.

For the linear slope estimation, the relative bias of the linear slope estimation was larger and more influenced by the increase in attrition rate compared with the intercept estimation. When the attrition rate was low, it reflected similar results from both models with the relative bias standing at 2.25%. However, when the attrition rate was increased, the extended model displayed an increased relative bias of 2.79%. The conventional model manifested a more pronounced increase, with the relative bias soaring to 4.20% under high attrition rate conditions.

The proposed extended GBTA's additional estimation of dropout probability for the latent class suffered from attrition exhibited minor bias, with a 0.57% relative bias at a low attrition rate and a 0.33% relative bias at a high attrition rate.

2. MAR

Figure 6 depicts the relative bias in the estimations of latent class proportions derived from both the conventional and the extended GBTA when the dropout mechanism for the relevant latent class is assumed to follow MAR. Detailed numerical results are provided in Table A6 in the Appendix.

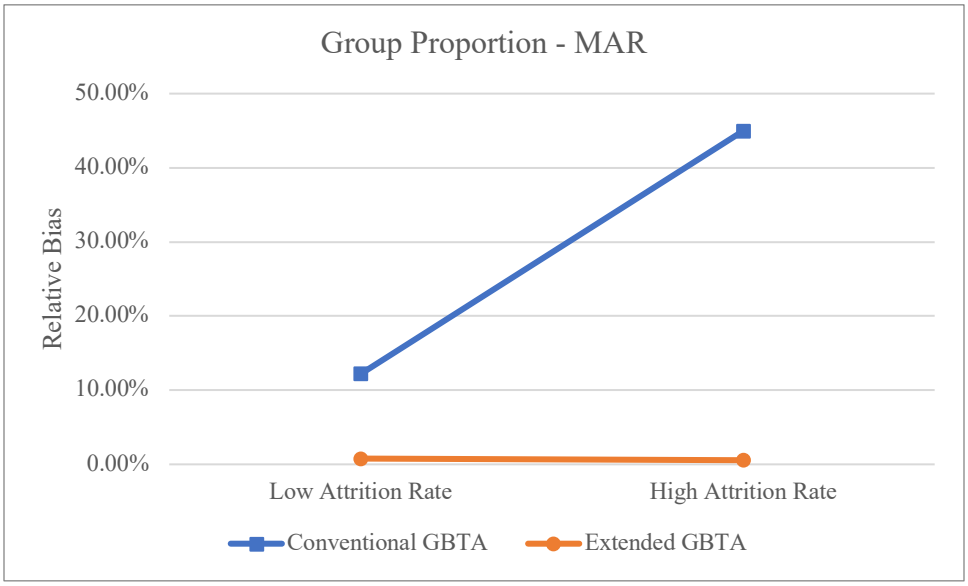

Figure 6. Relative biases of group proportion $\widehat{\pi^1}$ (π^1= 0.2) under the MAR mechanism when groups are not well separated.

A pronounced relative bias was observed in the estimation of latent group proportions utilizing the conventional GBTA. A 12.19% relative bias emerged when the dropout rate was around 5% of the latent class population per measurement. This relative bias amplified to 44.98% when the attrition rate was adjusted to a more substantial level, with approximately 20% of the participants in the latent class withdrawing at each measurement. Conversely, the proposed extended GBTA exhibited superiority in estimating latent class proportions compared to the conventional model. Irrespective of the attrition rate, the extended estimated the latent class proportion with a relative bias of less than 1%.

Figures 7 (a) and (b) present the estimation of growth parameters (i.e., the intercept and the linear slope) for the latent class experiencing attrition when the missing data mechanism is MAR.

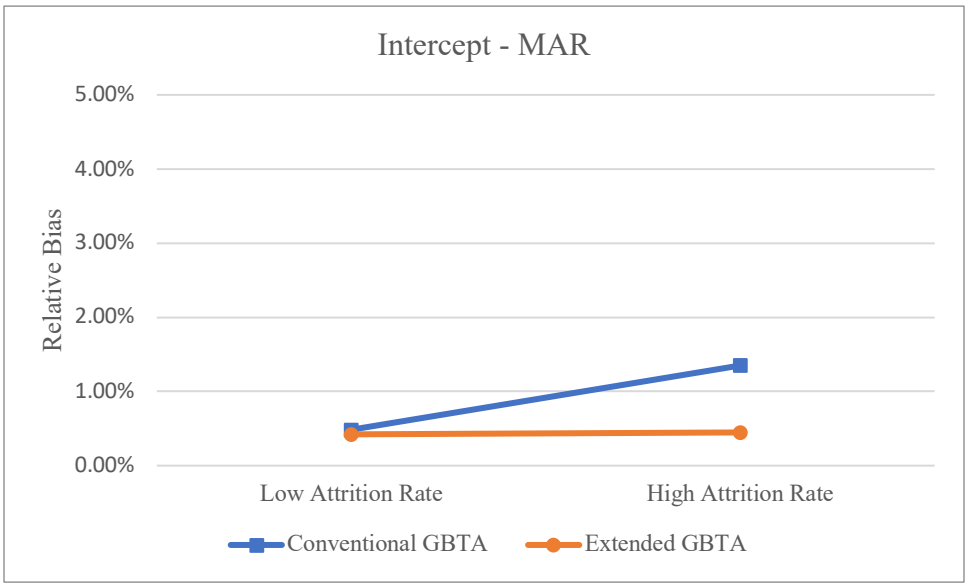

Figure 7 (a). Relative biases of intercept $\widehat{\beta_0^1}$ (β_0^1= 10) under the MAR mechanism when groups are not well separated.

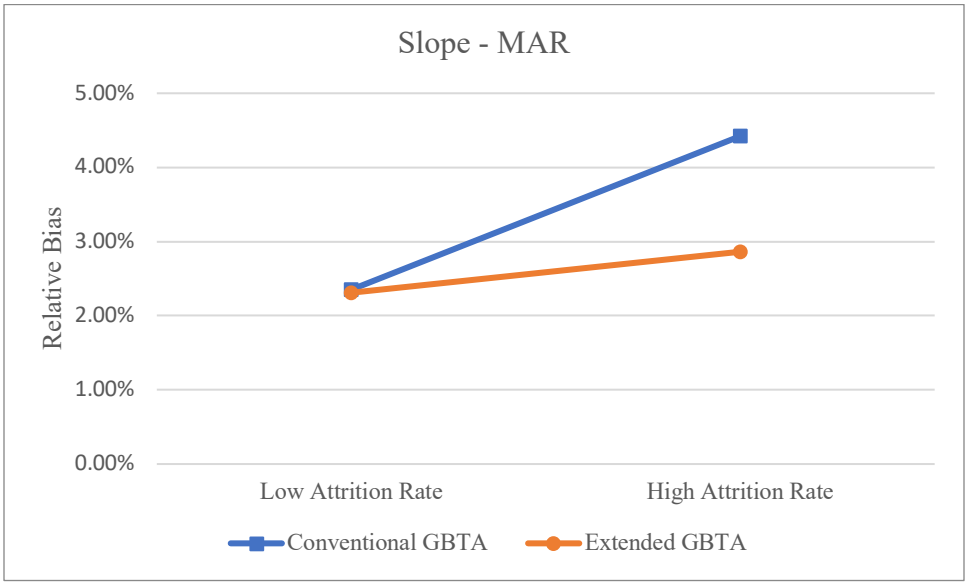

Figure 7 (b). Relative biases of intercept $\widehat{\beta_1^1}$ (β_1^1= 0.5) under the MAR mechanism when groups are not well separated.

Both models offered estimates for the growth parameters with relative bias below 5% across different attrition rates. Regarding the intercept estimation, the conventional model was more vulnerable to the attrition rate increase compared with the extended GBTA. The relative bias of the conventional model was 0.42% under the low attrition rate condition and increased to 1.35% under the high attrition rate condition. When it comes to the linear slope estimation, at low attrition rates, the relative bias from both models remained steady at 2.31%. Under high attrition rate conditions, the extended GBTA had a slight increase to 2.86%, whereas the conventional model witnessed an increase to 4.42%.

The extended model generated the estimates of the likelihood of dropout for the latent class. In instances of the low attrition rate, the five-year average relative bias of the dropout probability estimation was observed to be 3.01%. For the high attrition rate condition, the five-year average relative bias registered at 2.61%.

3. MNAR

Figure 8 elucidates the relative bias of the latent class proportion estimation derived from both the conventional GBTA and its extended form, under the assumption that the attrition mechanism for the latent class of interest is MNAR. For a comprehensive overview of the numerical results, refer to Table A7 in the Appendix.

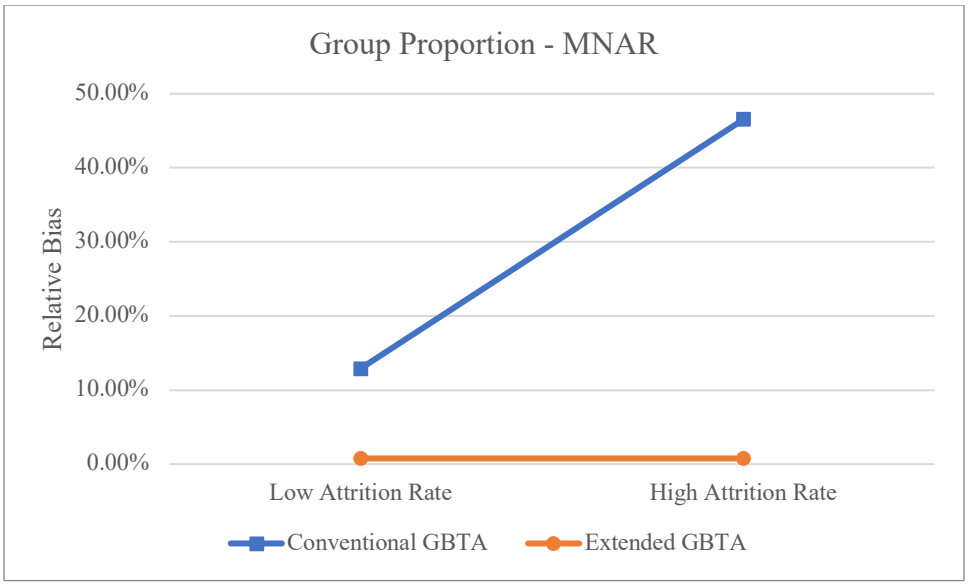

Figure 8. Relative biases of group proportion $\widehat{\pi^1}$ (π^1= 0.2) under MNAR mechanism when groups are not well separated

The conventional GBTA showed a significant relative bias in the estimation of latent group proportions. When the dropout rate accounted for approximately 5% of the latent class population per measurement, a pronounced relative bias of 12.91% was observed. This relative bias increased substantially to 46.52% when the attrition rate was adjusted to a higher level, where approximately 20% of the participants in the latent class withdrew at each measurement. The proposed extended GBTA demonstrated superior performance in estimating latent class proportions compared to the conventional model. Regardless of the attrition rate, the extended GBTA consistently estimated the latent class proportion with a relative bias of less than 1%.

Figure 9, sub-figures (a) and (b), depicted the estimations of growth parameters, namely the intercept and the linear slope, for the latent class experiencing attrition, with the missing data mechanism presumed to be MNAR.

