Summer 8-2019

USES OF CONTINUOUS-TIME MARKOV CHAIN TO DESCRIBE LONGITUDINAL PATIENT-REPORTED OUTCOMES FOR SURVIVAL PREDICTION AND DIMENSION REDUCTION

TING-YU CHEN
UTHealth School of Public Health

USES OF CONTINUOUS-TIME MARKOV CHAIN TO DESCRIBE LONGITUDINAL PATIENT-REPORTED OUTCOMES FOR SURVIVAL PREDICTION AND DIMENSION REDUCTION

by

TING-YU CHEN, MS

APPROVED:

WENYAW CHAN, PHD
ACADEMIC ADVISOR

ELAINE SYMANSKI, PHD

MICHAEL D. SWARTZ, PHD

CHARLES S. CLEELAND, PHD

DEAN, THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH

USES OF CONTINUOUS-TIME MARKOV CHAIN TO DESCRIBE LONGITUDINAL PATIENT-REPORTED OUTCOMES FOR SURVIVAL PREDICTION AND DIMENSION REDUCTION

by

TING-YU CHEN
MS, University of Texas School of Public Health, 2013

Presented to the Faculty of The University of Texas

School of Public Health

in Partial Fulfillment

of the requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS
SCHOOL OF PUBLIC HEALTH
Houston, Texas
August, 2019

ACKNOWLEDGEMENTS

I was fortunate receiving lots of help and support during my dissertation research. First, I would like to express my appreciation to my advisor Dr. Wenyaw Chan. His expertise in biostatistics and experience in the statistical application provide tremendous help during my research work. I also want to thank Dr. Elaine Symanski, Dr. Michael Swatrz, Dr. Charles Cleeland, and Dr. Qiuling Shi for their professional comments and guidance on my research. Additionally, I want to give my gratitude to Dr. Zhiling Wu for her assistance on the scientific writing editing for my dissertation.

Besides, I appreciate the support of my wife, Yi-Chien Lee, and my family. I also want to thank all of my friends for their backing during these years when I was pursuing my doctoral degree. Without all of these people's encouragement and contribution, I couldn't complete this journey.

USES OF CONTINUOUS-TIME MARKOV CHAIN TO DESCRIBE LONGITUDINAL PATIENT-REPORTED OUTCOMES FOR SURVIVAL PREDICTION AND DIMENSION REDUCTION

Ting-Yu Chen, MS, PhD
The University of Texas
School of Public Health, 2019

Dissertation Chair: Wenyaw Chan, PhD

A patient-reported outcome (PRO) is a type of outcome reported directly from patients, and it has been widely used in medical research and clinical trials to measure a patient's symptoms, health-related quality of life, physical functioning, and health status. Previous studies have linked PROs to survival outcomes, but most of them only used the PRO information at baseline or at a specific clinical time point [1, 2]. Even though some of these studies collected longitudinal PROs, only few of them evaluated the association between the longitudinal PROs and a survival outcome. One of the major challenges in longitudinal PRO studies is to address the individual heterogeneity in PRO repeated measurements. Due to the fact that PRO is reported directly from patients, and different patients may have different experiences, longitudinal PROs have been often observed with individual heterogeneity, yet current methods [3-5] are not able to account for the individual heterogeneity. Therefore, in this research, we developed three methods using two-state Continuous-Time Markov Chain (CTMC) to summarize longitudinal PRO. The primary summary used is the estimated state transition rates, which serve as summary statistics to depict longitudinal PRO patterns at the individual level. These transition rates can also be incorporated into survival models as predictors or into factor

analysis as observed variables. Specifically, in the first two papers, we developed prognostic models that contained baseline covariates and a longitudinal process in two survival models, Weibull Regression and Cox Proportional Hazard Regression, with different estimation approaches. Simulation studies were conducted to validate the proposed methods, and the proposed models were then applied to two PRO studies separately, with both using repeated PRO measurements during the treatment period in cancer patients to predict the survival outcomes that happened after the treatment. In the third paper, we then integrated two-state CTMC with factor analysis to evaluate the usage of CTMC in PRO symptom clustering. This study showed that CTMC could well summarize the longitudinal PRO information during the treatment period of cancer patients. The underlying construct of patient-reported symptoms had also met our expectations from clinical experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

1. BACKGROUND

1.1 Introduction of Patient-Reported Outcome

1.1.1 Patient-Reported Outcome in Cancer Research

Patient-reported outcome (PRO) is a type of outcome that is measured not from physicians or caregivers, but directly from patients themselves. It is based on the patients' perceptions of the disease and treatment [6, 7]. Generally, PRO includes measures of symptoms (single or multi-dimension), health-related quality of life (HRQoL), physical functioning, and health status [6, 8].

Patient-reported outcomes (PROs) have been widely used in biomedical research, especially in mental health, chronic illness, and oncology [9-12]. In clinical trials or observational studies, PROs have been used in many aspects and have provided different functionalities to researchers. For example, PROs have been used as a tool to determine the eligibility of patients in the enrollment process when the screening outcome can only be reported by patients [7]. Also, PROs have been used to confirm and monitor disease status. Because cancer patients always experience multiple symptoms (e.g., pain, fatigue, and distress) during their treatment, PROs can play a role as a surrogate to monitor patients' reactions to the treatment and provide responsive symptom information to clinicians [13]. Additionally, to explain low patient compliance rates in clinical trials, PRO measures are often used [7]. For example, patients who have severe symptoms may drop out from clinical trials early. PRO is also widely used as a study endpoint because it can provide a unique perspective on treatment effectiveness [14-16]. Studies have used PROs as primary endpoints or exploratory endpoints in clinical trials when the research questions can only be measured by PROs. When the treatment has a negative implication for other aspects of life (e.g., side effects), PROs can also assess the

impact of the treatment on the patient's quality of life. Furthermore, Basch et al. [17] showed that a patient-reported symptom monitoring group had better survival compared to a usual care group while controlling other demographic and clinical factors. Additionally, the development of a new biosimilar compound may need PROs to distinguish the benefit from the standard drugs when the biological benefits are comparable. Therefore, it is believed that PROs will be as important as other types of outcomes, such as clinical outcomes, physiological outcomes, and care-giver reported outcomes. In fact, in the past decades, PROs have been extensively applied to observational studies and clinical trials. The European Medicine Agency (EMA) published a guideline for the pharmaceutical industry, in which PRO is suggested to be included in oncology clinical trials to assist benefit-risk assessment and therapeutic claims [6, 18]. Similarly, the US Food and Drug Administration (FDA) released a guideline that recommended PROs for clinical trials and outlined the properties and components of PRO instruments [7].

1.1.2 Patient-Reported Outcome Instruments

To better understand the difference between symptoms and HRQoL and elements of a PRO instrument, this section will introduce properties of a PRO instrument and some commonly-used PRO instruments in oncology.

Patient-reported Outcome Symptoms and Health-Related Quality of Life

PRO symptoms and HRQoL are often studied and discussed together, yet they are not the same. A symptom is a one-dimensional property related to a disease or treatment status that is directly reported from patients. On the other hand, HRQoL is a multi-dimensional measure that

consists of several physical, psychological, social, economical factors, disease or treatment symptoms, and cultural set-up [19].

Properties of PRO Instruments

PRO instruments are the tools used to measure PROs, e.g., symptoms or HRQoL related to a disease or treatment. According to Deshpande et al. [19], a good PRO instrument should possess the following properties: First, it should be specific to the concept being measured [20]. Depending on the purpose of the instrument, a PRO instrument usually only measures one general concept of a disease or treatment. Second, a good PRO instrument should be based on an end-point model, a hierarchical model that considers all types of endpoints, non-PRO assessment and PRO assessment, to meet the requirements of a clinical trial's objective, study design, and data analysis plan [20]. For example, a clinical trial following an end-point model may contain three types of endpoints that can be measured from a bio-chemical exam (non-PRO assessment), physical exam (non-PRO assessment), and symptoms of the disease (PRO assessment). Third, a good PRO instrument should have conceptual equivalence, which means the instrument should measure the same concept equivalently in different languages and cultures. Fourth, a good PRO instrument should consist of a conceptual framework, an "item-domain-concept" structure that defines how concepts should be measured by the instrument [20]. For example, a general QoL concept can be defined by several domains (physical, emotional, and social), and each domain can be measured by several items. Fifth, a good PRO instrument should have easy and specific measurement properties and contain an optimum number of items. The scale system for the measurement and the number of total items in an instrument should be easy to understand and complete for patients. Last, a good PRO instrument should maintain

patient confidentiality. Patient confidentiality is important for all kinds of studies. So, a good instrument should only collect the information that is necessary for each study objective.

Common Oncology PRO Instruments

A number of PRO instruments have been designed for oncology studies. PRO instruments that are commonly used in cancer clinical trials are as follows:

EORTC-QLQ-C30

Developed by the European Organization for Research and Treatment of Cancer (EORTC), the EORTC quality of life core questionnaire (QLQ-C30) consists of 30 items [21, 22]. It is a general instrument that can be applied to all types of cancer patients through 5 functional scales (physical, role, emotional, cognitive, and social), 9 single items, and two items on global health status and QoL. It uses a 4-point numerical rating scale (not at all, a little, quite a bit, very much) in most questions and a 7-point numerical scale for global health status/QoL items. The raw score in EORTC-QLQ-C30 can be transformed to a 0-100 scaling by following its scoring procedure. A higher functional scale or QoL represents a healthy level of functioning or higher QoL; while a higher symptom item represents a higher level of symptomatology/problems. This instrument also has several modules to supplement the core EORTC-QLQ-C30. For example, QLQ-BR-23 contains 23 additional items measuring disease symptoms, side effects of a treatment, body image, sexual functioning, and future perspectives for breast cancer patients. A QLQ-H&N-35 is a module for head and neck cancer patients that measures symptoms and side effects of a treatment, social function, and body image/sexuality through 35 additional questions. The scoring approaches for these modules are identical with that for QLQ-C30. Thus, these modules can be easily applied to data analyses and reporting in

cancer studies. The interpretation of scoring can be done by reporting the raw score in each functional scale or single symptom/QoL item, by comparing the difference of score (e.g., Cohen's effect size) between groups, or by comparing scores at different time points.

EQ-5D-5L

The EuroQol Group developed EQ-5D-5L and released it in 2009 [23, 24]. It is a modified version based on the previous EQ-5D-3L, but EQ-5D-5L has improved the instrument's sensitivity and reduced the ceiling effects. This instrument contains 5 domains (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) using a 5-level scale (no problems, slight problems, moderate problems, severe problems, and extreme problems) and a visual analogue scale (EQ VAS) with a 0-100 scaling. Specifically, the EQ VAS is a scale that asks the patient to mark their health condition of that day on a ruler graph, ranging from "worst possible" to "best possible" health. The report for EQ-5D-5L may contain the distribution (frequency) of each domain, an overall index calculated from the domains (range in -0.1 to 1), and the EQ VAS.

PRO-CTCAE

The Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) [25, 26] was developed by National Cancer Institute (NCI) to capture the symptomatic adverse events (e.g., side effects of a treatment) in cancer patients enrolled in clinical trials. It is recommended to be used with a Common Terminology Criteria for Adverse Events (CTCAE) score graded by physicians. PRO-CTCAE includes an item library that consists of 124 items for 78 symptomatic toxicities. Selection of items from the library is usually based on previous studies or previous pre-clinical data. It uses a 0-4 scale for each item,

and each symptom is measured through three attributes: frequency, severity, and/or interference. The analysis of PRO-CTCAE depends on the design of each study, but every toxicity item is interpreted independently.