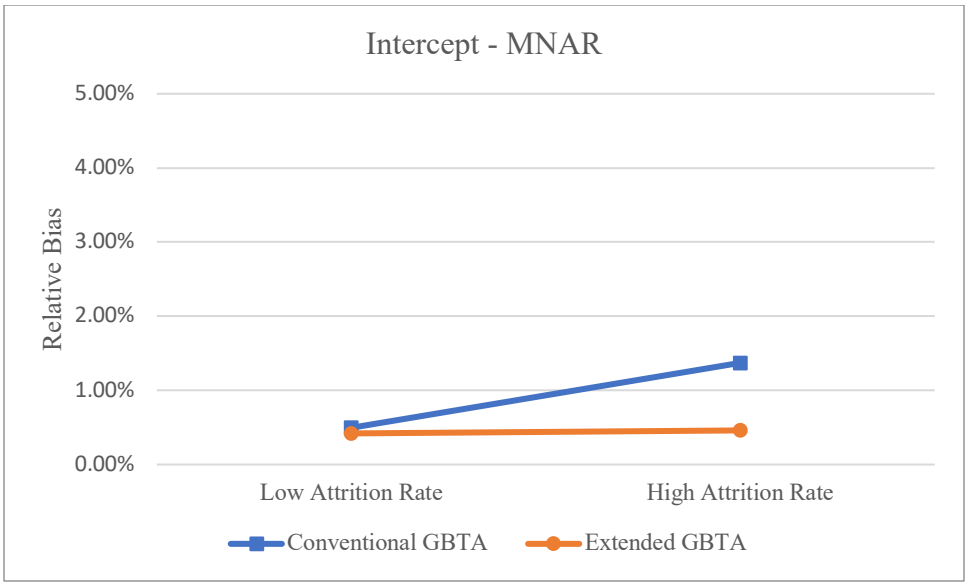

Figure 9 (a). Relative biases of intercept $\widehat{\beta_0^1}$ (β_0^1= 10) under MNAR mechanism when groups are not well separated

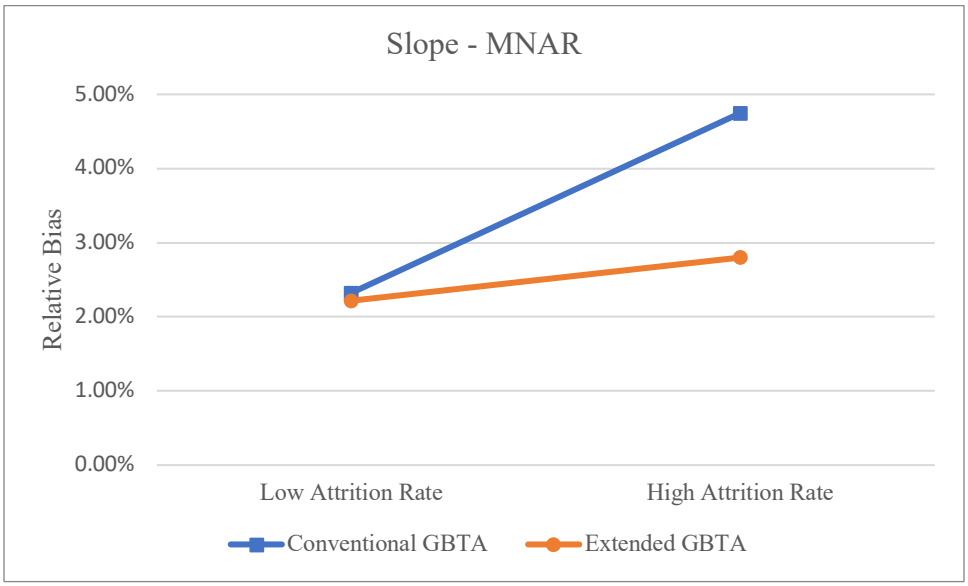

Figure 9 (b). Relative biases of intercept $\widehat{\beta_1^1}$ (β_1^1= 0.5) under MNAR mechanism when groups are not well separated

Both models employed in the analysis provided estimations for the growth parameters exhibiting relative bias values below 5% across various attrition rates. Under the low attrition rate condition, the relative bias for the intercept estimation from the conventional model was

determined to be 0.42%, which escalated to 1.35% under the high attrition rate condition. In terms of linear slope estimation, a larger relative bias was observed. For both models, when the attrition rates were low, the relative bias for slope estimations remained at around 2.30%. Under high attrition rate conditions, the extended GBTA model demonstrated a slight increase to 2.80% in relative bias, whereas the conventional model experienced a more pronounced increase to 4.75%.

Using the extended GBTA method, we can also estimate the likelihood of dropout for the class with attrition. In situations of a low attrition rate, the five-year average relative bias of the estimated dropout probability was 3.25%. Under conditions of high attrition, the five-year average relative bias was recorded at 3.71%.

3.5 Simulation summary

When the latent classes were well-separated, both the conventional and extended GBTA methods produced estimates of the class proportions and growth parameters that were approximately unbiased, with relative biases below 5%. These low biases were consistent across different missing data mechanisms and different attrition rates. The bias in the dropout probability estimation depends on the missing mechanisms, with larger relative bias for MNAR and smaller relative bias for MCAR than that for MAR.

When the latent classes are not well-separated, the extended GBTA method produced more accurate estimates than the conventional GBTA method for the class proportions and growth parameters. The advantage of the extended GBTA method over the conventional one was more pronounced when the attrition rate was high. The difference in the relative bias between the extended and conventional GBTA methods also depends on the missing mechanisms. The extended GBTA model consistently demonstrated high accuracy in class proportions, irrespective of missing data mechanisms. Regarding growth parameter estimates, both models correctly identified the average starting point (intercept) and rate of change (linear slope) across all missing data mechanisms. The relative bias of the likelihood of dropout estimation generated from the extended GBTA method remained with the 5% threshold regardless of the missing data mechanisms and the attrition rates.

In conclusion, the superiority of the extended GBTA methods became apparent when the initial classes were not well separated, especially in identifying the latent class proportions. The

conventional GBTA could do as effective as the extended model when the latent classes were well separated, and in estimating the growth parameters (i.e., the intercept and the linear slope).

4. Trajectory of frailty using data from the Manitoba Follow-up Study

In this section, we investigate distinctive frailty trajectories among older men and examine factors associated with the frailty trajectories using data from the Manitoba Follow-up Study (MFUS). The MFUS data set is described in section 4.1. In section 4.2, we describe how the frailty index is defined using health deficits items. In section 4.3, we report the results of frailty trajectory classes from both the conventional and the extended GBTA methods and discuss their implications. In section 4.4, we examine how living arrangements and marital status are associated with the frailty trajectory class membership.

4.1 Description of MFUS and sample used for trajectory analysis

The MFUS is the longest-running longitudinal study of cardiovascular disease and ageing in Canada, starting with a cohort of 3,983 men recruited from the Royal Canada Air Force in the year of 1948 (Tate et al., 2015). In 1996, the research focus of MFUS was expanded to explore the role of physical functioning in support of health and successful ageing. Since then, it has collected data on medical, functional, and psychosocial age-related information among survival MFUS members.

This thesis utilizes a subset of a ten-year longitudinal survey data collected survey from between 1996 to 2010 combined with the medical records of the participants. There were 10 waves of survey data collection over 14 years. In 1996, the baseline first survey on successful ageing questionnaire (SAQ) was first distributed. The core component of SAQ is a combination of Short Form-36 (SF-36), Assisted Daily Living (ADL), and Instrumental Assisted Daily Living (IADL). There were 2,055 questionnaires that were sent out in 1996, and 1,779 of them were returned. The SAQ was sent out again in 2000, 2002, and yearly after that.

As in most longitudinal studies, MFUS also suffers from missing data due to the multiple waves of data collection that increase the probability of non-response and participation attrition. In the period following 1996, the Manitoba Follow-up Study (MFUS) cohort, on average, has been characterized by an age exceeding 76 years. Given the progressive age of the cohort members, consistent participation in the longitudinal research becomes challenging. These challenges predominantly stem from factors such as, though not exclusive to, health deterioration, functional debilitation, and amplified risk of mortality. There are four non-monotone missing data patterns found in the MFUS. Firstly, there are missing items in the participant's responses to

the question each year. We assume those as MAR/MCAR and include those participants in the data analysis even if they missed some assessments. As part of the data pre-processing for the frailty index computation, the technique of single imputation was employed, which incorporated regression on the entirety of non-missing variables present within the questionnaires and medical records. This method was utilized to impute missing items from the questionnaires to facilitate the subsequent calculation of the frailty index. Secondly, there are missing responses to the whole questionnaire during the longitudinal observations. That is, those participants can be missing at one follow-up time and then measured again at one of the next. We operate under the assumption that these participants' data adhere to the MAR mechanism. Thirdly, there is a missing due to a high level of frailty when the death happened. As the cause of missing is a variable that could not be measured (i.e., potentially high frailty index score when the participant died), the missing data pattern is MNAR. Fourthly, participants are missing due to other reasons before their death and we assume the missing mechanism follows the MAR. In addressing the aforementioned three instances of missing data, a method of single imputation was employed, which was undertaken utilizing regression, predicated upon non-missing variables from both the questionnaire and medical records. The imputation resulted in the frailty index extending up to the year immediately preceding the participant's death. Since the research aim is to find the longitudinal trajectory of the frailty index, we only included those participants who had observations in 1996. Table 6 shows the summary of the participants of the MFUS each year since 1996.

Table 6. Summary of MFUS Members included in Frailty Trajectory Analysis (N = 1,839)

Year	Number of Participants	Number of Deaths
1996	1839	
2000	1495	
2002	1338	30
2004	1138	154
2005	1051	59
2006	956	79
2007	867	75
2008	770	84
2009	713	49
2010	629	71

Table 7 provides the baseline characteristics for the 1,839 participants in 1996. The average age was 76.8 years (SD = 3.8). The majority of participants (74.5%) received a high school education or more by 1996, and only 1.2% of the participants lived in personal care (Canada) or nursing homes (US). Over 73% of participants were married in 1996 and 20.4% were living alone in 1996.

Table 7. Baseline Characteristics of Participants (N = 1,839)

Variable Name	Number (%)
Age (Continuous)	
Mean ± SD	76.8±3.8
Median (IQR)	76(74-78)
Education	
Less than high school	320(17.4)
High school or more	1370(74.5)
Missing	201(10.9)
Living in Personal Care/Nursing Home	
Yes	23(1.2)
No	1701(92.5)
Missing	115(6.3)
Marital Status	
Single	22(1.2)
Married	1440(78.3)
Divorced/Separated	75(4.1)
Widowed	191(10.4)
Missing	111(6.0)
Living Alone	
Yes	375(20.4)
No	1337 (72.7)
Missing	127(6.9)

4.2 Creating a frailty index using data from the MFUS

We created a frailty index according to the available information in the SAQ questionnaire and medical records. We selected items for the frailty index following the standard procedure recommended by Searle et al. (2008); that is, items should meet the requirements of (1) relating to health status, (2) increasing prevalence with age, (3) not saturating too early, and (4) covering a range of systems. A total of 68 items are selected and they are from four categories: activities of daily living (ADL) (7 items), instrumental ADL (11 items), Short Form-36 (SF-36) (11 items),

and diseases (39 items). Table 8 lists some examples of 68 items. Table A1 in Appendix A presents the full list of items.