FACT-G

Functional Assessment of Chronic Illness Therapy (FACIT) developed Functional Assessment of Cancer Therapy – General (FACT-G) that collects core quality of life items from patients [27, 28]. It consists of 27 items in 4 domains (physical, social/family, emotional, and functional well-being), and it is usually applied to cancer and other chronic diseases (e.g., HIV/AIDS). It uses a 5-point scaling and contains some reversed items. A higher score in FACT-G (range in 0-108) means a better quality of life. FACIT also provides modules for various forms of cancer, which include an additional section with disease specific questions. For example, FACT-Hep is an instrument for patients with hepatobiliary cancer (liver, bile duct, and pancreas), and FACT-O is one for patients with ovarian cancer. In these modules, a Trial Outcome Index (TOI), a summary of 4 domains in FACT-G, and the subscale from a specific disease is usually reported to reflect the overall physical and functional condition of a patient.

SF-36

RAND Health Care developed Short Form Health Survey (SF-36) that contains 36 items to measure patients' generic quality of life [29, 30]. SF-36 is part of the Medical Outcome Study (MOS), a multi-site study that explaines variations in patient outcomes. SF-36 contains 8 health concepts: physical functioning, bodily pain, role limitations due to physical health problems, role limitations due to personal or emotional problems, emotional well-being, social functioning, energy/fatigue, and general health perceptions. In this instrument, it uses a scaling system

ranging from a 2-point scale to a 6-point scale. It interprets the results of each health concept (range in 0-100), and a higher score means a more favorable health state.

MDASI

The MD Anderson Symptom Inventory (MDASI) is a multi-symptom PRO instrument developed by The University of Texas MD Anderson Cancer Center [31, 32]. MDASI-Core contains 13 core symptom items and 6 interference items using a 0-10 numerical rating scale, where 0 represents no symptom present, and 10 represents the symptom that is as bad as a patient can imagine. The simple feature and design make this instrument easy to understand, and it can be completed in 3-5 minutes. MDASI also comprises multiple modules that are designed for different cancer diseases and treatment-specific purposes. For example, MDASI-BT is a module for brain tumor patients that includes items in MDASI-Core and additional items specifically related to brain tumor. The interpretation of MDASI depends on the study objective. Both single-item scoring (score in 0-10) and composite scoring (e.g., mean of most 5 severe symptom items) are widely used in clinical studies. In this dissertation, given the data availability, MDASI will be the only PRO instrument used to measure cancer patients' symptom burden during the treatment period.

Most PRO instrument manuals provide a recommendation for PRO result presentation or instructions of PRO instrument administration. Some instruments can report the result in single symptom item, while others can only report the subscale of each domain or a total index after designed transformation. Besides, the frequency of administering a PRO instrument is also important, and it depends on the recall period. For example, the recall period for EQ-5D-5L is the day when patients complete the survey. MDASI, however, has a recall period for the last 24 hours, and EORTC-C30, FACT-G, PRO-CTCAE have a recall periods for the past week (7

days). Some instruments suggest a weekly collection during the study; other depend on the need and design of the study. The more frequent a PRO instrument is distributed, the more costly the study will become. In some studies [33], PROs are collected only 2-3 times across the study period, and PRO results are compared in a cross-sectional manner. For example, a QoL score at baseline (before treatment) can be compared with the QoL score after the treatment. However, more researchers start to collect longitudinal PROs (more than 3 times of collection) in their studies [34, 35] because longitudinal PRO can better explain patients' experience on the cause of disease and treatment. The analysis of longitudinal PRO results will be discussed in the following section.

1.1.3 Longitudinal PRO Analysis

Longitudinal PROs can provide more symptom and HRQoL information realted to the disease, drugs, and operations involved. For example, a comprehensive deisgn of longitudinal PRO collections records PRO at baseline, treatment cycles (weekly or bi-weekly), and post-treatment follow-up (bi-weekly, monthly, or every 6 months) [36]. The PRO instruments; therefore, can measure symptom change over the study period, providing detailed information on the difference before and after the treatment; and capturing the variation of symptoms at a meaningful clinical time point.

The trajectory of longitudinal PRO may not always be linear. In most of the observational studies on cancer with chemotherapy [37, 38], a weekly measured PRO tends to show a non-linear pattern (waves) during the weeks when patients receive the chemotherapy. This type of data can indicate whether a symptom is changed by the influence of the disease or by the effectiveness/burden of the treatment.

Longitudinal PROs as Outcomes

There are many ways to analyze longitudinal PROs; however, most of the methods consider longitudinal PROs as outcome variables. For example, Fairclough [39] described the common considerations for analyzing a longitudinal PRO as an outcome in her book; and provided examples and programing codes with detailed instructions. Below are some examples of longitudinal PROs analyzed as outcomes:

Lin et al. [40] used a Linear Mixed-effect Model (LMM) to analyze longitudinal HRQoL data from patients who had lumber spine surgery in Taiwan. A series of PRO instruments were distributed (e.g., Taiwanese version of World Health Organization Quality of Life-BREF and Numerical Rating Scale for leg and back pain) to the patients 1 week before their lumber spine surgery, as well as during the first, sixth, and twelfth month after the surgery. Two domains in the HRQoL, physical health and social health, were fitted in the LMM with random intercept and random slope at a subject level to account for the variation within subjects. Lin et al. found neurological functions, sleep quality, and depressive symptoms were the key factors affecting the QoL. This is a typical example that treats longitudinal PRO as a repeated outcome, and how it is analyzed directly in a statistical model (LMM). In fact, because of the simplicity of the modeling and the clarity of interpretation and model assumption, LMM has been widely used in PRO research.

Another approach to analyzing longitudinal PROs is to transform longitudinal information into a one-dimensional variable and analyze it using regular models depending on the type of the transformed variable. For example, Shi et al. [41] showed an example using Group-Based Trajectory Modeling (GBTM) to transform longitudinal PROs into binary or ordinal variables, and then these variables were treated as outcomes in an (ordinal) logistic

regression. The GBTM is a model often used in psychology to identify clusters of subjects that follow a particular pattern (trajectory) related to the outcome over time. Shi et al. applied the 5 most severe symptom items (taking mean of 5 items) to GBTM and determined 2 or 4 groups of patients that followed the distinguished trajectories. Once the clustering of subjects was identified, the predictors to this outcome were examined by the Generalized Linear Model (GLM). Comparing with Lin et al.'s approach, this approach is based on the assumption that subjects have different patterns (groups) of longitudinal PROs, while Lin et al. [40] assumed their longitudinal PROs followed a general linear trend.

PROs as Covariates

Other studies treat PROs as covariates and investigate the association between PROs and one type of outcome. For example, Quinten et al. [42] suggested the patients-reported symptoms (measured from EORTC-QLQ-C30) plus clinical rating (CTCAE) could better predict overall survival. They used data from 14 clinical trials that compared models with or without PROs in the Cox Proportional Hazard Model using Harrell concordance index (C-index). It turned out that models with both patent-reported symptoms and clinical-rated variables showed a higher C-index value.

Armstrong et al. [43] also showed how PROs could be treated as covariates to predict the overall survival in patients with brain tumors. They linked the PROs (using both EORTC-QLQ-C30/BN20 and MDASI-BT) to examine the progression-free survival (PFS) and overall survival (OS) in a phase III clinical trial. They found that PROs were associated with the risk of OS and PFS. Besides, higher symptom burden on patients showed a higher risk of OS and PFS, while a higher QoL score showed a lower risk of OS and PFS.

However, most studies only used PROs at a specific time point as covariates in their model. None has applied longitudinal PROs as covariates directly in models to predict outcomes. Therefore, in this dissertation, we want to develop methods that incorporate longitudinal PROs as covariates in survival models.

1.2 Models for Longitudinal covariates

General Approaches to Repeated Measurements

Longitudinal data analysis has been developed for years when the structure of the data includes repeated measurements. The mixed-effects model and marginal model are the most popular methods applied to longitudinal data. Among these methods, the mixed-effects model has used quite widely because it models characteristics on subject specific interpretation. It can incorporate fixed-effects, random effects, and time-dependent covariates in the model while accounting for the correlation from repeated outcomes (continuous or categorical type).

In contrast, survival analysis (also known as "time-to-event" type of data) also collects information over time, but it models the survival outcome as a function of the covariates. Besides, time-dependent covariates can be included in the survival model when the covariates are assumed to be measured without error. For more information about general longitudinal analysis and survival analysis, please read Diggle et al. [44] and Klein and Moeschberger's [45] books.

Joint Models for Longitudinal Outcomes

Joint modeling is often used when two or more outcomes (/processes) are correlated or associated. The main advantage of this modeling technique is its ability to reduce the bias of

estimation [46]. When one of the outcomes is a time-to-event variable, modeling a longitudinal process and a time-to-event outcome can improve the inference for both outcomes [47, 48] because the relationship of both outcomes is considered. Several methods modeling a time-to-event outcome and longitudinal outcomes are discussed in Hickey's research [47] including the mechanism of joint modeling. One popular type of joint model for longitudinal data and the time-to-event outcome is expressing survival outcomes in hazard function (e.g., Cox Proportional Hazard (Cox PH) model) [49]. The difference between groups (e.g., a binary variable) or a unit increment in a continuous variable can be easily interpreted as a risk increase or decrease to the survival event. For example, Tsiatis and Davidian [5] proposed a joint model modeling longitudinal data (as covariates) as a function of time in the Cox PH model; and estimating parameters from the partial-likelihood.

Another class of longitudinal data in joint modeling deals with situations when the outcome is not time-to-event. For example, Wang et al. [50] proposed a joint model that is a combination of a longitudinal model and a generalized linear model (GLM). The first submodel is a random-effect model for repeated measurements, and the second one (considered as the primary model) is a GLM that takes random effects from the first submodel as covariates. The limitation of this model is that the normality assumption for longitudinal covariates (random effects in the primary model) is not applicable to most real world situations. Wang and Huang [51] and Li [52] modified the method and proposed similar methods that did not require the distribution assumption for the random effects, but through the longitudinal process. In summary, these models [50-52] manage the longitudinal process as covariates and treat the process as a continuous variable in the primary model (e.g., a GLM). It is true that one or more

features from the longitudinal submodel may also be evaluated in the primary models; however, not all of these features will represent individual behaviors.

Functional Data Analysis

Another category of analyzing longitudinal covariates is through dimension reduction. When the longitudinal data include multiple variables, an approach to analyze this type of data is using Functional Data Analysis (FDA) while treating longitudinal data as smoothed curves, surfaces, or even hypersurfaces [53-56]. The concept here is to extract the proper information from the larger dimensions and compress it into smaller features. Functional Principal Component Analysis (FPCA) is one of the FDA techniques that select meaningful components from the multiple longitudinal covariates. FPCA also needs to select suitable basis functions (usually expects an orthogonal relationship amongs the features) and to determine the number of eigenfunctions to be chosen. For more information, Ramsy [55] provides a comprehensive introduction and detailed statistical methodology in FDA.