Table 8. Some examples of items used to construct the frailty index

Category	Variables	Coding
ADL	Unable to do light housework	Yes=1/No=0
	Unable to do heavy housework	Yes=1/No=0
	Unable to do laundry	Yes=1/No=0
	…	
IADL	Unable to go up/down the stairs	Yes=1/No=0
	Unable to wash/bath/groom	Yes=1/No=0
	Unable to get about the house	Yes=1/No=0
	…	
SF-36	Unable to do vigorous activities	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to do moderate activities	Limited=1/Limited a Little=0.5/Not limited=0
	Lifting groceries	Limited=1/Limited a Little=0.5/Not limited=0
	…	
Diseases	Cancer	Yes=1/No=0
	Asthma	Yes=1/No=0
	Stroke	Yes=1/No=0
	…	

For some items in the frailty index, re-coding was required to grade the rank into a score between 0 (where no deficit is present) and 1 (where the deficit is maximally expressed by a given variable). For items with continuous scores, the cutoff points could be used (e.g., for the PSA, the recognized cutoff point is 4.0 ng/ml). An individual's frailty index score was calculated by dividing the sum of deficits present by the maximum possible score. The frailty index ranged from 0 (no deficit) to 1 (all selected deficits present).

To determine the frailty index for a participant at the time of their admission to personal care or nursing home facilities, or when they were assisted by another individual in completing the SAQ,

we adopted an alternative strategy. Instead of imputing missing data within the SAQ, we made use of a substitute value: the mean frailty index calculated for participants living in personal care or nursing homes, with the missing data points excluded. Table 9 shows the number of participants recorded on the questionnaire entering the Personal Care Homes (in the Canadian setting) or Nursing Homes (in the US setting) for each year.

Table 9. Number of participants in personal care/nursing homes for each year

Year	Number of Participants Entering Personal Care
1996	23
2000	9
2002	13
2004	24
2005	18
2006	16
2007	11
2008	21
2009	9
2010	19

Figure 10 shows the histogram of the frailty index in 1996. It shows a skewed distribution, which might suggest the presence of latent classes within the population.

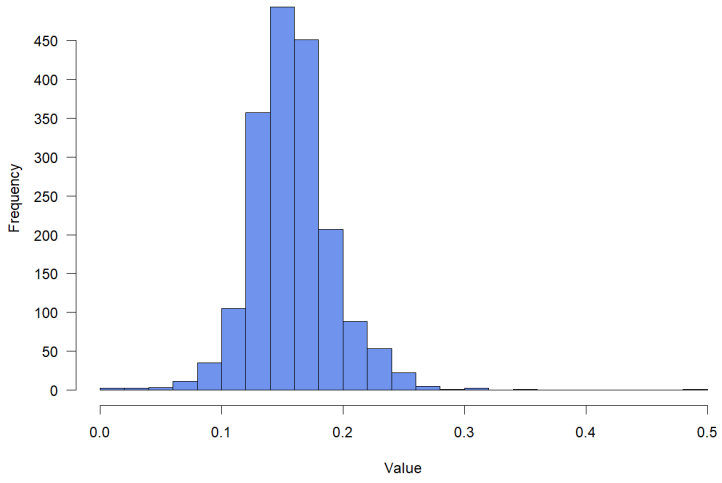

Figure 10. The distribution of baseline frailty index in 1996

Figure 11 shows a spaghetti plot of frailty index (individual trajectories over time) for six randomly selected individuals. This plot highlights the variations among participants, making it easier to identify the potential presence of latent classes within our population.

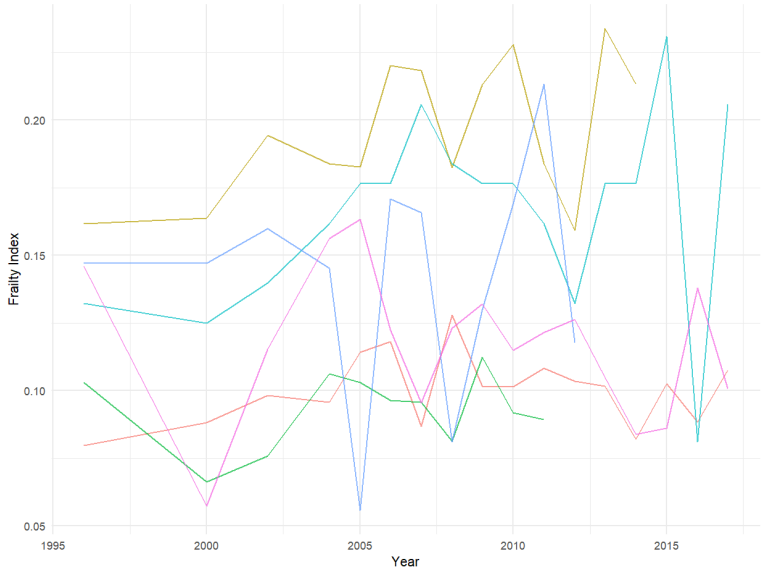

Figure 11. Individual frailty index growth plot over time for six randomly selected participants

4.3 Data analysis of the trajectory of frailty

4.3.1 Conventional group-based trajectory analysis of frailty

The conventional group-based trajectory analysis (GBTA) was first used to identify the distinctive clusters of individual trajectories of frailty within our sample. In this GBTA, the dependent variable was the frailty index at each data collection from 1996 to 2010. The censored normal distribution was used to model the trajectories to account for the censoring at the lower and upper bounds of the frailty index scale. The time variable is coded as years since 1996 at the data collection. The model specified an association extending to a polynomial function to link time to the frailty index for each latent class. The analytic process was initiated with a two-class, progressively adding an additional latent class, until the optimal number of sub-groups was discovered, which best balances goodness-of-fit and parsimony.

For the model selection, the output was scrutinized to find if there were out-of-bound estimates (e.g., negative variances). Secondly, BIC was used to compare the GBTA with a different

number of latent classes. The model was preferred when its BIC value is small. Thirdly, if two models with a different number of sub-groups have similar BIC values, we used entropy values to distinguish the performance of the two models. Entropy is a criterion deriving from a relation linking the likelihood and the classification likelihood of a mixture to measure the precision latent cluster numbers from a mixture model (Celeux & Soromenho, 1996). The model will be preferred with a high value of entropy (close to 1). Further, there are also other measures of performance to compare the model, including a relatively high posterior probability of individual membership (close to 1) and relatively large latent class proportions (above 3% of the population).

The conventional GBTA with 2, 3, 4, and 5 latent classes were compared as well as the trajectory shapes for each latent class. Firstly, each latent class was subjected to scrutinizing the respective trajectory shapes. Each latent class was initially modeled using a polynomial relationship between time and the frailty index, after which the estimated outcomes were compared against the actual trajectory. This comparison facilitated iterative refinement of the polynomial relationship to achieve the alignment with the true trajectory shape for each latent class. Secondly, having ascertained the optimal trajectory shape for each latent class, the Bayesian Information Criterion (BIC) was employed as the primary evaluative metric to ascertain the optimality of the number of latent classes. It is worth emphasizing that the Proc Traj uses a different calculation for the BIC values which makes the larger BIC values instead of smaller indicates better fit.

After evaluating multiple classes, the 4-class conventional GBTA emerged as the most efficient when considering goodness of it, parsimony, and stability. This assessment was supported by the highest BIC value of 21283.29 (please note that SAS PROC Traj has different definition of BIC, with large value indicating better fit), which outperformed the 2-class and 3-class GBTA. Although the 5-class model had a slightly better BIC, it was not chosen due to its complexity, underlining the preference for the simpler, yet adequately fitting, 4-class model. The statistical model of 4-class conventional GBTA is described below.

$$\text{Class 1 - Low(L)}: y_{it}^{(1)} = \beta_0^{(1)} + \epsilon_{it}, \pi^{(1)}$$

$$\text{Class 2 - Moderate(M)}: y_{it}^{(2)} = \beta_0^{(2)} + \epsilon_{it}, \pi^{(2)}$$

$$\text{Class 3 - High (H): } y_{it}^{(3)} = \beta_0^{(3)} + \beta_1^{(3)} \cdot T_{it} + \epsilon_{it}, \pi^{(3)}$$

$$\text{Class 4 - High Increase(HI): } y_{it}^{(4)} = \beta_0^{(4)} + \beta_1^{(4)} \cdot T_{it} + \beta_2^{(4)} \cdot T_{it}^2 + \epsilon_{it}, \pi^{(4)}$$

$$\sum_{j=1}^{4} \pi^{(j)} = 1$$

$$\epsilon_{it} \sim \text{Censored Normal}(0, \sigma_\varepsilon)$$

The $y_{it}^{(j)}$ is the class-specific frailty index for individual i from latent class j ($j = 1, 2, 3, or\ 4$) at measurement t. The T ($T = 0, 1, 2, \dots, 14$) is the realization of the chronological period extending from 1996 through 2010. The estimated results for this model are described in Table 11.

Table 11. Model estimation results for 4-class conventional GBTA.

Variable	Coefficient estimate (SD)	P-value
Class I – Low		
Group proportion	21.47%(2.68%)	<0.001
Intercept	0.13(0.0015)	<0.001
Class II – Moderate		
Group proportion	47.46%(2.17%)	<0.001
Intercept	0.16(0.0016)	<0.001
Class III – High		
Group proportion	24.49%(2.37%)	<0.001
Intercept	0.19(0.0018)	<0.001
Class IV – High Increase		
Group proportion	5.58%(0.85%)	<0.001
Intercept	0.21(0.0040)	<0.001
Linear slope	0.0051(0.0011)	0.0079
Quadratic slope	-0.00021(0.00008)	0.011

The latent classes, derived from the conventional GBTA in this study, were assigned descriptive names to capture their growth characteristics. The "Low" latent class exemplifies a stable trajectory throughout the observational period, maintaining a frailty index at a relatively low level. This class comprises 21.47% of the total population in 1996. Similarly, the "Moderate" latent class exhibits a stable trajectory over the course of the observational years, but with a higher frailty index comparted to the “Low” class. In 1996, the “Moderate” latent class

constituted 48.46% of the total study population. The "High" latent class, on the other hand, delineates the highest stable level in the frailty index over the observational period. In 1996, this group commenced with a relatively elevated frailty level, averaging an estimated frailty index value of 0.19. This class accounted for 24.49% of the total participant count in 1996. Lastly, the "High Increase" latent class starts with the highest frailty level at the onset of the study, with an average frailty index of 0.21. With the highest initial level of frailty, the average frailty level in "High Increase" class is projected to experience an additional increase of 33.33% over the subsequent 14 years. All the estimators were statistically significant.

4.3.2 The extended group-based trajectory analysis of frailty

The extended GBTA was also conducted to examine the frailty trajectory classes. For comparison, we start with specifying the same number of latent classes in the extended GBTA.

$$P\left(Y_i \middle| T_{it}, j; \beta^{(j)}, \theta_t^{(j)}\right) = \prod_{t=1}^{\tau_i - 1} p\left(y_{it} \middle| \omega_{it} = 0, T_{it}, j; \beta^{(j)}\right)\left(1 - \theta_t^{(j)}\right) \theta_{\tau_i}^{(j)}$$

When the attrition due to the death of subject i in class j happens at the measurement τ_i, the indicator ω_{τ_i} equals one at a probability of $\theta_{\tau_i}^{(j)}$. And the probability of attrition not happening is $1 - \theta_t^{(j)}$. The likelihood of the individuals' observed outcomes given the class membership j equals the likelihood of the observed outcomes before attrition multiplied by the rate of staying in cohort $1 - \theta_t^{(j)}$, and multiplies the group-specific dropout probability $\theta_{\tau_i}^{(j)}$.