1.3 Markov Chain as Covariate Processes

Markov Chain models have been used to study the progression process for neurological diseases such as Alzheimer's disease [57, 58] and Parkinson's disease [59, 60] because the nature of these diseases' process can be assumed to follow the Markov Property. The Markov Property is that for a collection of events, the state of a future event only depends on the current state, not the past state. Therefore, the Markov Chain, a collection of variables over time that follows the Markov Property, is usually considered as an outcome in the study. Karisson [61] used the first-order Markov model to describe the change of sleep level (lighter sleep and deeper sleep) from

patients who had insomnia with comparison of treatment effect. However, not many studies in the literature modeled the longitudinal process as a covariate and integrated it in the outcome of interest.

Ho [62] modeled a longitudinal DNA-damage process as a Discrete Time Markov Chain (DTMC) to predict the occurence of lung cancer in a case-control study. This method used the information that is repeatedly measured from genetic testing and modeled it as DTMC with transition probabilities estimated, and then applied these traisition probabilities in a logistic regression to esitmate the odds of the lung cancer. Later on, Rubin et al. [63] extended this concept to the Continuous Time Markov Chain (CTMC) and modeled longitudinal intracranial pressure (ICP) with CTMC as covariates in the logistic regression to predict the status of patients with traumatic brain injuries. In constrast to Ho and Rubin's research, our proposed models in this dissertation will model longitudinal PROs as CTMC and then use the transition rates as covariates in a survival model framework. To the best of our knowledge, no research published has ever applied CTMC to the longitudinal PRO data analysis.

1.4 Public Health Significance

The patient-reported outcome (PRO) is a tool that has been widely applied to biomedical research and clinical trials, and it is expected to be used more intensively in the future. Hassett et al. [36] outlined the high-priority topics that were discussed during the 2012 American Society of Clinical Oncology (ASCO) Quality Care Symposium, and developing patient-centered quality measures is one of them. Symptom measurement in cancer patients also needs to be conducted more comprehensively before, during, and after the treatment. Therefore, the methodology that can utilize the information of longitudinal PROs in cancer studies is important; especially when

no methodology has linked longitudinal PROs with survival type of outcome in single statistical technique. The proposed methods in this dissertation fill the gap in current PRO cancer research to fully link the patients-reported symptom burden experience in the treatment period with the overall survival.

Furthermore, our prognostic models have significant impacts on clinical research and real world clinical practices especially when the concept of personalized medicine is a common goal to achieve. In particular, our prognostic models can add value to treatment decisions, quantify the disease status, identify the patient's condition, and predict clinical outcomes. Basch et al. [17, 64] reported that routine symptom monitoring could prolong the survial time when compared to a usual care group without PRO monitoring because caregivers could respond to the patients when PRO symptom burden increment was detected. The caregivers could then provide symptom management conselling, supportive medications, chemotherapy dose modifications, and referrals to doctors. As a result, a methodology that can analyze PRO monitoring data (a longitudinal PRO structure) is essential and is useful when the primary outcome of interest is survial. We expect the proposed methodologies to provide clinicians meaningful and interpretable results on the association between longitudinal PROs and survival outcome. It is also believed that these methods could improve quality care of cancer patients in the future.

1.5 Specific Aims

The specific aims of this research study are described as follows:

Aim 1. To develop a fully-parametric time-to-event model using a two-state continuous-time Markov Chain (CTMC) as predictors and other clinical/demographic risk factors as covariates.

In this aim, a joint model of full-parametric survival model that follows a Weibull distribution and a two-state CTMC was developed when the primary study outcome is the "time" to an event (e.g., overall survival time). This approach models a longitudinal patient-reported outcome (PRO) measured during the treatment period as a CTMC to predict a survival outcome (survival time) that occurred after the treatment period. The focus of this aim is to evaluate the prognostic value of a single longitudinal PRO score (a single item score) dichotomized as none/mild or moderate/severe to overall survival time. Other covariates, such as demographic and clinical variables, can be adjusted in the survival submodel as well as in the CTMC submodel. The sets of covariates adjusted in each submodel can contain different components of variables. This method was applied to an observational study of non-small cell lung cancer (NSCLC) patients (named as NSCLC study), recruited from 2004-2009 at The University of Texas MD Anderson Cancer Center.

Aim 2. To develop a semi-parametric time-to-event model using a two-state continuous-time Markov Chain (CTMC) as predictors and other clinical/demographic risk factors as covariates.

In this aim, a joint approach of semi-parametric survival model and a two-state CTMC was developed when the primary study outcome is "risk" of the event (e.g., the risk of death in overall survival or risk of a disease progression). A longitudinal PRO measured during treatment period was modeled as a CTMC to predict a survival risk using Cox proportional hazard model when the survival event occurs after the treatment period. This aim targets the association between a longitudinal PRO score and risk of a survival event. Similarly, as that in Aim 1, the sets of covariates included in the survival submodel and CTMC submodel can consist of

different variables. This method was applied to an observational study of head and neck (HN) cancer patients (named as HN study), recruited from 2006-2007 at The University of Texas MD Anderson Cancer Center.

Aim 3. To conduct factor analysis of multiple two-state continuous-time Markov Chains (CTMCs) on individual longitudinal patient-reported outcomes.

Each longitudinal PRO was modeled as two-state CTMC, and their transition rates were utilized for factor analysis that identified the latent factor on PRO items. Unlike most studies that applied factor analysis to PRO items at a single occasion, we summarized each longitudinal PRO as CTMC by its transition rates and applied to factor analysis to reduce dimensions of longitudinal PROs. This method was applied to data from HN studies. The factors of symptoms provided hidden information on the sources of symptom burden and classification of symptom items such as treatment-related or disease-induced items.

2. Journal Article 1

A Joint Model of Time to Event and Continuous-Time Markov Chain

Statistics in Medicine

Keywords: Continuous-Time Markov Chain (CTMC), longitudinal data analysis, patient-reported outcome (PRO), survival analysis, Weibull regression

2.1 Background

In biomedical research, linking a patient's longitudinal measurements to time-to-event types of outcome is essential since patients' repeated measurements may contain more information on impact from disease and/or treatment. Available approaches so far include a proportional hazard model with time-dependent covariates [1-3] and a joint model that comprises a longitudinal model and a survival model [4, 5]. In these two approaches, the former evaluates time-dependent covariate effects by truncating time-to-event to pieces based on the time points when covariates change, but it does not take into account any information of the longitudinal pattern of covariates. The latter approach, on the other hand, models a longitudinal covariate as an overall trend and specifies the correlated nature of both outcomes, and the longitudinal submodel may include another set of covariates that are correlated to the longitudinal outcome resulting in a more complex model structure. However, when individual heterogeneity exists in longitudinal measurements, neither of these two approaches could adequately fit the data. Under such circumstance, a more flexible method, such as Markov Chain [6, 7] at an individual patient level, may be more suitable to explain the distinct trajectories at individuals. Therefore, in this study, we propose a joint Continuous-Time Markov Chain (CTMC) and Weibull model (also called Weibull Regression model [2, 3]) that treats the transition rates of CTMC as covariates in

Weibull Model to predict a survival outcome that happens after the longitudinal process, while both submodels permit other baseline information as covariates.

Joint Models for Longitudinal Covariates and Time-to-event Outcome

Joint modeling is regularly applied when two or more outcomes or processes are associated, and their relationship with other covariates needs to be simultaneously considered. The major advantage of this modeling technique is its ability to reduce estimation bias [8]. Moreover, modeling a longitudinal process and a time-to-event outcome can improve the inference for both outcomes [5, 9] because this model is able to consider the relationship of both outcomes. In the literature, several methods modeling a time-to-event outcome and longitudinal outcomes including the mechanism of the joint modeling have been discussed in Hickey's research [9]. Tsiatis and Davidian [4] also proposed a joint model modeling longitudinal data (as covariates) as a function of time in the Cox PH model and estimated parameters from the partial-likelihood. Nevertheless, these models do not address change of the longitudinal process over time at the individual level.

Markov Chain in Medical Research

Markov Chain models have been used to study the progression process for neurological diseases such as Alzheimer's disease (AD) [10, 11] and Parkinson's disease [12, 13] because the nature of these diseases' process can be appropriately assumed to follow the Markov Property, in which the states of future events only depend on the current state, not the past state. Therefore, Markov Chain, a collection of variables over time that follows the Markov Property, is usually considered as a longitudinal outcome in medical studies. For example, the first-order Markov model has been used to describe the change of sleep level (lighter sleep vs. deeper sleep) from patients who had insomnia, and the treatment showed an improved time in sleep [14]. Wu [15]

proposed a joint model that consists of two three-state CTMCs, and applied it to two neuropsychological markers in AD to understand the transition probability among disease severity levels. Yu et al. [16] illustrated a similar joint model with two two-state CTMCs in a stress study, and explored the association between mothers' stress level and children's illness status. However, not many Markov Chain studies in the literature modeled the longitudinal process as a covariate and integrated it to the outcome of interest.

One of the few Markov Chain studies that examined the longitudinal process as a covariate was found in a dissertation work. Ho [17] modeled the longitudinal DNA-damage process as a Discrete Time Markov Chain (DTMC) to predict the occurrence of lung cancer in a case-control study. This method used the information that is repeatedly measured from genetic testing and modeled them as a DTMC with transition probabilities estimated, and the estimated transition probabilities are then applied into a logistic regression to estimate the odds of developing lung cancer. This concept was later extended to the CTMC that modeled longitudinal intracranial pressure (ICP) as a CTMC with the transition rates as covariates in a logistic regression to predict the status of patients with traumatic brain injuries [18]. However, to our knowledge, no studies have modeled longitudinal process as covariates through Markov Chain and linked to the time-to-event outcome.

Therefore, in this paper, to provide a solution to the condition when individual heterogeneity was observed in a longitudinal process with a time-to-event outcome, we propose a joint model of a Markov Chain and a Weibull regression that describes a longitudinal process as CTMC and treats the generated transition rates as covariates in the Weibull regression model to predict a survival event occurring after the longitudinal process. We will apply this model to a completed patient-reported outcome (PRO) study using the MD Anderson Symptom Inventory

(MDASI) [19] for lung cancer patients who received chemotherapy at the MD Anderson Cancer Center. In this application, the repeated patient-reported symptom severity during the chemotherapy periods are handled as a CTMC and are used to evaluate its association with post-treatment overall survival in the Weibull regression model.

The remainder of this paper is organized as follows: In Section 2, we describe the proposed joint model including the likelihood function and its estimation procedures. Section 3 provides validation of the estimating procedure through simulation. Application of the proposed model to the real world PRO research is demonstrated in Section 4. In the last section, we discuss the advantages and limitations of the proposed model.

2.2 Methods

We develop a joint model which consists of a combination of a time-to-event model and a two-state CTMC. Through this joint model, we are able to consider simultaneously the effect of baseline covariates on the time-to-event outcome as well as on the longitudinal covariate process that is dependent on other covariates and is associated with the time-to-event outcome.

2.2.1 The Joint Weibull Model and Two-state CTMC Model with Covariates

In the CTMC submodel, let T be a time-to-event outcome, and $Z(t)$ be a homogeneous CTMC with a state space $S = \{1, 2\}$ characterized by the transition rates λ and μ. The transition rates matrix can then be defined as

$$Q = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

where λ represents the transition rate from State 1 to State 2, while μ represents the transition rate from State 2 to State 1.