The group-specific dropout probability $\theta_t^{(j)}$ is estimated using the logit function,

$$\theta_t^{(j)}\left(y_{it-1}; \gamma_j\right) = \frac{1}{1 + e^{\gamma_0 + \gamma_1 y_{it-1}}}$$

where we assume that the probability of attrition due to truncation (e.g., death) is a logit function of the before-attrition frailty level y_{it-1}.

We further assume that attrition happens to each latent class. The optimal number of latent classes, as identified by the conventional GBTA, is employed as the initial values for the extended model. Moreover, to explore possible improvements in model fit, we evaluate the extended model with the number of latent classes adjusted by an increment and decrement of one

unit to the initial optimal number. That is, given the optimal number of latent groups from the conventional GBTA is four, we examine the extended GBTA with three, four, and five latent classes. The choice of the optimal number is guided primarily by BIC.

Upon evaluating extended GBTA with 3, 4, and 5 latent classes, the 4-class extended GBTA emerged as the best-fitted model. This conclusion was substantiated by the higher BIC value, as reported by Proc Traj, which outperformed the 3-class and 5-class. The statistical model of 4-class extended GBTA is described below.

$$\text{Class 1 - Low(L)}: y_{it}^{(1)} = \beta_0^{(1)} + \epsilon_{it}, \pi^{(1)}$$

$$\text{Class 2 - Moderate(M)}: y_{it}^{(2)} = \beta_0^{(2)} + \epsilon_{it}, \pi^{(2)}$$

$$\text{Class 3 - Moderate Increase(MI)}: y_{it}^{(3)} = \beta_0^{(3)} + \beta_1^{(3)} \cdot T_{it} + \epsilon_{it}, \pi^{(3)}$$

$$\text{Class 4 - High Increase(HI)}: y_{it}^{(4)} = \beta_0^{(4)} + \beta_1^{(4)} \cdot T_{it} + \beta_2^{(4)} \cdot T_{it}^2 + \epsilon_{it}, \pi^{(4)}$$

$$\sum_{j=1}^{4} \pi^{(j)} = 1$$

$$\theta_t^{(j)}(y_{it-1}; \gamma_j) = \frac{1}{1 + e^{\gamma_0 + \gamma_1 y_{it-1}}}, j = 1, 2, 3, 4$$

$$\epsilon_{it} \sim \text{Censored Normal}(0, \sigma_\varepsilon)$$

The $y_{it}^{(j)}$ is the class-specific frailty index for individual i from latent class j ($j = 1, 2, 3, or\ 4$) at measurement t generated by the extended GBTA. The T ($T = 0, 1, 2, \dots, 14$) is the realization of the chronological period extending from 1996 through 2010. The estimated results for the growth trajectory are shown in Table 12. Owing to the analysis limiting the cause of attrition to death, it is plausible to interpret the probability of participant dropout in the latent class as the class-specific death rate.

Table 12. Model estimation results for 4-class extended GBTA.

Variable	Coefficient estimate (SD)	P-value
Group I – Low		
Group proportion	17.27%(2.51%)	<0.001
Trajectory – Intercept	0.13(0.0016)	<0.001
Dropout – Intercept*	-1.11 (0.51)	0.030
Dropout – Slope*	-9.82(4.49)	0.029
Group II – Moderate		
Group proportion	39.74%(3.38%)	<0.001
Trajectory – Intercept	0.16(0.0016)	<0.001
Dropout – Intercept	-1.23(0.66)	0.032
Dropout – Slope	-6.29(3.94)	0.11
Group III – Moderate Increase		
Group proportion	35.95%(3.27%)	<0.001
Trajectory - Intercept	0.17(0.0019)	<0.001
Trajectory - Linear slope	0.0020(0.00018)	<0.001
Dropout – Intercept	0	<0.001
Dropout – Slope	-9.36(2.78)	0.008
Group IV – High Increase		
Group proportion	7.02%(1.03%)	<0.001
Trajectory - Intercept	0.21(0.0036)	<0.001
Trajectory - Linear slope	0.0046(0.0010)	<0.001
Trajectory - Quadratic slope	-0.00018(0.000080)	0.024
Dropout – Intercept	-2.12(0.66)	0.0013
Dropout – Slope	2.87(2.54)	0.26

The latent classes identified through the extended GBTA were assigned identical descriptive labels to those used in the conventional GBTA, aligning their growth characteristics. The extended GBTA identified a "Low" latent class that exhibited a stable level of frailty over time, measured at a frailty index of 0.13. This “Low” latent class accounted for 17.27% of the population. The model also identified a "Moderate" latent class, accounting for 39.74% of the population, and characterized by a higher yet stable frailty level compared to the “Low” latent class, marked at a frailty index of 0.16. An increasing latent class named “Moderate Increase” was discerned by the extended GBTA, beginning with an initial frailty index similar to that of the “Moderate” latent class (0.17) in 1996. The “Moderate Increase” class represented 35.95% of the total population. After 14 years of observation, the “Moderate Increase” class experienced an average increase in the frailty index of 16.47%. A more substantial increase rate was observed in the "High Increase" latent class. This class (7.02% of the study cohort), started with the

maximum frailty level in 1996 (0.21). By 2010, the "High Increase" latent class had reached an average peak in the frailty index of 0.24, representing a 30.67% increase from the initial level.

As mentioned above, the cause of dropout was limited to mortality in the analysis. Consequently, the dropout probability could be alternatively conceptualized as the mortality rate for each latent class up until 2010, and this is then projected up to 2015. Detailed results are shown in Table 13.

Table 13. Five-year mortality rate for the 4-class extended group-based trajectory analysis.

Latent class	Five-year Mortality Rate			
	1996-2000	2001-2005	2006-2010	2011-2015
Low	0.050	0.21	0.36	0.48
Moderate	0.065	0.25	0.35	0.53
Moderate Increase	0.34	0.40	0.48	0.64
High Increase	0.25	0.34	0.57	0.70

We observed that the five-year mortality rate increased over time for all latent classes. The "Low" and "Moderate" latent classes consistently displayed a lower level of mortality rate, commencing at 0.05 and 0.07 respectively, and reaching approximately 0.35 in the 2006-2010 observational period. This is consistent with their estimated lower and stable level of frailty index. The two latent classes have a similar projected mortality rate in the 2011-2015 period. The "Moderate Increase" latent class started with a significantly higher mortality level (0.34), which continued to increase over the following years, reaching 0.64 in the projected 2011-2015 observational periods. Although the "High Increase" latent class initially showed a relatively lower mortality rate during the first five-year observational span (0.25), it exhibited a rapid increase in the mortality rate in the subsequent years up to 2015. This is consistent with its averagely its higher average level of frailty index.

4.3.3 Model comparison between the extended and the conventional GBTA

Both the conventional and the extended GBTA yielded a 4-class mode. However, the estimation of group proportion as well as the shape and sizes for each latent class manifested differences, especially for the group proportion estimation. Table 14 shows the comparison of the estimation

results from both the conventional and the extended GBTA. Figure 16 shows estimated trajectories of both models.

Table 14. Model estimation comparison from the two models.

Variable	Conventional GBTA		Extended GBTA	
	Coefficient (SD)	P-value	Coefficient (SD)	P-value
Group I – Low				
Group proportion	21.47%(2.68%)	<0.001	17.27%(2.51%)	<0.001
Intercept	0.13(0.0015)	<0.001	0.13(0.0016)	<0.001
Group II – Moderate				
Group proportion	47.46%(2.17%)	<0.001	39.74%(3.38%)	<0.001
Intercept	0.16(0.0016)	<0.001	0.16(0.0016)	<0.001
Group III – High/Moderate Increase				
Group proportion	24.49%(2.37%)	<0.001	35.95%(3.27%)	<0.001
Intercept	0.19(0.0018)	<0.001	0.17(0.0019)	<0.001
Linear slope			0.0020(0.00018)	<0.001
Group IV – High Increase				
Group proportion	5.58%(0.85%)	<0.001	7.02%(1.03%)	<0.001
Intercept	0.21(0.0040)	<0.001	0.21(0.0036)	<0.001
Linear slope	0.0051(0.0011)	0.0079	0.0046(0.0010)	<0.001
Quadratic slope	-0.00021(0.00008)	0.011	-0.00018(0.000080)	0.0024

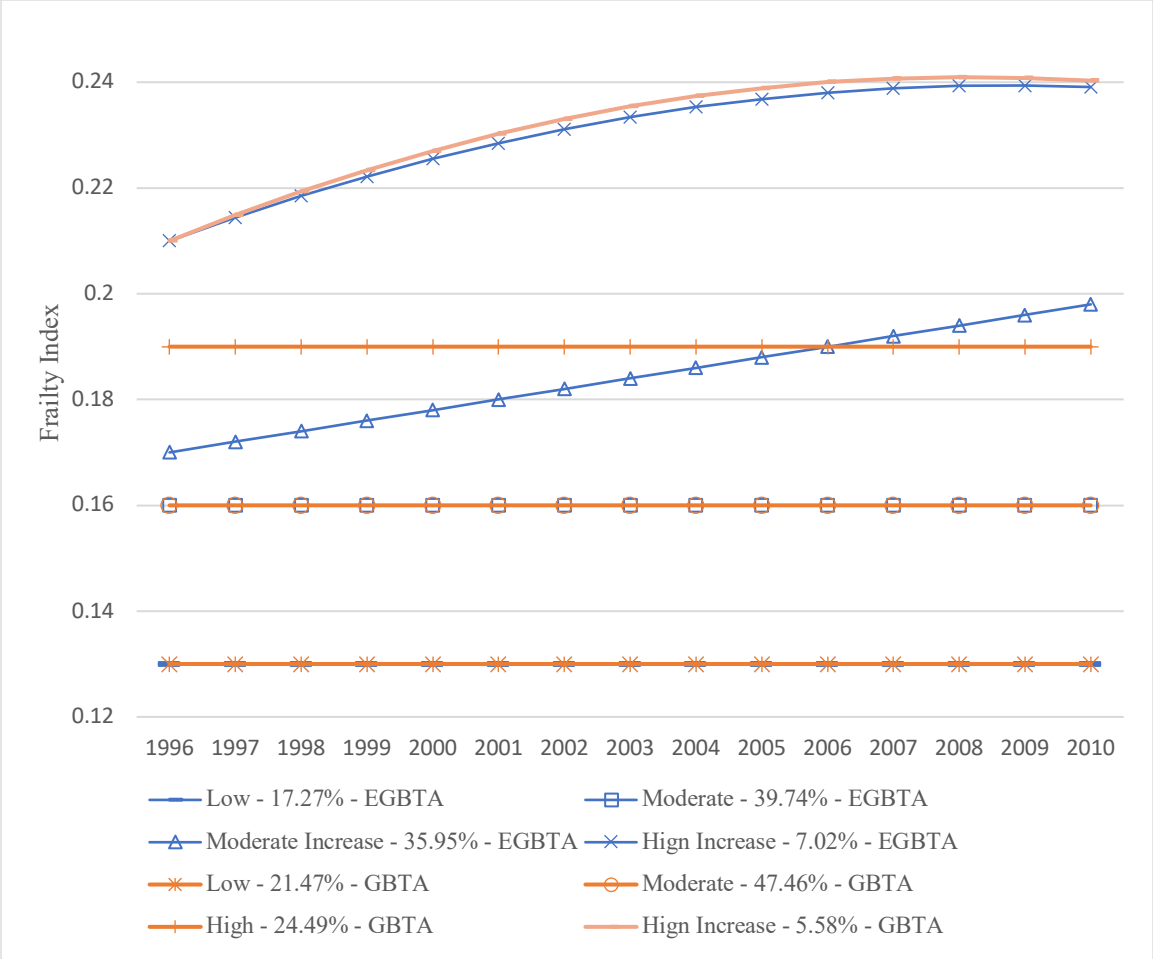

Figure 12. The model comparison between the conventional and the extended GBTA model.

There was a notable difference in the estimation of group proportions for the latent classes between the conventional and extended GBTA models. As mentioned above, the conventional model indicated the "Moderate" group was the largest among the population, accounting for 47.46% of the population. However, the extended model estimated a smaller proportion for the "Moderate" group, at 39.74% of the population. On the other hand, the extended model estimated the "Moderate Increase" latent class with a larger proportion of the population, at 35.95% of the population. This proportion was substantially larger than the estimation from the conventional GBTA model (24.49%), which instead indicated a constant trajectory shape represented by the descriptive name "High".

Upon reflecting on the simulation findings, it was observed that when the two latent classes were not well-separated initially (with the similar intercepts), the group proportion estimation derived

from the conventional GBTA was significantly biased. In contrast, the extended model exhibited resilience to scenarios where latent classes were not well-separated, consistently producing nearly unbiased estimations for the latent class proportions. This situation was observable for both the "Moderate" and "Moderate Increase" latent classes. The "Moderate" latent class, derived from the extended GBTA, demonstrated a starting value of 0.16, close to the starting value of 0.17 for the "Moderate Increase" latent class. This observation aligns with the scenario of the latent classes not being well-separated, as specified in the simulation design. Based on the simulation findings, the group proportion estimation derived from the extended GBTA ("Moderate Increase": 39.74%; "High Increase": 35.95%) appears to be more reliable when compared to the estimation from the conventional GBTA.

In summary, we have leveraged extended GBTA method to model the frailty trajectories using the MFUS data set. Our conclusions are based on the results obtained from earlier simulation studies, which indicated the superiority of extended GBTA in discerning a 4-class frailty trajectory. Our analysis revealed that 17.27% of the participants maintained a consistently low level of frailty over the years ("Low" latent class), demonstrating a negligible progression in frailty. Contrarily, about 39.74% of the participants exhibited a higher level of frailty that remained constant throughout the years ("Moderate" latent class). Additionally, 35.95% of the participants, who initially displayed a similar level of frailty to those in the "Moderate" latent class, saw their frailty levels increase over time ("Moderate Increase" latent class). A minority group, approximately 7.02% of the study population, started with a significantly high frailty level that worsened progressively towards the end of the observation period ("High Increase" latent class). In terms of mortality rate, the highest rates were observed in the "Moderate Increase" and "High Increase" latent classes.

4.4 Linking age, living arrangements, and marital status to the group membership

Building upon the findings from our extended GBTA, we further explored the potential association between the participant group membership and variables such as age, living arrangements, and marital status utilizing multinomial logistic regression. The outcome variable was the predicted group membership gleaned from the extended GBTA, conceptualized as a 4-category nominal variable. The predictors encompassed baseline age (considered as a continuous variable), living arrangement (binary), and marital status (binary), each recorded in 1996.

Participants with missing values in any of these three predictors were excluded from the analysis, resulting in a final cohort size of 1,627 participants.

Let outcome classes denoted as j, and set "Low" latent class $j = 1$, "Moderate" latent class $j = 2$, "Moderate Increase" latent class $j = 3$, and "High Increase" latent class $j = 4$. The "Low" latent class was set to be the reference class when running the multinomial logistic regression. The x_1, x_2, and x_3 denote factors of age, living arrangements (living alone = 1), and marital status (married = 1) at the baseline. and let $P(y = j|x_1, x_2, x_3)$ denote the probability of group membership in group j given the three predictors. The logit function of the conditional probability is assumed to follow a linear combination of intercept and predictors. The statistical model is shown below.

$$logit\big(P(y = j|x_1, x_2, x_3)\big) = \lambda_{j0} + \lambda_{j1} \cdot x_1 + \lambda_{j2} \cdot x_2 + \lambda_{j3} \cdot x_3$$

where λ_{j0}, λ_{j0}, λ_{j0}, and λ_{j0} are parameters for the jth outcome class $(j = 2, 3, 4)$. To ease interpretation, it is common to exponentiate these coefficients. For example, the parameters $e^{\lambda_{j0}}$ is the estimated odds of the participant being in the latent class j versus the reference class when all predictors are zero. The estimated odds and the 95% confidence interval (CI) with significance level is shown in Table 15.

Table 15. Model estimation results for multinomial logistic regression

Characteristics	Latent Classes from the Extended GBTA			
	Low	Moderate	Moderate Increase	High Increase
Age	1	0.962 (0.922, 1.003) *	1.050 (1.009, 1.093) ***	1.044 (0.979, 1.113)
Living alone	1	0.773 (0.545, 1.096)	0.740 (0.524, 1.047) *	0.663 (0.359, 1.226)
Married	1	1.177 (0.801, 1.728)	1.295 (0.882, 1.899)	1.479 (0.743, 2.943)

The effect of age on latent class membership is statistically significant for the "Moderate Increase" latent class when compared to the "Low" latent class as the reference group. Specifically, for each one-year increment in age, there is a 5% increase in the odds of an individual being classified in the "Moderate Increase" latent class rather than the "Low" latent class.

The impact of living alone on latent class membership is only marginally statistically significant for the "Moderate Increase" latent class when comparing it to the "Low" latent class.

Specifically, participants who were living alone (assigned value of one) had about 26% lower odds of being in the “Moderate” frailty index latent class compared to the “Low” latent class, compared to those who were living alone in 1996, assuming age and marital status remained constant.

After analyzing the impact of marital status, it has been determined that there is no statistically significant relationship between the class membership estimation and the participants’ marital status in 1996.

5. Discussion

The findings of the thesis project have demonstrated the advantages of the extended Group-based Trajectory Analysis (GBTA) over the conventional GBTA for longitudinal data featuring non-ignorable missing or dropouts.

Our simulation studies reveal that when latent classes are not well-separated, the conventional GBTA exhibits substantial bias in the estimation of class proportions, especially when the attrition rate is high. However, the extended GBTA maintains accuracy in estimating the proportion of latent classes under varying missing data mechanisms and attrition rates. When the latent classes are well-separated, both the extended and conventional GBTA can give accurate estimate of the class proportions. When analyzing Manitoba Follow-Up Study (MFUS) data, we found that the predictions from the extended GBTA were closer to expectations, even though it remains uncertain whether the extended or conventional GBTA provides more accurate estimations of frailty trajectory class proportions.

Haviland et al. (2011) also found that that the latent class proportion estimation generated by conventional GBTA could be heavily biased due to differential attrition rates, particularly when the classes were initially not well separated. However, they only considered one missing mechanism - Missing Completely at Random (MCAR), and a constant dropout rate over time. This thesis project extended their findings to more complex missing data mechanisms like MAR and MNAR. Additionally, this thesis contributes to the current literature by expanding our understanding of the effect of nonrandomly missing data on modeling developmental trajectories.

In investigation of frailty trajectories using the MFUS data, both the conventional and the extended GBTA identified four trajectory classes: two classes with stable levels of frailty index (one with low level of frailty index and one with moderate level) and two classes with increased level of frailty over time (one with moderate level of frailty index and one with high level at baseline). Two increasing classes reported higher mortality rates over time.

Stow et al. (2018) investigated the frailty trajectories one-year before death and found three distinct trajectories of frailty: rapid rise, moderate increase, and stability. The percentages of participants fitting each of these categories were 2.2%, 21.2%, and 76.6% respectively. Their findings also suggested that the group with rapidly increasing frailty experienced a 180% surge in mortality, while the group with moderately increasing frailty saw a 65% rise compared to the

stable frailty group. However, Stow et al. (2018) only conducted the conventional GBTA, and their results could be biased due to the lack of consideration of the nonignorable missing mechanism.

Our analysis also indicated that older age was associated with a higher probability of being in the two increased frailty trajectory classes. This finding aligns with Welstead et al. (2020).

Our analyses revealed that those people who lived alone had decreased odds of being the "Moderate" frailty class as opposed to the "Low" frailty class. However, this finding contradicts a systematic review by Kojima, Taniguchi, et al. (2020), which synthesized results from 203 studies demonstrating a significant association between living alone and frailty risk. This discrepancy may result from the predominance of cross-sectional studies in the systematic review, which contrast with our longitudinal study design. Furthermore, the cohort used for this thesis, derived from the MFUS, consisted of former members of the Royal Canadian Air Force who have maintained relative health over the years. This factor may skew our findings compared to those based on a more generalized population.

We found no relationship between the marital status and frailty trajectories, contradicting the findings by Kojima, Walters, et al. (2020). This discrepancy may stem from our analysis used marital status at baseline. As participants aged, a significant shift occurred in marital status, with an observed increase in the widowed category and a decrease in the married category. However, incorporating time-varying covariates in class enumeration exceeds the capabilities of the proposed statistical method.

Significance. Our research has uncovered four distinct latent classes of frailty growth among older men. These findings have significant implications for the formulation of public health policy and intervention strategies. By identifying and categorizing these different patterns of frailty trajectories, we can develop more focused and personalized interventions for each specific pattern. The association between age and the membership of a particular latent class underscores the complexity and multifaceted nature of frailty growth trajectories. This correlation suggests that as men age, there could be varying risk levels and patterns of frailty progression. Recognizing such age-specific variations is crucial for health policymakers, as it enables them to design age-appropriate interventions that are appropriate for each age group, rather than using a one-size-fits-all approach.

The proposed extended GBTA offered the applied researcher with an innovative method for addressing non-random attrition caused by truncated events. Researchers who utilize longitudinal studies to analyze the ageing process often encounter the challenge of non-ignorable missing data due to death, especially when studying a population that is already advanced in age at the initiation of the study. Our proposed method could address this issue by incorporating the non-ignorable missingness into the enumeration of latent classes. This allows us to accurately estimate the true size of the latent classes and provides an alternative approach for analyzing the longitudinal time-event data that may involve the presence of latent classes.

Understanding the latent classes of frailty growth and their associated factors provides a more nuanced understanding of the heterogeneous nature of frailty progression among older men. It is paramount for public health policymakers to recognize and act upon these findings. The introduction of the extended model provides the applied researcher with a valuable tool for addressing the non-ignorable missingness in their analyses.