In the Weibull submodel, let X be a vector of covariates directly related to the time-to-event outcome T. Also, vector $\boldsymbol{v}$ is the vector of covariates related to T through the transition rate λ, and vector $\boldsymbol{w}$ is the vector of covariates related to T through the transition rate μ. The survival function of T can then be expressed as

$$S(t) = exp(-(\boldsymbol{X'\beta} + b_1\lambda + b_2\mu)t^{\gamma}) = exp\left(-\left(\boldsymbol{X'\beta} + b_1 e^{\boldsymbol{v'g}} + b_2 e^{\boldsymbol{w'h}}\right)t^{\gamma}\right)$$

where $\boldsymbol{\beta} = \left(\beta_0, \beta_1, \cdots, \beta_p\right)^T$ is the effect of the baseline covariates vector $\boldsymbol{X}$ on the outcome T. γ is the Weibull shape parameter. $b_1$and b_2 are the overall multiplicative effects of the transition rates λ and μ, respectively, on the survival outcome. Because of the stationary assumption of CTMC, the probability of moving from one state to another in a unit time t is the same at any time point. Note that, because λ and μ are both positive, we rewrite the transition rates as an exponential form and allow the regression expression on each form's exponent. Therefore, $\boldsymbol{g}$ is the effect of the covariate vector $\boldsymbol{v}$ on the transition rate λ, and $\boldsymbol{h}$ is the effect of the covariate vector $\boldsymbol{w}$ on the transition rate μ.

2.2.1.1 Interpretation of the transition rates

The association of the covariates λ and μ with the outcome T can be interpreted as follows: For each unit increment in the transition rate λ, the outcome T is expected to change by a multiplicity of e^{b_1}, while holding other covariates constant. Similarly, for each unit increment in the transition rate μ, the outcome T is expected to change by a multiplicity of e^{b_2}, while other covariates remain as constants. The interpretation of other parameters including $\beta, v,$ and w is provided in Section 2.4.3.2 and Section 2.4.3.3.

2.2.2 The Likelihood Function of the Proposed Joint Model

The joint likelihood function for the longitudinal covariate process and the Weibull outcome model for n subjects can be written as

$$L(\boldsymbol{h}, \boldsymbol{g}, \boldsymbol{b}, \boldsymbol{\beta}, \gamma) =$$

$$\prod_{i=1}^{n} \left\{ \begin{array}{c} \prod_{u=1}^{K} [P(Z(0) = u)]^{I_u(Z(0))} \\ \prod_{k=1}^{J_i} \prod_{l=1}^{K} \prod_{m=1}^{K} \left[P_{Z(s_{i,k-1})=l,\ Z(s_{i,k})=m}(s_{i,k} - s_{i,k-1}) \right]^{I_l(Z(s_{i,k-1}))I_m(Z(s_{i,k}))} \\ \left[e^{-(e^{-\boldsymbol{X}'\boldsymbol{\beta}+\boldsymbol{q}_i'\boldsymbol{b}})t^{\gamma}} \right]^{\delta_i} \left[\gamma (e^{-\boldsymbol{X}'\boldsymbol{\beta}+\boldsymbol{q}_i'\boldsymbol{b}}) t^{\gamma-1} e^{-(e^{-\boldsymbol{X}'\boldsymbol{\beta}+\boldsymbol{q}_i'\boldsymbol{b}})t^{\gamma}} \right]^{1-\delta_i} \end{array} \right\}$$

$Z(s)$ is a family of random variables that describes the state of a continuous process at time s with state space $S = \{1, 2\}$. $P_{ab}()$ is the probability of transition from a to b for a fixed time interval, which can be explicitly expressed in transition rates λ and μ. Also, $\boldsymbol{q}_i = (\lambda_i, \mu_i) = (e^{\boldsymbol{w}'\boldsymbol{h}}, e^{\boldsymbol{v}'\boldsymbol{g}})$ is the vector of the transition rates. Note that the set of covariates $(\boldsymbol{v}, \boldsymbol{w})$ can be identical or different from the vector $\boldsymbol{X}$. Besides, K is the state that specifies in state space S, and J_i is the number of transitions measured in the CTMC process for subject i, and δ_i is the indicator of occurrence of a survival event.

2.2.3 Estimation and Initial Values for Iterative Procedures

The maximum likelihood method was used to maximize the formula in Section 2.2.2 with respect to $\boldsymbol{\beta}$, $\boldsymbol{b}$, $\boldsymbol{h}$, and $\boldsymbol{g}$ in a one-stage procedure. This is a nonlinear optimization problem where there is no closed-form solution for parameter estimators. Newton-Raphson optimization was used in the study to maximize the likelihood and obtain the estimators of interest. It did not require any calculation of derivatives for each parameter, so it was adequate for problems with non-smooth functions and applicable to most statistical software. For analyses and simulation in the study, PROC NLMIXED in SAS 9.4 (SAS Institute, Cary NC) was applied.

To speed up the computational time and increase the likelihood of reaching a global maximum, we selected the initial values for numerical estimation by implementing the following two-step algorithm:

1. A usual two-state CTMC model with covariate sets $\boldsymbol{v}$ and $\boldsymbol{w}$ was fit to find the parameter value sets $\boldsymbol{g}$ and $\boldsymbol{h}$. At this step, time-to-event outcome T and baseline characteristics X were ignored.
2. Following Step 1, the transition rates λ and μ were calculated with covariates $\boldsymbol{v}$ and $\boldsymbol{w}$ and parameter values $\boldsymbol{g}$ and $\boldsymbol{h}$. The transition rates were then incorporated to the Weibull Model with survival outcome T to obtain the estimators for $\boldsymbol{\beta}$, $\boldsymbol{b}$, and γ. Once this model was fitted, the initial values for all parameters were ready for the joint model.

2.3 Simulation and Validation of the estimation procedure

2.3.1 Description of the Simulation Study

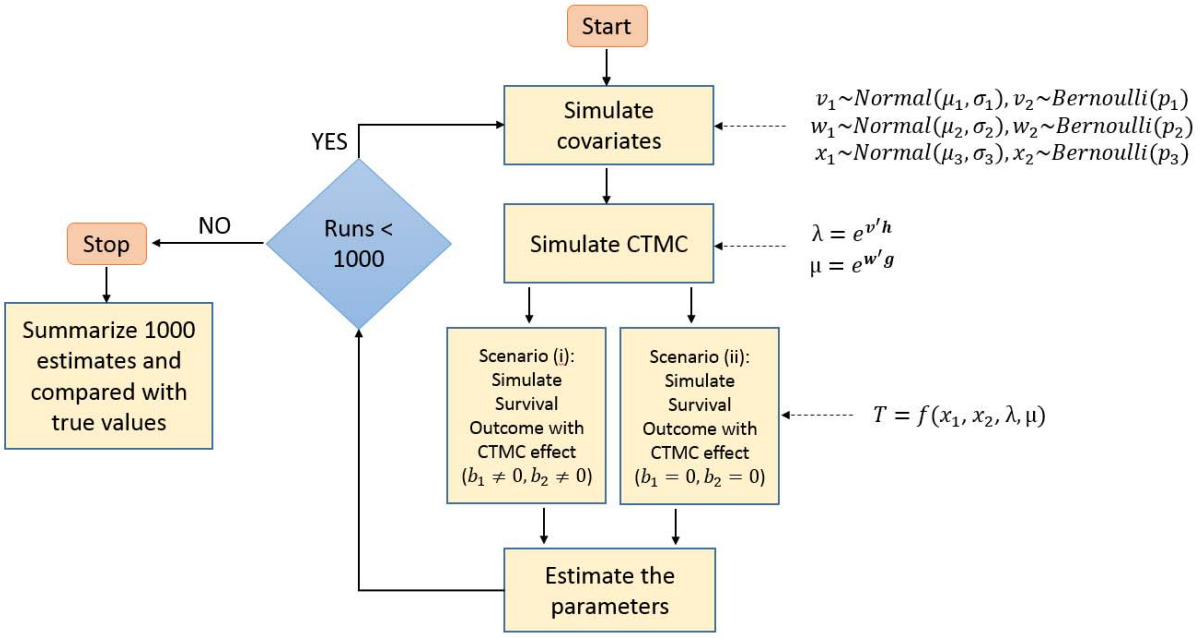
To validate the estimation procedure of the joint models, we also conducted simulations to examine our proposed model. The "true" parameters in this simulation study were obtained from the application of the proposed model to PRO studies. The data simulation procedure was a two-step process (see Figure 2.1). The covariates were first simulated following each variable's specified distribution. The transition rates and the Weibull distribution dependent on these rates were then simulated accordingly. A complete description and a flowchart of the simulation mechanism are as follows:

2.3.1.1 Simulation of covariates

The goal of this study is to develop prognostic models to predict survival in patients with cancer. Therefore, the type of covariates and the distributions of chosen for the simulation study

should mimic those in PRO studies. For simplicity, we simulated a few covariates that might be encountered in real-world data.

Figure 2.1. Simulation procedure of Markov Chain and time-to-event outcome.



1. Covariates in Markov Chain: Two covariates, one continuous (v_1) and one binary (v_2) were simulated as predictors directly associated with the transition rate λ. v_1 followed a normal distribution, and v_2 followed a Bernoulli distribution. Similarly, one continuous variable (w_1) and one binary variable (w_2) were simulated for the transition rate μ.
2. Baseline covariates in the survival model: In the Weibull model, we simulated two covariates (one continuous (x_1) and one binary (x_2)) to explain the effect of baseline characteristics on the survival outcome. x_1 followed a normal distribution, and x_2 followed a Bernoulli distribution.

2.3.1.2 Simulation of Markov chain

The simulation for a two-state CTMC required fixed transition rates λ and μ for each subject. The transition rates were calculated based on the covariate sets $\boldsymbol{v} = (v_1, v_2)$ and $\boldsymbol{w} =$

(w_1, w_2) generated from the procedures presented in Section 2.3.1.1. Once λ and μ were obtained, the CTMC was simulated, and its states were observed at prespecified time points for each subject in the study [20].

2.3.1.3 Simulation of time-to-event outcome

The survival outcome was simulated based on the covariates x_1, x_2, λ, and μ obtained from the previous steps. The Weibull distribution type of survival time T can be expressed as [21]

$$T = \alpha^{-1}[h_0 \times \exp(-\boldsymbol{\beta}'\boldsymbol{X})]^{1/\gamma}$$

where h_0 is a pre-specified baseline hazard; α is the scale parameter; and γ is the shape parameter. The dependency of covariates has been adjusted through log-linear transformation, indicating the survival time T followed the Weibull distribution with varying scale parameter $\alpha(x) = \alpha \cdot \exp(\boldsymbol{\beta}'\boldsymbol{X})$. The shape parameter γ was set at 1.2.

2.3.2 Implementation and simulation results

We examined our proposed joint model when: (i) the transition rates were associated with the survival outcome, and (ii) the transition rates did not affect the survival outcome. In both scenarios, we ran 1,000 times of simulations with 150 subjects in the simulated data, yet we burned-in the first ten times of the simulated data. The length of the Markov Chain was set as 30 weeks.

In Scenario (i), when the transition rates were associated with the main outcome (i.e. $b_1 \neq 0$, $b_2 \neq 0$), the proposed joint model performed well in both accuracy and precision (see Table 2.1). Except that the estimate of the transition rate μ had a slightly higher bias, the estimates of the parameter were close to the true value. The coverage probability for each parameter ranged from 83% to 96%, which was acceptable. Besides, the standard deviation (SD)

of the estimates across all simulation runs were close to the average of the square root of estimated variance (SE) for each run, indicating the number of simulation runs was appropriate.