Limitation. The current study has several limitations. Firstly, In the process of estimating our extended model, we utilized the SAS Proc Traj procedure. It is crucial to acknowledge its inherent limitation, as the Proc Traj can only employ prior observations in modeling the attrition process. This limitation significantly impacts scenarios where the mechanism of missingness is MNAR, a situation where the missing data depend on unobserved values. Such a restriction can potentially amplify the bias in the estimated results. Secondly, when employing the conventional and extended GBTA to investigate frailty growth using the MFUS data, we limited our inclusion to baseline measurements of living arrangements and marital status, although most participants did not experience changes in these factors during the observation period used for modeling. This limitation inherent in the extended GBTA is due to its inability to incorporate time-varying factors related to latent class membership. Lastly, despite the MFUS data set being rich in information and currently the longest running longitudinal research in Canada, it focuses exclusively on male participants. Furthermore, the target population of the MFUS is former members of the Royal Canadian Air Force, which increases the challenges of achieving external validity for this thesis.

Future Directions. The current thesis study provides several suggestions for future research. Considering dropout as a terminal event, we can also conduct latent class joint models of the

longitudinal data (e.g., frailty index) and survival outcomes (e.g., death). Comparisons between latent class joint models and extended GBTA approach might be interesting, which could potentially overcome the limitation of using SAS Proc Traj procedure.

For the statistical model development, the extended GBTA could be further advanced by incorporating the time-varying covariates associated with class membership and incorporating the would-be value for estimating the attrition process. Also, further development in the GBTA could allow for participant's latent class membership changing during the observation period.

For the frailty trajectory analysis, socio-demographic factors such as marital status and living arrangements need to be further analyzed, especially under the longitudinal study designs. The current consistent findings are primarily derived from studies employing cross-sectional designs, thereby limiting their ability to account for the influence of these factors on frailty development. Furthermore, using a more representative cohort to conduct the analysis could strengthen the external validity of the statistical findings.

Attrition is common in longitudinal studies, especially among the older population who are susceptible to health and functional decline, illness, and death. The proposed study has shown the importance of taking nonignorable attrition into consideration when conducting trajectory analysis using longitudinal data. Analyzing the frailty trajectories and their linkages to death will help researchers develop a more precise understanding of the heterogeneity within the population regarding frailty changes over time. Understanding different risk factors associated with diverse frailty trajectories will inform us about who is at the greatest risk and how the potential intervention promoting general welfare should be implemented in the future.

Reference

Aranda, M. P., Ray, L. A., Snih, S. A., Ottenbacher, K. J. & Markides, K. S. (2011). The Protective Effect of Neighborhood Composition on Increasing Frailty Among Older Mexican Americans. *Journal of Aging and Health*, *23*(7), 1189–1217. https://doi.org/10.1177/0898264311421961

Arnold, M., J., M., Alexander & Kenneth, R. (2001). Accumulation of Deficits as a Proxy Measure of Aging. *The Scientific World Journal*, *1*, 323–336. https://doi.org/10.1100/tsw.2001.58

Baek, W. & Min, A. (2022). Frailty index and gender-specific mortality in Korean adults: Findings from the Korean Longitudinal Study of Aging (2006–2018). *Journal of Advanced Nursing*, *78*(8), 2397–2407. https://doi.org/10.1111/jan.15168

Bessa, B., Ribeiro, O. & Coelho, T. (2018). Assessing the social dimension of frailty in old age: A systematic review. *Archives of Gerontology and Geriatrics*, *78*, 101–113. https://doi.org/10.1016/j.archger.2018.06.005

Bollen, K. A. & Curran, P. J. (2006). *Latent Curve Models: A Structural Equation Perspective*. 1–15. https://doi.org/10.1002/0471746096.ch1

Brigola, A. G., Alexandre, T. da S., Inouye, K., Yassuda, M. S., Pavarini, S. C. I. & Mioshi, E. (2019). Limited formal education is strongly associated with lower cognitive status, functional disability and frailty status in older adults. *Dementia & Neuropsychologia*, *13*(2), 216–224. https://doi.org/10.1590/1980-57642018dn13-020011

Celeux, G. & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, *13*(2), 195–212. https://doi.org/10.1007/bf01246098

Cesari, M., Gambassi, G., Kan, G. A. van & Vellas, B. (2014). The frailty phenotype and the frailty index: different instruments for different purposes. *Age and Ageing*, *43*(1), 10–12. https://doi.org/10.1093/ageing/aft160

Cesari, M., Prince, M., Thiyagarajan, J. A., Carvalho, I. A. D., Bernabei, R., Chan, P., Gutierrez-Robledo, L. M., Michel, J.-P., Morley, J. E., Ong, P., Manas, L. R., Sinclair, A., Won, C. W., Beard, J. & Vellas, B. (2016). Frailty: An Emerging Public Health Priority. *Journal of the American Medical Directors Association*, *17*(3), 188–192. https://doi.org/10.1016/j.jamda.2015.12.016

Choi, J., Ahn, A., Kim, S. & Won, C. W. (2015). Global Prevalence of Physical Frailty by Fried's Criteria in Community-Dwelling Elderly With National Population-Based Surveys. *Journal of the American Medical Directors Association*, *16*(7), 548–550. https://doi.org/10.1016/j.jamda.2015.02.004

Clegg, A., Young, J., Iliffe, S., Rikkert, M. O. & Rockwood, K. (2013). Frailty in elderly people. *The Lancet*, *381*(9868), 752–762. https://doi.org/10.1016/s0140-6736(12)62167-9

Collard, R. M., Boter, H., Schoevers, R. A. & Voshaar, R. C. O. (2012). Prevalence of Frailty in Community-Dwelling Older Persons: A Systematic Review. *Journal of the American Geriatrics Society*, *60*(8), 1487–1492. https://doi.org/10.1111/j.1532-5415.2012.04054.x

Curran, P. J., Obeidat, K. & Losardo, D. (2010). Twelve Frequently Asked Questions About Growth Curve Modeling. *Journal of Cognition and Development*, *11*(2), 121–136. https://doi.org/10.1080/15248371003699969

Dent, E., Kowal, P. & Hoogendijk, E. O. (2016). Frailty measurement in research and clinical practice: A review. *European Journal of Internal Medicine*, *31*, 3–10. https://doi.org/10.1016/j.ejim.2016.03.007

Ferrucci, L., Guralnik, J. M., Studenski, S., Fried, L. P., Cutler, G. B., Walston, J. D. & Group, T. I. on F. W. (2004). Designing Randomized, Controlled Trials Aimed at Preventing or Delaying Functional Decline and Disability in Frail, Older Persons: A Consensus Report. *Journal of the American Geriatrics Society*, *52*(4), 625–634. https://doi.org/10.1111/j.1532-5415.2004.52174.x

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., Seeman, T., Tracy, R., Kop, W. J., Burke, G., McBurnie, M. A. & Group, C. H. S. C. R. (2001). Frailty in Older AdultsEvidence for a Phenotype. *The Journals of Gerontology: Series A*, *56*(3), M146–M157. https://doi.org/10.1093/gerona/56.3.m146

Gale, C. R., Cooper, C. & Sayer, A. A. (2015). Prevalence of frailty and disability: findings from the English Longitudinal Study of Ageing. *Age and Ageing*, *44*(1), 162–165. https://doi.org/10.1093/ageing/afu148

Goldstein, H., Bryk, A. S. & Raudenbush, S. W. (1993). Hierarchical Linear Models: Applications and Data Analysis Methods. *Journal of the American Statistical Association*, *88*(421), 386. https://doi.org/10.2307/2290750

Hardy, S. E., Allore, H. & Studenski, S. A. (2009). Missing Data: A Special Challenge in Aging Research. *Journal of the American Geriatrics Society*, *57*(4), 722–729. https://doi.org/10.1111/j.1532-5415.2008.02168.x

Haviland, A. M., Jones, B. L. & Nagin, D. S. (2011). Group-based Trajectory Modeling Extended to Account for Nonrandom Participant Attrition. *Sociological Methods & Research*, *40*(2), 367–390. https://doi.org/10.1177/0049124111400041

Hofer, S. M. & Sliwinski, M. J. (2006). Handbook of the Psychology of Aging. *Part One: Concepts, Theory, and Methods in the Psychology of Aging*, *Psychology and Aging 10 1995*, 15–37. https://doi.org/10.1016/b978-012101264-9/50005-7

Hoogendijk, E. O., Hout, H. P. J. van, Heymans, M. W., Horst, H. E. van der, Frijters, D. H. M., Groenou, M. I. B. van, Deeg, D. J. H. & Huisman, M. (2014). Explaining the association between educational level and frailty in older adults: results from a 13-year longitudinal study in the Netherlands. *Annals of Epidemiology*, *24*(7), 538-544.e2. https://doi.org/10.1016/j.annepidem.2014.05.002

Hwang, A.-C., Lee, W.-J., Huang, N., Chen, L.-Y., Peng, L.-N., Lin, M.-H., Chou, Y.-J. & Chen, L.-K. (2021). Longitudinal changes of frailty in 8 years: comparisons between physical frailty and frailty index. *BMC Geriatrics*, *21*(1), 726. https://doi.org/10.1186/s12877-021-02665-1

Ibrahim, J. G. & Molenberghs, G. (2009). Missing data methods in longitudinal studies: a review. *TEST*, *18*(1), 1–43. https://doi.org/10.1007/s11749-009-0138-x

John, P. D. St., Jiang, D. & Tate, R. B. (2018). Quality of Life Trajectories Predict Mortality in Older Men: The Manitoba Follow-Up Study. *Journal of Aging and Health*, *30*(2), 247–261. https://doi.org/10.1177/0898264316673714

John, P. D. St., McClement, S. S., Swift, A. U. & Tate, R. B. (2019). Older Men's Definitions of Frailty – The Manitoba Follow-up Study. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, *38*(1), 13–20. https://doi.org/10.1017/s0714980818000405

Jones, B. L. & Nagin, D. S. (2007). Advances in Group-Based Trajectory Modeling and an SAS Procedure for Estimating Them. *Sociological Methods & Research*, *35*(4), 542–571. https://doi.org/10.1177/0049124106292364

Kojima, G. (2015). Prevalence of Frailty in Nursing Homes: A Systematic Review and Meta-Analysis. *Journal of the American Medical Directors Association*, *16*(11), 940–945. https://doi.org/10.1016/j.jamda.2015.06.025

Kojima, G. (2016). Frailty as a predictor of hospitalisation among community-dwelling older people: a systematic review and meta-analysis. *Journal of Epidemiology and Community Health*, *70*(7), 722. https://doi.org/10.1136/jech-2015-206978

Kojima, G., Iliffe, S. & Walters, K. (2017). Frailty index as a predictor of mortality: a systematic review and meta-analysis. *Age and Ageing*, *47*(2), 193–200. https://doi.org/10.1093/ageing/afx162

Kojima, G., Taniguchi, Y., Kitamura, A. & Fujiwara, Y. (2020). Is living alone a risk factor of frailty? A systematic review and meta-analysis. *Ageing Research Reviews*, *59*, 101048. https://doi.org/10.1016/j.arr.2020.101048

Kojima, G., Walters, K., Iliffe, S., Taniguchi, Y. & Tamiya, N. (2020). Marital Status and Risk of Physical Frailty: A Systematic Review and Meta-analysis. *Journal of the American Medical Directors Association*, *21*(3), 322–330. https://doi.org/10.1016/j.jamda.2019.09.017

Kulminski, A., Yashin, A., Arbeev, K., Akushevich, I., Ukraintseva, S., Land, K. & Manton, K. (2007). Cumulative index of health disorders as an indicator of aging-associated processes in the elderly: Results from analyses of the National Long Term Care Survey. *Mechanisms of Ageing and Development*, *128*(3), 250–258. https://doi.org/10.1016/j.mad.2006.12.004

Laird, N. M. & Ware, J. H. (1982). Random-Effects Models for Longitudinal Data. *Biometrics*, *38*(4), 963. https://doi.org/10.2307/2529876

Lee, D. R., Kawas, C. H., Gibbs, L. & Corrada, M. M. (2016). Prevalence of Frailty and Factors Associated with Frailty in Individuals Aged 90 and Older: The 90+ Study. *Journal of the American Geriatrics Society*, *64*(11), 2257–2262. https://doi.org/10.1111/jgs.14317

Lengyel, C. O., Jiang, D. & Tate, R. B. (2017). Trajectories of nutritional risk: The Manitoba follow-up study. *The Journal of Nutrition, Health & Aging*, *21*(6), 604–609. https://doi.org/10.1007/s12603-016-0826-7

Liu, Z., Han, L., Gahbauer, E. A., Allore, H. G. & Gill, T. M. (2018). Joint Trajectories of Cognition and Frailty and Associated Burden of Patient-Reported Outcomes. *Journal of the American Medical Directors Association*, *19*(4), 304-309.e2. https://doi.org/10.1016/j.jamda.2017.10.010

Lo, Y. (2001). Testing the number of components in a normal mixture. *Biometrika*, *88*(3), 767–778. https://doi.org/10.1093/biomet/88.3.767

McIsaac, D. I., Taljaard, M., Bryson, G. L., Beaulé, P. E., Gagné, S., Hamilton, G., Hladkowicz, E., Huang, A., Joanisse, J. A., Lavallée, L. T., MacDonald, D., Moloo, H., Thavorn, K., Walraven, C. van, Yang, H. & Forster, A. J. (2018). Frailty as a Predictor of Death or New Disability After Surgery. *Annals of Surgery*, *Publish Ahead of Print*(NA;), NA; https://doi.org/10.1097/sla.0000000000002967

McNeish, D. & Matta, T. (2018). Differentiating between mixed-effects and latent-curve approaches to growth modeling. *Behavior Research Methods*, *50*(4), 1398–1414. https://doi.org/10.3758/s13428-017-0976-5

Meredith, W. & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*(1), 107–122. https://doi.org/10.1007/bf02294746

Mitnitski, A., Song, X., Skoog, I., Broe, G., Cox, J. L., Grunfeld, E. & Rockwood, K. (2005). Relative Fitness and Frailty of Elderly Men and Women in Developed Countries and Their Relationship with Mortality. *Journal of the American Geriatrics Society*, *53*(12), 2184–2189. https://doi.org/10.1111/j.1532-5415.2005.00506.x

Muthén, B. (2001). *New methods for the analysis of change.* 291–322. https://doi.org/10.1037/10409-010

Myers, V., Drory, Y., Goldbourt, U. & Gerber, Y. (2014). Multilevel socioeconomic status and incidence of frailty post myocardial infarction. *International Journal of Cardiology*, *170*(3), 338–343. https://doi.org/10.1016/j.ijcard.2013.11.009

Nagin, D. S. (1999). Analyzing Developmental Trajectories: A Semiparametric, Group-Based Approach. *Psychological Methods*, *4*(2), 139–157. https://doi.org/10.1037/1082-989x.4.2.139

Nagin, D. S. & Odgers, C. L. (2010a). Group-Based Trajectory Modeling in Clinical Research. *Annual Review of Clinical Psychology*, *6*(1), 109–138. https://doi.org/10.1146/annurev.clinpsy.121208.131413

Nagin, D. S. & Odgers, C. L. (2010b). Group-Based Trajectory Modeling in Clinical Research. *Annual Review of Clinical Psychology*, *6*(1), 109–138. https://doi.org/10.1146/annurev.clinpsy.121208.131413

Newman, A. B. (2010). An Overview of the Design, Implementation, and Analyses of Longitudinal Studies on Aging. *Journal of the American Geriatrics Society*, *58*, S287–S291. https://doi.org/10.1111/j.1532-5415.2010.02916.x

Ng, T. P., Feng, L., Nyunt, M. S. Z., Larbi, A. & Yap, K. B. (2014). Frailty in Older Persons: Multisystem Risk Factors and the Frailty Risk Index (FRI). *Journal of the American Medical Directors Association*, *15*(9), 635–642. https://doi.org/10.1016/j.jamda.2014.03.008

Nicholson, J. S., Deboeck, P. R. & Howard, W. (2017). Attrition in developmental psychology. *International Journal of Behavioral Development*, *41*(1), 143–153. https://doi.org/10.1177/0165025415618275

Ntanasi, E., Yannakoulia, M., Mourtzi, N., Vlachos, G. S., Kosmidis, M. H., Anastasiou, C. A., Dardiotis, E., Hadjigeorgiou, G., Megalou, M., Sakka, P. & Scarmeas, N. (2020). Prevalence and Risk Factors of Frailty in a Community-Dwelling Population: The HELIAD Study. *Journal of Aging and Health*, *32*(1), 14–24. https://doi.org/10.1177/0898264318801735

Peek, M. K., Howrey, B. T., Ternent, R. S., Ray, L. A. & Ottenbacher, K. J. (2012). Social Support, Stressors, and Frailty Among Older Mexican American Adults. *The Journals of Gerontology: Series B*, *67*(6), 755–764. https://doi.org/10.1093/geronb/gbs081

Pereira, A. A., Borim, F. S. A. & Neri, A. L. (2017). Risk of death in elderly persons based on the frailty phenotype and the frailty index: a review study. *Revista Brasileira de Geriatria e Gerontologia*, *20*(02), 273–285. https://doi.org/10.1590/1981-22562017020.160102

Pérez-Zepeda, M. U., Godin, J., Armstrong, J. J., Andrew, M. K., Mitnitski, A., Kirkland, S., Rockwood, K. & Theou, O. (2020). Frailty among middle-aged and older Canadians: population norms for the frailty index using the Canadian Longitudinal Study on Aging. *Age and Ageing*, *50*(2), 447–456. https://doi.org/10.1093/ageing/afaa144

Pilotto, A., Ferrucci, L., Franceschi, M., D'Ambrosio, L. P., Scarcelli, C., Cascavilla, L., Paris, F., Placentino, G., Seripa, D., Dallapiccola, B. & Leandro, G. (2008). Development and Validation of a Multidimensional Prognostic Index for One-Year Mortality from Comprehensive Geriatric Assessment in Hospitalized Older Patients. *Rejuvenation Research*, *11*(1), 151–161. https://doi.org/10.1089/rej.2007.0569

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111. https://doi.org/10.2307/271063

Raina, P. S., Wolfson, C., Kirkland, S. A., Griffith, L. E., Oremus, M., Patterson, C., Tuokko, H., Penning, M., Balion, C. M., Hogan, D., Wister, A., Payette, H., Shannon, H. & Brazil, K. (2009). The Canadian Longitudinal Study on Aging (CLSA)*. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, *28*(3), 221–229. https://doi.org/10.1017/s0714980809990055

Ram, N. & Grimm, K. J. (2009). Methods and Measures: Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, *33*(6), 565–576. https://doi.org/10.1177/0165025409343765

Rockwood, K & Mitnitski, A. (2012). HOW MIGHT DEFICIT ACCUMULATION GIVE RISE TO FRAILTY? *Journal of Frailty & Aging*, *1*(1), 1–4. https://doi.org/10.14283/jfa.2012.2

Rockwood, Kenneth & Mitnitski, A. (2007). Frailty in Relation to the Accumulation of Deficits. *The Journals of Gerontology: Series A*, *62*(7), 722–727. https://doi.org/10.1093/gerona/62.7.722

Rockwood, Kenneth, Song, X., MacKnight, C., Bergman, H., Hogan, D. B., McDowell, I. & Mitnitski, A. (2005). A global clinical measure of fitness and frailty in elderly people. *Canadian Medical Association Journal*, *173*(5), 489–495. https://doi.org/10.1503/cmaj.050051

Romero-Ortuno, R. (2013). An alternative method for Frailty Index cut-off points to define frailty categories. *European Geriatric Medicine*, *4*(5), 299–303. https://doi.org/10.1016/j.eurger.2013.06.005

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika*, *63*(3), 581. https://doi.org/10.2307/2335739

Rubin, D. B. (2008). Multiple Imputation for Nonresponse in Surveys. *Wiley Series in Probability and Statistics*, 1–26. https://doi.org/10.1002/9780470316696.ch1

Sachs, E., John, P. St., Swift, A. & Tate, R. (2021). Self-Rated Frailty and Mortality in Old Men: The Manitoba Follow-Up Study. *The Journal of Frailty & Aging*, *10*(1), 44–48. https://doi.org/10.14283/jfa.2020.14

Sager, M. A., Rudberg, M. A., Jalaluddin, M., Franke, T., Inouye, S. K., Landefeld, C. S., Siebens, H. & Winograd, C. H. (1996). Hospital Admission Risk Profile (HARP): Identifying Older Patients at Risk for Functional Decline Following Acute Medical Illness and Hospitalization. *Journal of the American Geriatrics Society*, *44*(3), 251–257. https://doi.org/10.1111/j.1532-5415.1996.tb00910.x

Searle, S. D., Mitnitski, A., Gahbauer, E. A., Gill, T. M. & Rockwood, K. (2008). A standard procedure for creating a frailty index. *BMC Geriatrics*, *8*(1), 24–24. https://doi.org/10.1186/1471-2318-8-24

Siriwardhana, D. D., Hardoon, S., Rait, G., Weerasinghe, M. C. & Walters, K. R. (2018). Prevalence of frailty and prefrailty among community-dwelling older adults in low-income and middle-income countries: a systematic review and meta-analysis. *BMJ Open*, *8*(3), e018195. https://doi.org/10.1136/bmjopen-2017-018195

Stow, D., Matthews, F. E. & Hanratty, B. (2018). Frailty trajectories to identify end of life: a longitudinal population-based study. *BMC Medicine*, *16*(1), 171. https://doi.org/10.1186/s12916-018-1148-x

Tate, R. B., Cuddy, T. E. & Mathewson, F. A. (2015). Cohort Profile: The Manitoba Follow-up Study (MFUS). *International Journal of Epidemiology*, *44*(5), 1528–1536. https://doi.org/10.1093/ije/dyu141

Vermeulen, J., Neyens, J. C., Rossum, E. van, Spreeuwenberg, M. D. & Witte, L. P. de. (2011). Predicting ADL disability in community-dwelling elderly people using physical frailty indicators: a systematic review. *BMC Geriatrics*, *11*(1), 33. https://doi.org/10.1186/1471-2318-11-33

Walston, J., Hadley, E. C., Ferrucci, L., Guralnik, J. M., Newman, A. B., Studenski, S. A., Ershler, W. B., Harris, T. & Fried, L. P. (2006). Research Agenda for Frailty in Older Adults: Toward a Better Understanding of Physiology and Etiology: Summary from the American Geriatrics Society/National Institute on Aging Research Conference on Frailty in Older Adults. *Journal of the American Geriatrics Society*, *54*(6), 991–1001. https://doi.org/10.1111/j.1532-5415.2006.00745.x