Table 2.5 Simulation results of the proposed model (N=150 subjects, R=1,000, Weeks=30)

Covariates in joint model	Parameter true value	Estimate	Bias	SD[a]	SE[b]	Coverage Probability for 95% CI
Intercept	β_0 =2.5	2.31	-0.19	1.88	2.15	0.85
x_1 (binary)	β_1 =0.3	0.30	0.00	0.19	0.19	0.94
x_2 (continuous)	β_2 =1	1.00	0.00	0.09	0.09	0.93
γ (Weibull shape)	γ =1.1	1.13	0.03	0.08	0.08	0.94
λ	b_1 =-0.5	-0.70	-0.20	0.99	0.90	0.96
w_0 (intercept)	c_0 =-0.7	-0.81	-0.11	0.11	0.11	0.83
w_1 (binary)	c_1 =0.1	0.09	-0.01	0.10	0.09	0.94
w_2 (continuous)	c_2 =0.3	0.27	-0.03	0.07	0.06	0.91
μ	b_2 =0.6	1.26	0.66	4.28	4.98	0.86
v_0 (intercept)	d_0 =-0.8	-0.85	-0.05	0.11	0.11	0.85
v_1 (binary)	d_1 =-0.15	-0.14	0.01	0.09	0.09	0.87
v_2 (continuous)	d_2 =-0.2	-0.19	0.01	0.40	0.41	0.86

[a]Standard deviation of the point estimates.
[b]Standard error, calculated from the average of squared root of estimated variance for each run.

Moreover, to evaluate the performance of the proposed model in the condition of null transition rates effect on the outcome (i.e. Scenario (ii)), we set $b_1 = 0$ and $b_2 = 0$ when we simulated the survival outcome. Table 2.2 shows the result of Scenario (ii). Similar to Scenario (i), the estimates were close to the true parameters except for the transition rate μ, and SD was also close to SE at each parameter while the coverage probability ranged from 0.85 to 0.96. As a result, the proposed model could become a regular Weibull model when there was no effect from the transition rates.

We also tested the simulated data in each scenario in the regular Weibull model for model performance. When the transition rates exert no effect on the survival outcome (Scenario (ii)), the regular Weibull model should perform better as compared to the model performance in Scenario (i). The bias in Scenario (i) should be larger, and the coverage probability would be

lower. Table 2.3 shows the summary of the Weibull model performance in each scenario. Overall, the model performance in Scenario (i) included higher bias, especially in the intercept with a lower coverage probability at 88%. However, the rest of the parameters, x_1, x_2, and γ, had the same performance in each scenario.

Table 6.2 Simulation results using the proposed model to analyze data with no Markov Chain effect included on the survival outcome (N=150 subjects, R=1,000, Weeks=30)

Covariates in joint model	Parameter true value	Estimate	Bias	SD[a]	SE[b]	Coverage Probability for 95% CI
Intercept	β_0 =2.5	2.19	-0.31	1.89	2.23	0.85
x_1 (binary)	β_1 =0.3	0.30	0.00	0.20	0.19	0.94
x_2 (continuous)	β_2 =1	1.00	0.00	0.09	0.09	0.93
γ (Weibull shape)	γ =1.1	1.13	0.03	0.08	0.08	0.94
λ	$b_1 = 0$	-0.01	-0.01	0.98	0.91	0.96
w_0 (intercept)	c_0 =-0.7	-0.81	-0.11	0.11	0.11	0.84
w_1 (binary)	c_1 =0.1	0.09	-0.01	0.10	0.09	0.93
w_2 (continuous)	c_2 =0.3	0.27	-0.03	0.07	0.06	0.91
μ	$b_2 = 0$	0.76	-0.76	4.33	5.18	0.86
v_0 (intercept)	d_0 =-0.8	-0.85	-0.05	0.11	0.11	0.85
v_1 (binary)	d_1 =-0.15	-0.14	0.01	0.09	0.09	0.86
v_2 (continuous)	d_2 =-0.2	-0.19	0.01	0.41	0.42	0.86

[a]Standard deviation of the point estimates.
[b]Standard error, calculated from the average of squared root of estimated variance for each run.

Table 2.7 Simulation results of Weibull model in two scenarios (N=150 subjects, R=1,000, Weeks=30)

Covariates in joint model	Parameter true value	Estimate	Bias	SD[a]	SE[b]	Coverage Probability for 95% CI
Scenario (i). Markov Chain effect on survival outcome. ($b_1 \neq 0$, $b_2 \neq 0$)						
Intercept	β_0 =2.5	2.38	-0.12	0.17	0.16	0.88
x_1 (binary)	β_1 =0.3	0.30	0.00	0.19	0.19	0.95
x_2 (continuous)	β_2 =1	1.01	0.01	0.09	0.09	0.94
γ (Weibull shape)	γ =1.1	1.11	0.00	0.08	0.08	0.95
Scenario (ii). No Markov Chain effect on survival outcome. ($b_1 = 0$, $b_2 = 0$)						
Intercept	β_0 =2.5	2.49	-0.01	0.17	0.16	0.94
x_1 (binary)	β_1 =0.3	0.30	0.00	0.19	0.19	0.95
x_2 (continuous)	β_2 =1	1.01	0.01	0.09	0.09	0.93
γ (Weibull shape)	γ =1.1	1.11	0.01	0.08	0.08	0.95

[a]Standard deviation of the point estimates.
[b]Standard error, calculated from the average of squared root of estimated variance for each run.

2.4 Application

2.4.1 Study Population and Description of the Proposed Joint Model

We also applied the proposed model to a study called *NSCLC study* in this paper, which includes non-small cell lung cancer (NSCLC) patients recruited between 2004-2009 at The University of Texas MD Anderson Cancer Center (MDACC) thoracic medical oncology clinic. The patients were at least 18 years old and they all had with late-stage IIIB and IV NSCLC. The inclusion criteria of the study included English speaking ability, performance score between 0-2, and first-line chemotherapy scheduled. A total of 94 patients were eligible for the study. The PROs were collected using the MD Anderson Symptom Inventory (MDASI) before and after chemotherapy treatment cycles. Two additional common symptom items in NSCLC, coughing and constipation, were also collected through MDASI. The patients were scheduled to complete the MDASI assessment at baseline, chemotherapy treatment (weekly), and post-treatment follow-up. The survival data were then collected from the MDACC medical record system or followed up by the research coordinators.

In a previous publication [22] using the NSCLC data, it was stated that baseline coughing score (a binary variable using score 4 as cut-point) could predict the overall survival with a hazard ratio (HR) at 8.69 (95% CI: 3.53 - 21.38) while adjusting for patients' performance score and race in the model. Fatigue and shortness of breath were also found associated with the overall survival in the study. The symptoms change (treated as continuous variables) between baseline and at the end of first chemotherapy cycle showed the HR was 2.41 (95% CI: 1.32 – 4.37) and 2.30 (95% CI: 1.19 – 4.43) for fatigue and shortness of breath, respectively.

The motivation of this application is to fully utilize the longitudinal PRO data in the survival analysis when individual heterogeneity in longitudinal PRO symptoms is observed. Our

proposed joint model with Markov Chain and the Weibull model may provide a different solution to this type of study by using PROs at baseline and complete chemotherapy cycles to predict survival outcome happened after the treatment period.

2.4.1.1 Markov Chain Submodel

We used fatigue in this application because the previous study showed that changes in fatigue score were associated with overall survival [22]. We dichotomized the MDASI-fatigue scoring system (score 0 - 10) into a binary variable using a cut-point at 4, where score greater than or equal to 4 indicated symptoms at the moderate/severe state, and score less than 4 indicated symptoms at the none/mild state [23, 24]. For each symptom, we excluded subjects who had never reported a symptom change across the two states (moderate/severe vs. none/mild). The transition from none/mild to moderate/severe was called worsening rate represented by λ, and the transition from moderate/severe to none/mild was called as improving rate, and represented by μ. The longitudinal fatigue data were also transformed into the transition types of data to fit the Markov Chain submodel.

2.4.1.2 Main outcome model (Weibull model)

The post-treatment survival is the outcome in the Weibull regression model. It is calculated from the end of treatment date to the date of death, or the last date followed up. The Weibull regression model included individual transition rates (λ and μ must be included), demographic variables, and some clinical variables.

2.4.2 Model selection

During the model selection process, both Markov Chain submodel and the Weibull submodel could adjust with demographic variables and clinical variables. Demographic variables included age, gender (female vs. male), race (white vs. non-white), marital status

(currently married vs. not married), and job (currently employed vs. not employed). Clinical variables included baseline cancer stage (IV vs. III), previous treatment (Yes/No), solid tumor (Yes/No), Charlson Comorbidity Index (CCI, 1 or higher vs. 0), Eastern Cooperative Oncology Group Performance Status (ECOG-PS, 2 vs 0-1), ever smoke (Yes/No), and length of smoking before study.

For the model selection process, we applied a backward selection strategy in two steps:

Step 1. Model selection in the Markov Chain submodel. We considered all variables (demographic or clinical) in the transition rates and selected an optimal model when only transition rates were included as covariates in the Weibull model.

Step 2. Model selection in the Weibull model. After Step 1, we conducted another model selection process in the Weibull model which included both demographic variables and clinical variables. However, the significance of covariates in transition rates might change after the covariates were incorporated in the Weibull model.

2.4.3 Results of the joint model

We excluded one patient who had no changes in fatigue score between the two states (none/mild vs. moderate/severe), and we also excluded patients who had missing data in any of the demographic variables and clinical variables during the model selection. As a result, there were 81 patients used in the analysis. 70% of these patients died in the study, with a median post-treatment survival time of 43 weeks. Besides, the average treatment period was 17.2 weeks, ranging from 6 to 42 weeks.

Table 2.4 shows the results of the optimal model of the joint Markov Chain and Weibull model. Worsening rate was associated with covariates such as age, previous surgery, and ECOG-PS, while improving rate was associated with the length of smoking, previous treatment,

and CCI. The Weibull submodel included both transition rates adjusted by other covariates like age, cancer stage, race, gender, previous treatment, and ECOG-PS. The worsening rate λ showed a negative association with the survival time, but the improving rate μ showed a positive association with on the survival outcome.

2.4.3.1 Interpretation of the transition rate effect on the outcome

For a unit increment in the worsening rate (λ), meaning the time from none/mild to moderate/severe was shorter, or the patient reported more often in the moderate/severe state during the chemotherapy cycles, the overall survival time was expected to change by a multiplicity of $\exp(-0.04) = 0.96$ while holding other covariates as constants. This indicated a 4% decrease for a unit increment in the worsening rate when the time unit in the Markov Chain was a week (7 days). On the other hand, a unit increase in improving rate (μ), meaning a patient reported more frequently in the none/mild state during treatment cycles, the survival time would change by a multiplicity of $\exp(2.19) = 8.94$ while adjusting with other covariates. Even though both transition rates were not statistically significant at a 0.05 level, the directions of estimates for transition rates were estimated as expected. A higher worsening rate shortened the survival time, and a higher improving rate would prolong the survival time. We expect the proposed model to perform better in estimation when the length of Markov Chain (longitudinal PRO observations) is longer. Both transition rates followed the proportional hazard assumption.

2.4.3.2 Interpretation of covariates to the outcome

When interpreting the effect of covariates on the survival outcome, special attention was paid because some covariates were included in both the Weibull submodel and in the Markov Chain submodel. For example, age was selected in both the Weibull model and the worsening rate in the Markov Chain model. Since age was standardized before it was used in the model,

when a unit of a standard deviation increase on age (i.e. 9.1 years), the survival time would change by a multiplicity of $\exp(0.33 + (-0.04 \times \exp(-0.70 + 0.001 \times 1))) = 1.36$ when holding other covariates as constants.

For the covariates selected only for the Weibull submodel, the interpretation is the same as that for any standard Weibull model. For instance, female patients (variable gender, using male as the reference group) had 46% ($\exp(0.48) = 1.62$) longer survival time than male patients while controlling other variables as constants.

Table 2.8 Result of the joint Markov Chain and Weibull model applied to the NSCLC study

Covariates in the joint model	Estimate	SE*	Pr > |t|
Weibull submodel			
Intercept	4.84	1.05	< 0.01
Age[a]	0.33	0.22	0.14
Stage[b]	-2.09	0.61	< 0.01
Race[c]	-0.83	0.55	0.14
Gender[d]	0.48	0.40	0.24
Previous Treatment[e]	-0.80	0.59	0.17
ECOG-PS[f]	0.58	0.43	0.17
Weibull shape	0.89	0.12	< 0.01
λ	-0.04	0.51	0.93
μ	2.19	1.62	0.18
Markov Chain submodel for transition rate λ			
w_0 (intercept)	-0.70	0.15	< 0.01
Age[a]	0.001	0.12	0.92
Previous Surgery[g]	1.29	0.32	< 0.01
ECOG-PS[f]	0.30	0.18	0.10
Markov Chain submodel for transition rate μ			
v_0 (intercept)	-0.44	0.20	0.03
Smoking length[h]	0.08	0.09	0.42
Previous Treatment[e]	0.43	0.19	0.03
CCI[i]	-0.52	0.20	0.01

*SE: Standard error, [a]per 9.1 years, [b]stage IV vs. III, [c]white vs. non-white, [d]female vs. male, [e]Yes vs. No, [f]2 vs. 0-1, [g]Yes vs. No, [h]per 7.9 years, [i]1+ vs. 0

2.4.3.3 Interpretation of covariates to the transition rates

In order to account for the baseline covariates effect on the longitudinal PRO trend in the Markov Chain model (i.e., the transition rates), our proposed model allows covariates to be put into the Markov Chain submodel. For example, the dichotomized Charlson Comorbidity Index (CCI) was selected in the improving rate with a negative sign (est $= -0.52$), indicating a lower improving rate ($\exp(-0.52) = 0.59$) would be expected when the CCI was larger than and equal to 1. This association is expected because a higher CCI means the worse condition on the patient and then reduce the improving rate.

2.5 Discussion

In previous sections, we have developed a new methodology that can incorporate longitudinal covariates in a survival framework. We proposed a joint Markov Chain and Weibull model that could adjust a longitudinal covariate in the Weibull model when individual heterogeneity is found in the longitudinal covariate. Our simulation results have shown the model estimation procedure is valid and model performance is good in both accuracy and precision. We also applied the proposed method in a PRO study that examined longitudinal PRO during the treatment period with the post-treatment overall survival of lung cancer patients. Even though the transition rates did not show statistical significance, the direction of estimates for transition rates and other covariates aligned with our expectations.

When comparing our proposed joint modeling to the time-dependent covariate approach in survival model or another joint modeling approach that combine a longitudinal model with a survival model, our proposed approach has a fundamental difference in ability to model the data we analyzed from those other approaches. In our proposed approach, we modeled the longitudinal process (e.g., longitudinal PRO during the treatment cycles) using a CTMC, and

included the generated transition rates as covariates similar to other baseline covariates. The survival event would only happen after the longitudinal process, so we use the post-treatment survival time (calculated from time of treatment completion to the survival event) in the application. The usage of the transition rates, therefore, will behave like other baseline covariates, and these transition rates can perform as predictors to the survival time. Other joint modeling approaches or time-dependent survival model are not appropriate in analyzing this type of dataset because they usually require a longer longitudinal pattern of data and for it continue until the survival event happened.

There are some advantages in this joint Markov Chain and Weibull model. First, modeling longitudinal data in Markov Chain can reduce the dimension of covariates in the main effect model. By using two transition rates in the Weibull submodel to explain the longitudinal effect of the covariate on survival outcome, we are able to interpret the longitudinal trend at the individual level. Second, CTMC tends to be more suitable for clinical data since they are often collected at unequal time intervals due to patients' missing appointments or scheduling problems. Third, the Weibull model is a common parametric approach in the survival analysis that can explain the association between a factor and survival outcome. The nature of this parametric approach also makes it easily to apply the joint model approach using popular statistical software such as SAS or R by outlining the likelihood of the model. Additionally, to our best knowledge, this is the first time that a method integrates CTMC with a Weibull model.

One of the disadvantages of this method is that we may lose some information from the longitudinal covariates when dichotomizing a continuous score into a binary variable. However, this particular disadvantage could also be a benefit for PRO research. For instance, some PRO instruments [19, 25-27] scoring can be considered as an ordinal or a continuous variable, but it is

always a challenge to interpret a unit difference in the original scoring system. Therefore, in the NSCLC study, we transformed the original score from 0 to 10 into a two-level variable based on previous research [24], leading this method to be more clinically meaningful.

Besides, the identifiability problem in Markov Chain could be a challenge when the joint model method is applied to real-world data. If the longitudinal trends (transition between the two states in Markov Chain) are the same in individuals and the number of covariates adjusted is limited, the identifiability problem may cause estimation difficulty and incompleteness in model optimization (e.g., Newton-Raphson optimization).

Moreover, Like other joint model approaches, initial values are important for model optimization. In the simulation process to validate our methodology, the estimates of parameter were the same when using our initial value approach (see Section 2.2.3) or setting initial values as non-informative value as 1. These identical results indicate the proposed computational approach is robust to the specification of initial values. The optimization (i.e., Newton-Raphson optimization) was able to find the global maximum of the likelihood function.

Furthermore, there are other limitations to this model. The length of Markov Chain is also crucial to the model performance. Section 2.3.2 showed the simulation result in 30 unit (in weeks) of time. We also tested the model performance in a longer length of Markov Chains and observed smaller bias and better coverage probability. Besides, in the application, we excluded one subject reported no change in fatigue between none/mild and moderate/severe level. The result of our approach may be biased if the proportion of this group of patients is large.

In conclusion, we built a joint two-state CTMC model and a Weibull model to the situation when a longitudinal covariate effect exists on time-to-event types of outcome. Our proposed method is especially useful when individual heterogeneity is observed in longitudinal

covariates. The model is appropriate for clinical data collection setting, and it is easy to implement using standard statistical software. Our method provides a more flexible alternative to modeling PRO data when patients' longitudinal measurements show distinct trajectories.

Reference

1. Fisher, L.D. and D.Y. Lin, *Time-dependent covariates in the Cox proportional-hazards regression model.* Annual review of public health, 1999. **20**(1): p. 145-157.
2. Hosmer Jr, D.W. and S. Lemeshow, *Applied survival analysis: regression modelling of time to event data (1999).* Eur Orthodontic Soc, 1999: p. 561-2.
3. Klein, J.P. and M.L. Moeschberger, *Survival analysis: techniques for censored and truncated data*. 2006: Springer Science & Business Media.
4. Tsiatis, A.A. and M. Davidian, *A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error.* Biometrika, 2001. **88**(2): p. 447-458.
5. Tsiatis, A.A. and M. Davidian, *Joint modeling of longitudinal and time-to-event data: an overview.* Statistica Sinica, 2004: p. 809-834.
6. Pinsky, M. and S. Karlin, *An introduction to stochastic modeling*. 2010: Academic press.
7. Ross, S.M., *Stochastic processes (Vol. 2)*. 1996, John Wiley & Sons New York.
8. Molenberghs, G. and G. Verbeke, *Longitudinal data analysis.* Wiley StatsRef: Statistics Reference Online, 2014: p. 1-28.
9. Hickey, G.L., et al., *Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues.* BMC medical research methodology, 2016. **16**(1): p. 117.

10. Sukkar, R., et al. *Disease progression modeling using hidden Markov models*. in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. 2012. IEEE.

11. Yesavage, J.A., et al., *Modeling the prevalence and incidence of Alzheimer's disease and mild cognitive impairment.* Journal of psychiatric research, 2002. **36**(5): p. 281-286.

12. Costin, H. and O. Geman, *Parkinson's disease prediction based on multistate markov models.* International Journal of Computers Communications & Control, 2013. **8**(4): p. 525-537.

13. Faglioni, P., et al., *Parkinson's disease affects automatic and spares intentional verbal learning a stochastic approach to explicit learning processes.* Cortex, 1995. **31**(4): p. 597-617.

14. Karlsson, M.O., et al., *A pharmacodynamic Markov mixed-effect model for the effect of temazepam on sleep.* Clinical Pharmacology & Therapeutics, 2000. **68**(2): p. 175-188.

15. Wu, C.-H., *Analysis of bivariate longitudinal discrete data: A joint continuous-time Markov chains approach.* 2016.

16. Yu, X., et al., *Joint models of dynamics of mothers' stress and children's disease.* Communications in Statistics-Simulation and Computation, 2018: p. 1-16.

17. Ho, C.-H., *Logistic regression with Markov chains as covariates*. 2011: The University of Texas School of Public Health.

18. Rubin, M.L., et al., *A joint logistic regression and covariate-adjusted continuous-time Markov chain model.* Statistics in medicine, 2017. **36**(28): p. 4570-4582.

19. Center, T.U.o.T.M.A.C. *The MD Anderson Symptom Inventory*. 2018 [cited 2018 Nov. 15]; Available from: https://www.mdanderson.org/research/departments-labs-

institutes/departments-divisions/symptom-research/symptom-assessment-tools/md-anderson-symptom-inventory.html.

20. Li, Y.-P. and W. Chan, *Analysis of Longitudinal Multinomial Outcome Data.* Biometrical Journal, 2006. **48**(2): p. 319-326.

21. Kleinman, K. and N.J. Horton, *SAS and R: Data management, statistical analysis, and graphics*. 2009: Chapman and Hall/CRC.

22. Wang, X.S., et al., *Prognostic value of symptom burden for overall survival in patients receiving chemotherapy for advanced nonsmall cell lung cancer.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2010. **116**(1): p. 137-145.

23. Berger, A.M., et al., *Cancer-related fatigue, version 2.2015.* Journal of the National Comprehensive Cancer Network, 2015. **13**(8): p. 1012-1039.

24. Wang, X.S., et al., *Prevalence and characteristics of moderate to severe fatigue: a multicenter study in cancer patients and survivors.* Cancer, 2014. **120**(3): p. 425-432.

25. Aaronson, N.K., et al., *The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology.* JNCI: Journal of the National Cancer Institute, 1993. **85**(5): p. 365-376.

26. Care, R.H. *36-Item Short Form Survey (SF-36)*. 2018 [cited 2018 Nov. 15]; Available from: https://www.rand.org/health/surveys_tools/mos/36-item-short-form.html#section-navigation.