Ware, J. E. & Sherbourne, C. D. (1992). The MOS 36-ltem Short-Form Health Survey (SF-36). *Medical Care*, *30*(6), 473–483. https://doi.org/10.1097/00005650-199206000-00002

Welstead, M., Jenkins, N. D., Russ, T., Luciano, M. & Muniz-Terrera, G. (2020). A Systematic Review of Frailty Trajectories: Their Shape And Influencing Factors. *The Gerontologist*, *61*(8), gnaa061-. https://doi.org/10.1093/geront/gnaa061

Woods, N. F., LaCroix, A. Z., Gray, S. L., Aragaki, A., Cochrane, B. B., Brunner, R. L., Masaki, K., Murray, A., Newman, A. B. & Initiative, W. H. (2005). Frailty: Emergence and Consequences in Women Aged 65 and Older in the Women's Health Initiative Observational

Study. *Journal of the American Geriatrics Society*, *53*(8), 1321–1330. https://doi.org/10.1111/j.1532-5415.2005.53405.x

Xue, Q.-L. (2011). The Frailty Syndrome: Definition and Natural History. *Clinics in Geriatric Medicine*, *27*(1), 1–15. https://doi.org/10.1016/j.cger.2010.08.009

Appendix

Table A1. The full list of the candidate variables for the frailty index from the MFUS (N = 68)

Category	Variables	Coding
ADL	Unable to do light housework	Yes=1/No=0
	Unable to do heavy housework	Yes=1/No=0
	Unable to do laundry	Yes=1/No=0
	Unable to do a hot meal	Yes = 1/No=0
	Unable to shovel and yardwork	Yes=1/No=0
	Unable to do shopping	Yes=1/No=0
	Unable to manage financial affairs	Yes=1/No=0
IADL	Unable to go up/down the stairs	Yes=1/No=0
	Unable to wash/bath/groom	Yes=1/No=0
	Unable to get about the house	Yes=1/No=0
	Unable to go out of doors in good weather	Yes=1/No=0
	Unable to get in and out of bed	Yes=1/No=0
	Unable to dress and putting shoes on	Yes=1/No=0
	Unable to eat	Yes=1/No=0
	Unable to take medication or treatment	Yes=1/No=0
	Unable to use toilet	Yes=1/No=0
	Unable to use of telegraph	Yes=1/No=0
	Unable to get up out of a chair and walk 3 meters	Yes=1/No=0
SF-36	Unable to do vigorous activities	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to do moderate activities	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to lift groceries	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to climb several flights of stairs	Limited=1/Limited a Little=0.5/Not limited=0

	Unable to climb one flight of stairs	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to bend, keel, or stop	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to walk more than a mile	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to walk several blocks	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to walk one block	Limited=1/Limited a Little=0.5/Not limited=0
	Unable to bath or dress yourself	Limited=1/Limited a Little=0.5/Not limited=0
	Self-rated health	Poor=1/Fair=0.75/Good=0.5/Very Good=0.25/Excellent=0
Diseases	Cancer	Yes=1/No=0
	Myocardial infarction - clinical + ECG + enzymes	Yes=1/No=0
	Myocardial infarction - ECG only	Yes=1/No=0
	Myocardial infarction - clinical + enzymes + non-specific ECG	Yes=1/No=0
	Unstable angina (acute coronary insufficiency)	Yes=1/No=0
	Angina pectoris - classical heberden's	Yes=1/No=0
	Positive exercise test with pain	Yes=1/No=0
	Angina pectoris - plus RS-T and T changes	Yes=1/No=0
	Abnormal coronary angiogram	Yes=1/No=0
	ECG evidence of MI after cardiac surgery	Yes=1/No=0
	Sudden death	Yes=1/No=0
	Valvular heart disease – rheumatic or non-rheumatic or etiology unknown	Yes=1/No=0
	Cardiomyopathy	Yes=1/No=0
	Congenital heart disease	Yes=1/No=0

Thoracic aorta	Yes=1/No=0
Abdominal aorta	Yes=1/No=0
Definite stroke	Yes=1/No=0
Suspected Stroke	Yes=1/No=0
Reported Stroke	Yes=1/No=0
Definite transient ischemic attacks (TIA)	Yes=1/No=0
Suspect TIA	Yes=1/No=0
Reported TIA or cerebral embolism	Yes=1/No=0
Cerebral aneurysm	Yes=1/No=0
Subdural haematoma	Yes=1/No=0
Head injury	Yes=1/No=0
Chest Pain	Yes=1/No=0
Paroxysmal Tachycardia or Atrial Fibrillation or Flutter	Yes=1/No=0
Sick sinus syndrome	Yes=1/No=0
Peripheral artery surgery - includes cerebral, renal, and femoral	Yes=1/No=0
Arthritis	Yes=1/No=0
Cervical spine disease	Yes=1/No=0
Lumbar disc disease	Yes=1/No=0
Asthma, chronic bronchitis, or emphysema	Yes=1/No=0
Pulmonary fibrosis	Yes=1/No=0
Anemia	Yes=1/No=0
Polycythemia	Yes=1/No=0
Gastric ulcer, peptic ulcer, or duodenal	Yes=1/No=0
Renal failure	Yes=1/No=0
Hyperfunction or hypofunction	Yes=1/No=0

Table A2. Simulation results under the MCAR mechanism when groups are well separated.

Performance Measure	MCAR	GBTA			EGBTA		
		Beta_0	Beta_1	Group Proportion (%)	Beta_0	Beta_1	Group Proportion (%)
Bias	Low Attrition Rate	0.042	0.011	0.0077	0.042	0.011	0.00017
	High Attrition Rate	0.045	0.014	0.0303	0.045	0.014	0.00000
Relative Bias	Low Attrition Rate	0.42%	2.25%	0.04%	0.42%	2.25%	0.00%
	High Attrition Rate	0.45%	2.77%	0.15%	0.45%	2.77%	0.00%
MSE	Low Attrition Rate	0.0028	0.00019	0.0003	0.0028	0.00019	0.000
	High Attrition Rate	0.0033	0.00030	0.0017	0.0033	0.00030	0.000

Table A3. Simulation results under the MAR mechanism when groups are well separated.

Performance Measure	MAR	GBTA			EGBTA		
		Beta_0	Beta_1	Group Proportion (%)	Beta_0	Beta_1	Group Proportion (%)
Bias	Low Attrition Rate	0.042	0.011	0.0063	0.042	0.011	0.00024
	High Attrition Rate	0.045	0.014	0.0261	0.045	0.014	0.00000
Relative Bias	Low Attrition Rate	0.42%	2.28%	0.03%	0.42%	2.28%	0.00%
	High Attrition Rate	0.45%	2.86%	0.13%	0.45%	2.86%	0.00%
MSE	Low Attrition Rate	0.0028	0.00019	0.0002	0.0028	0.00019	0.000
	High Attrition Rate	0.0033	0.00032	0.0013	0.0033	0.00032	0.000

Table A4. Simulation results under the MNAR mechanism when groups are well separated.

Performance Measure	MNAR	GBTA			EGBTA		
		Beta_0	Beta_1	Group Proportion (%)	Beta_0	Beta_1	Group Proportion (%)
Bias	Low Attrition Rate	0.042	0.011	0.0086	0.042	0.011	0.00000
	High Attrition Rate	0.047	0.014	0.0340	0.047	0.014	0.00000
Relative Bias	Low Attrition Rate	0.42%	2.24%	0.04%	0.42%	2.24%	0.00%
	High Attrition Rate	0.47%	2.78%	0.17%	0.47%	2.77%	0.00%
MSE	Low Attrition Rate	0.0028	0.00018	0.0003	0.0028	0.00018	0.000
	High Attrition Rate	0.0034	0.00030	0.0020	0.0035	0.00030	0.000

Table A5. Simulation results under the MCAR mechanism when groups are not well separated.

Performance Measure	MCAR	GBTA			EGBTA		
		Beta_0	Beta_1	Group Proportion (%)	Beta_0	Beta_1	Group Proportion (%)
Bias	Low Attrition Rate	0.046	0.012	2.1177	0.042	0.011	0.14928
	High Attrition Rate	0.006	0.021	8.0578	0.045	0.014	0.11086
Relative Bias	Low Attrition Rate	0.46%	2.33%	10.59%	0.42%	2.25%	0.75%
	High Attrition Rate	0.06%	4.20%	40.29%	0.45%	2.79%	0.55%
MSE	Low Attrition Rate	0.0036	0.00021	4.5946	0.0027	0.00019	0.036
	High Attrition Rate	0.0187	0.00067	65.1694	0.0032	0.00030	0.021

Table A6. Simulation results under the MAR mechanism when groups are not well separated.

Performance Measure	MAR	GBTA			EGBTA		
		Beta_0	Beta_1	Group Proportion (%)	Beta_0	Beta_1	Group Proportion (%)
Bias	Low Attrition Rate	0.048	0.012	2.4371	0.042	0.012	0.15637
	High Attrition Rate	0.135	0.022	8.9957	0.045	0.014	0.11197
Relative Bias	Low Attrition Rate	0.48%	2.35%	12.19%	0.42%	2.31%	0.78%
	High Attrition Rate	1.35%	4.42%	44.98%	0.45%	2.86%	0.56%
MSE	Low Attrition Rate	0.0037	0.00022	6.0504	0.0028	0.00020	0.038
	High Attrition Rate	0.0232	0.00078	81.2558	0.0033	0.00033	0.024

Table A7. Simulation results under the MNAR mechanism when groups are not well separated.

Performance Measure	MNAR	GBTA			EGBTA		
		Beta_0	Beta_1	Group Proportion (%)	Beta_0	Beta_1	Group Proportion (%)
Bias	Low Attrition Rate	0.050	0.012	2.5813	0.042	0.011	0.15295
	High Attrition Rate	0.137	0.024	9.3041	0.046	0.014	0.11063
Relative Bias	Low Attrition Rate	0.50%	2.32%	12.91%	0.42%	2.22%	0.76%
	High Attrition Rate	1.37%	4.75%	46.52%	0.46%	2.80%	0.55%
MSE	Low Attrition Rate	0.0039	0.00021	6.7881	0.0027	0.00018	0.037
	High Attrition Rate	0.0242	0.00080	86.8447	0.0034	0.00032	0.021

Table A7. Simulation results under the MNAR mechanism when groups are not well separated.

Performance Measure	MNAR	GBTA			EGBTA		
		Beta_0	Beta_1	Group Proportion (%)	Beta_0	Beta_1	Group Proportion (%)
Bias	Low Attrition Rate	0.050	0.012	2.5813	0.042	0.011	0.15295
	High Attrition Rate	0.137	0.024	9.3041	0.046	0.014	0.11063
Relative Bias	Low Attrition Rate	0.50%	2.32%	12.91%	0.42%	2.22%	0.76%
	High Attrition Rate	1.37%	4.75%	46.52%	0.46%	2.80%	0.55%
MSE	Low Attrition Rate	0.0039	0.00021	6.7881	0.0027	0.00018	0.037
	High Attrition Rate	0.0242	0.00080	86.8447	0.0034	0.00032	0.021