27. Group, E. *EQ-5D*. 2018 [cited 2018 Nov. 15]; Available from: https://euroqol.org/.

3. Journal Article 2

The Relationship Between the Longitudinal Covariates During the Treatment Period as a Markov Chain and The Post-Treatment Survival in Cox Regression

Statistical Methods in Medical Research

Keywords: Continuous-Time Markov Chain (CTMC), longitudinal data, patient-reported outcome (PRO), survival analysis, Cox PH regression

3.1 Background

Patient-reported outcome (PRO) has become a popular mechanism for measuring patient experience in medical research when more and more clinical trials use it as their primary outcome or exploratory outcome [1-3]. Government agencies such as US Food and Drug Administration (FDA) and the European Medicine Agency (EMA) have provided guidelines for PRO research and specified the considerations of PRO application [4-7]. So far, PRO has been linked to survival outcomes and research has shown that PRO monitoring during chemotherapy can improve survival in patients with advanced cancer diagnosis [8]. Also, baseline PRO has been used to predict survival outcome in previous studies [9, 10]. Nevertheless, when PRO contains repeated measurements, the effect of longitudinal PRO on the survival outcome becomes complex, and its relevant statistical models need more sophisticated considerations. Due to the nature of PRO that collects clinical response directly from the patients and can be varied across patients, it is common to observe distinct trajectories of PRO on each patient in a longitudinal study. However, current methods, such as a linear mixed model approach in joint modeling [11], are not able to account for distinct trajectories in individuals. Therefore, in this paper, we propose a joint approach that utilizes the transition rates of a Continuous-Time

Markov Chain (CTMC) generated from longitudinally observed PROs during the treatment period as covariates in a Cox Proportional Hazard Regression (Cox PH model) [12] to evaluate the effect of longitudinal PRO on post-treatment survival.

Longitudinal PRO Analysis

Longitudinal PROs can provide more symptom and health-related quality of life (HRQoL) information related to the disease and treatment involved. A comprehensive design collecting longitudinal PRO data at baseline, treatment visits (weekly or bi-weekly), and post-treatment follow-up (bi-weekly, monthly, or every six months) [13]. Therefore, PRO instruments can measure symptom change over the whole study period, provide detailed information on the differences before, during, and after treatment, and capture variation of symptoms. The trajectory of longitudinal PRO may not always be linear. In fact, in most of the observational studies on cancer with chemotherapy [14, 15], a weekly measured PRO during the chemotherapy cycles tends to show a non-linear pattern (waves). This type of non-linear data can actually indicate whether a symptom is changed by the influence of the disease and/or by the effectiveness/burden of a treatment.

Longitudinal PROs as Outcomes

There are many approaches to analyzing longitudinal PROs, and most of them consider longitudinal PROs as outcome variables. Lin et al. [16], for example, used a Linear Mixed-effect Model (LMM) to analyze the longitudinal HRQoL data from patients who had lumbar spine surgery in Taiwan. They found neurological functions, sleep quality, and depressive symptoms were the key factors affecting quality of life. In fact, LMM has been widely used in longitudinal PRO research because the model accounts for within-subject correlation, random data

missingness, and is easily interpreted. For more information, Fairclough has described how to analyze a longitudinal PRO as an outcome [17].

Another approach to analyzing longitudinal PROs is to transform longitudinal information into a one-dimensional variable and analyze it using regular models depending on the type of the transformed variable. For example, Shi et al. [18] transformed longitudinal PROs into binary or ordinal variables using Group-Based Trajectory Modeling (GBTM), and then treated these variables as outcomes in an (ordinal) logistic regression. Compared with Lin et al.'s approach [16] which assumed the longitudinal PROs followed a general linear trend, Shi et al.'s approach was based on the assumption that the subjects had different patterns of longitudinal PROs.

PROs as Covariates to Survival Outcome

A number of other studies also treated PROs as covariates and investigated the association between PROs and outcomes: Quinten et al. [19] suggested patients-reported symptoms (measured by EORTC-QLQ-C30) plus clinical rating (CTCAE) could more accurately predict overall survival. Also, Armstrong et al. [9] linked the PROs (using both EORTC-QLQ-C30/BN20 and MDASI-BT) to examine the progression-free survival (PFS) and overall survival (OS) in patients with brain tumors. They found that higher symptom burden on patients had a higher risk of OS and PFS, yet a higher QoL score was associated with a lower risk of OS and PFS. In addition, Wang et al. [10] found baseline coughing and symptom worsening from baseline to the end of the first chemotherapy cycle on fatigue and shortness of breath (measured by MD Anderson Symptom Inventory, MDASI) increased the risk of death in a study on late-stage non-small cell lung cancer patients. However, most studies conducted so far

only used PROs at a specific time point as covariates in their models. None of them has directly applied longitudinal PROs as covariates in the model to predict survival outcomes.

Markov Chain Model

One useful statistical approach to modeling a longitudinal covariate process is Markov Chain, and it has already been used to describe the progression of Alzheimer's disease [20, 21] and Parkinson's disease [22, 23] when the diseases' process follows the Markov Property, in which the state of the future events only depends on the current state, not any past state. Previous studies have included a Markov Chain as covariates in their models for cross-sectional outcomes. Ho [24] modeled a longitudinal DNA-damage process as a Discrete Time Markov Chain (DTMC) to predict the occurrence of lung cancer in a case-control study. Rubin et al. [25] later extended this concept to a Continuous Time Markov Chain (CTMC) and modeled longitudinal intracranial pressure (ICP) with CTMC as covariates in a logistic regression model to predict the future status of patients with traumatic brain injuries.

In this paper, we developed a joint approach that uses CTMC as a longitudinal covariate to predict a survival type of outcome that happens after the longitudinal process in a two-stage procedure. The model description, the likelihood function, and the estimation procedure are illustrated in Section 3.2. Sections 3.3 shows the simulation results of the proposed method, while Section 3.4 shows an application on a study with head and neck cancer patients. The advantages and challenges of the proposed method and conclusion of the study are presented in Section 3.5.

3.2 Methods

3.2.1 Continuous-Time Markov Chain

Ross [26] and Pinsky and Karlin [27] have illustrated Continuous-Time Markov Chain with detailed specification and methodology. In this paper, we used the transition rates of a two-state time-homogeneous CTMC as covariates into a Cox PH model. For a CTMC, the transition probabilities can be determined by solving a system of differential equations. When the number of the states is equal to 2, the explicit formulas of transition probabilities are as follows:

$$P_{11}(t) = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

$$P_{12}(t) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

$$P_{21}(t) = \frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

$$P_{22}(t) = \frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)t}$$

Where $P_{ij}(t)$ is the transition probability from state i to state j, $i, j = 1,2$. λ is the transition rate from State 1 to State 2, and μ is the transition rate from State 2 to 1. Given that there are N subjects in the study sample, the parameters λ and μ can be estimated using the maximum likelihood method. The likelihood function can then be written as

$$L = \prod_{n=1}^{N} \left(\prod_{i=1}^{2} P[Z(0) = i]^{I_i[Z_n(0)]} \right)$$

$$\left(\prod_{k=2}^{T_n} \prod_{i=1}^{2} \prod_{j=1}^{2} \left[P_{Z(t_{n,k-1})=i,\, Z(t_{n,k})=j}\left(t_{n,k} - t_{n,k-1}\right) \right]^{I_i[Z(t_{n,k-1})] I_j[Z(t_{n,k})]} \right)$$

$$= \prod_{n=1}^{N} \left(P[Z(0) = 1]^{I_1[Z_n(0)]} P[Z(0) = 2]^{I_2[Z_n(0)]} \right) \cdot$$

$$\left(\prod_{k=2}^{T_n}\left[\frac{\mu}{\lambda+\mu}+\frac{\lambda}{\lambda+\mu}e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})}\right]^{I_1[Z(t_{n,k-1})]I_1[Z(t_{n,k})]}\right.$$
$$\left[\frac{\lambda}{\lambda+\mu}-\frac{\lambda}{q_{12}+\mu}e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})}\right]^{I_1[Z(t_{n,k-1})]I_2[Z(t_{n,k})]}$$
$$\left[\frac{\mu}{\lambda+\mu}-\frac{\mu}{\lambda+\mu}e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})}\right]^{I_2[Z(t_{n,k-1})]I_1[Z(t_{n,k})]}$$
$$\left.\left[\frac{\lambda}{\lambda+\mu}+\frac{\mu}{\lambda+\mu}e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})}\right]^{I_2[Z(t_{n,k-1})]I_2[Z(t_{n,k})]}\right)$$

where $i = 1, 2$,

$$I_i[Z(t)] = \begin{cases} 1, if\ Z(t) = i \\ 0, otherwise \end{cases}$$

and $P[Z(0) = i]$ is the probability when the initial state is equal to i.

In clinical studies, researchers are usually interested in studying the effect of different subject characteristics on a particular outcome. To achieve this goal, in a Markov model, the dependency of covariates on the transition rates can be modeled through a log link: $\lambda(v) = \mathrm{e}^{\boldsymbol{v}'\boldsymbol{g}}$, and $\mu(w) = \mathrm{e}^{\boldsymbol{w}'\boldsymbol{h}}$, where $\boldsymbol{g}$ represents a vector of coefficients for a covariate vector $\boldsymbol{v}$, and $\boldsymbol{h}$ represents a vector of coefficients for a covariate vector $\boldsymbol{w}$. Note that the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ can be identical, completely different, or partially overlapped in their components.

3.2.2 Cox Proportional Hazard Model (Cox PH Model)

The Cox PH model has been widely used in clinical research because it is easy to interpret a survival type of outcome as a risk of disease. Klein and Moeschberger [28] and Hosmer and Lemeshow [29] have described the model specification and estimation procedure of the Cox PH model in detail. Here, we briefly outline the components of the Cox PH model:

Let T denote the time to the event, and δ be the event indicator ($\delta = 1$ if the event occurs, and $\delta = 0$ if the event is right-censored). $\boldsymbol{Z}$ is a vector of baseline covariates which may affect the survival distribution of T. The Cox PH model can be expressed as

$$h(t|\boldsymbol{Z}) = h_0(t)\exp(\boldsymbol{\beta}'\boldsymbol{Z})$$

where $h(t|Z)$ is the hazard at time t given the covariate set $\boldsymbol{Z}$. Also, $h_0(t)$ is the baseline hazard, and $\boldsymbol{\beta}$ is a vector of coefficients with respect to covariate set $\boldsymbol{Z}$. Since the baseline hazard, $h_0(t)$, is usually not the interest of a study, we can only estimate the covariate effect $\boldsymbol{\beta}$ through the partial likelihood method. The partial likelihood for the Cox PH model can be written as

$$L(\beta) = \prod_{i=1}^{D} \frac{\exp(\beta' Z_i)}{\sum_{j \in R(t_i)} \exp(\beta' Z_j)}$$

where D represents the set of subjects when an event occurs, and $R(t_i)$ is the set of subjects at risk at time t_i.

3.2.3 A Joint Two-Stage Approach

We developed a prognostic modeling approach that is a combination of the Cox PH model with a two-state CTMC. This approach allowed us to study the effect of a longitudinal covariate process on a survival type of outcome in terms of the hazard rate while controlling for other baseline covariates in both CTMC and the Cox PH model.

Let T be a time-to-event outcome, and Z(t) be a homogeneous CTMC with a state space $S = \{1, 2\}$ characterized by the transition rates λ and μ. X is a vector directly related to the outcome T. $\boldsymbol{v}$ is a vector of covariates related to T through the transition rate λ, and $\boldsymbol{w}$ is a vector of covariates related to T through the transition rate μ. The proposed joint approach can be expressed as

$$h(t|Z) = h_0(t)\exp(\boldsymbol{X}'\boldsymbol{\beta} + \boldsymbol{q_i'b}) = h_0(t)\exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_r X_r + b_1\lambda + b_2\mu)$$

$$= h_0(t)\exp(\beta_0 + \beta_1 X_1 + \cdots + \beta_r X_r + b_1 e^{\boldsymbol{v}'\boldsymbol{g}} + b_2 e^{\boldsymbol{w}'\boldsymbol{h}})$$

where $\boldsymbol{\beta}$ is the effect of the baseline covariate vector $\boldsymbol{X}$ on the outcome T; $\boldsymbol{b} = (b_1, b_2)$ is the effect of unobservable transition rates $\boldsymbol{q_i} = (\lambda_i, \mu_i)$ on the outcome, and $h_0(t)$ is the baseline hazard.

We can acquire the estimates through the partial likelihood function. For a study with N subjects, the partial likelihood of the proposed model can be expressed as

$$L(\boldsymbol{h}, \boldsymbol{g}, \boldsymbol{b}, \boldsymbol{\beta}) =$$

$$\prod_{i=1}^{D} \frac{\exp(\boldsymbol{X'\beta} + \boldsymbol{q_i'b})}{\sum_{j \in R(t_i)} \exp(\boldsymbol{X'\beta} + \boldsymbol{q_j'b})} = \prod_{i=1}^{D} \frac{\exp(\boldsymbol{X'\beta} + b_1 e^{\boldsymbol{v'g}} + b_2 e^{\boldsymbol{w'h}})}{\sum_{j \in R(t_i)} \exp(\boldsymbol{X'\beta} + b_1 e^{\boldsymbol{v'g}} + b_2 e^{\boldsymbol{w'h}})}$$

The notations here are the same as those in Section 3.2.1 and Section 3.2.2.

In terms of the estimation procedure, we proposed a two-stage approach as follows:

Stage 1. Estimate the unobservable transition rates λ and μ for a longitudinal covariate in a two-state CTMC. First, we transformed a longitudinal covariate into transition type of data and applied the method that has been described in Section 3.2.1 and built the complete likelihood function. With the log link function, we incorporated baseline covariates at each transition rate. After that, we used the maximum likelihood method to obtain the estimates of transition rates λ_i and μ_i at the individual level.

Stage 2. Model the transition rates λ and μ as covariates in the Cox PH model. We used the partial likelihood method to evaluate the covariate effects on the survival outcome including transition rates from Stage 1 and baseline covariates.

3.2.4 Interpretation of Transition Rates

Transition rates represented the speed of transition from one state to another in a longitudinal process. Thus, evaluating the longitudinal covariate effect on the survival outcome and interpreting the transition rate correctly are important to this joint approach. The association

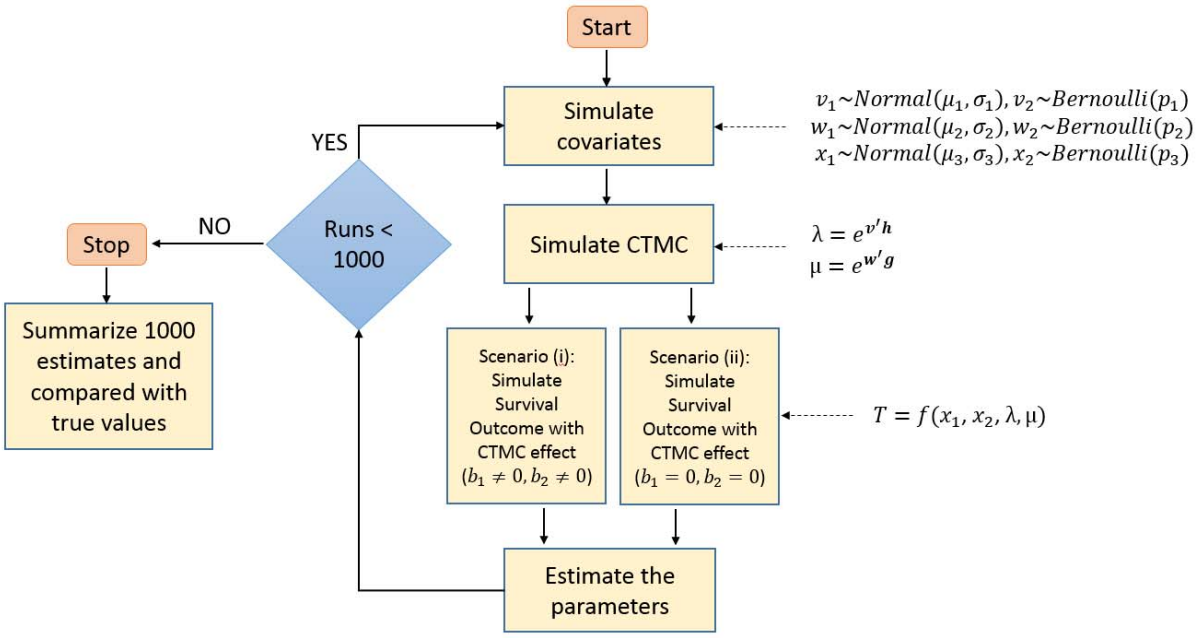
of the transition rates λ and μ with the hazard rate at time t, $(h(t|Z))$, can be interpreted as follows: For each unit increment in the transition rate λ, the hazard rate is expected to change by a multiplicity of e^{b_1}, while holding other covariates as constants. Similarly, for each unit increment in the transition rate μ, the hazard rate is expected to change by a multiplicity of e^{b_2}, while other covariates are held as constant.

3.3 Validation of the Estimation Procedure and Simulation

3.3.1 Description of the Simulation Study

We examined our proposed joint approach through simulations to validate our two-stage estimation procedure. The parameter values were obtained from a PRO study [10] where limited covariates were adjusted in the model as described below. Three majot steps were involved in the simulation procedure: simulation of covariates, simulation of Markov Chain, and simulation of time-to-event outcomes. Figure 3.1 outlines the scheme of the simulation procedure.

Figure 3.1. Simulation procedure of Markov Chain and time-to-event outcomes



3.3.1.1 Simulation of Covariates

To reflect the situation of real-world data, we incorporated a set of covariates (a binary variable and a continuous variable) in both CTMC and the Cox PH model.

In CTMC: Two covariates, one continuous (v_1) and one binary (v_2) are simulated as predictors directly associated with the transition rate λ. v_1 followed a Normal distribution, and v_2 followed a Bernoulli distribution. Similarly, one continuous variable (w_1) that followed a Normal distribution, and one binary variable (w_2) that followed a Bernoulli distribution were simulated for the transition rate μ.

In the Cox PH model: Two covariates (one continuous type (x_1) and one binary (x_2) type) were also simulated to explain the effect of baseline characteristics on the survival outcome. Similar to the setting of CTMC, x_1 followed a Normal distribution, and x_2 followed a Bernoulli distribution.

3.3.1.2 Simulation of Markov Chain

Once the covariate sets $\boldsymbol{v} = (v_1, v_2)$ and $\boldsymbol{w} = (w_1, w_2)$ were generated from Section 3.3.1.1, the individual transition rates λ_i and μ_i could be calculated through the log link function described in Section 3.2.1. Next, CTMC was also simulated [30] based on individual transition rates. As a result, each subject had their own simulated longitudinal observations with covariates adjusted.

3.3.1.3 Simulation of Time-To-Event Outcome

When the covariates x_1, x_2, λ, and μ were simulated and calculated from the previous steps, we then simulated survival outcome that followed a Weibull distribution. Survival time T which followed a Weibull distribution could be expressed as

$$T = [h_0 \times \exp(-(\beta_1 x_1 + \beta_2 x_2 + b_1 \lambda + b_2 \mu))]^{-1/\gamma}$$

where h_0 is a pre-specified baseline hazard [29, 31], and γ is the shape parameter in the Weibull distribution. The dependency of covariates had been adjusted through a log-linear transformation, so the scale parameter $\alpha(x)$ varied with covariates. The shape parameter γ was set at 1.2 in the simulation.

3.3.2 Implementation and Simulation Results

Two simulation situations, Scenario (i) and Scenario (ii), were set up to examine the proposed joint approach. In Scenario (i), we simulated the survival outcome when CTMC had an association with the survival outcome. In contrast, we assumed a situation when there was no CTMC effect on the survival outcome in Scenario (ii), and set $b_1 = 0$ and $b_2 = 0$. For each scenario, we ran 1,000 replicates, and for each of replicate, the length of CTMC was set at 10 weeks with 150 subjects at each run.

Table 3.1 shows the results of the proposed joint approach in Scenario (i) when $b_1 \neq 0$, and $b_2 \neq 0$. The bias, difference between the estimate and true parameter value, of all parameters was small, and the coverage probabilities were appropriate and ranged from 0.90 to 0.96. The standard deviation (SD) of all the estimates across all simulation runs were close to the average of the square root of the estimated variance (SE) at each run for all parameters, indicating the number of simulation runs was appropriate.

In Scenario (ii), when there was no transition effect on the survival outcomes (i.e. $b_1 = 0$ and $b_2 = 0$), the proposed method detected the null transition rate effect in the simulated data (see Table 3.2). The biases for all the parameters were small, and coverage probabilities were good with a range from 90% to 97%. In terms of stability, SD and SE in Scenario (ii) were close to each other for all the parameters. Therefore, the proposed approach became the regular Cox PH model without the covariate effect from the transition rates.

Table 3.1. Simulation results of the proposed model (N=150 subjects, R=1,000, weeks=10)

Covariates in joint model	Parameter true value	Estimate	Bias	SD[a]	SE[b]	Coverage Probability for 95% CI
x_1 (binary)	β_1 = 0.03	0.30	0.00	0.21	0.21	0.95
x_2 (continuous)	β_2 = -0.2	-0.21	-0.01	0.09	0.09	0.96
λ	b_1 = 0.6	0.74	0.14	0.84	0.70	0.93
w_0 (intercept)	c_0 = -0.5	-0.62	-0.12	0.19	0.19	0.90
w_1 (binary)	c_1 = 0.1	0.10	0.00	0.16	0.16	0.96
w_2 (continuous)	c_2 = 0.3	0.28	-0.02	0.11	0.11	0.95
μ	b_2 = -0.5	-0.55	-0.05	0.98	0.67	0.92
v_0 (intercept)	d_0 = -0.4	-0.43	-0.03	0.30	0.29	0.97
v_1 (binary)	d_1 = -0.16	-0.15	0.01	0.16	0.16	0.95
v_2 (continuous)	d_2 = 0.5	0.47	-0.03	0.25	0.24	0.96

[a]Standard deviation of the point estimates.
[b]Standard error, calculated from the average of squared root of estimated variance for each run.

Table 3.2. Simulation results using the proposed model to analyze data with no Markov Chain effect included on the survival outcome (N=150 subjects, R=1,000, weeks=10)

Covariates in joint model	Parameter true value	Estimate	Bias	SD[a]	SE[b]	Coverage Probability for 95% CI
x_1 (binary)	β_1 = 0.03	0.30	0.00	0.21	0.21	0.95
x_2 (continuous)	β_2 = -0.2	-0.21	-0.01	0.09	0.09	0.96
λ	b_1 = 0	0.00	0.00	0.81	0.70	0.94
w_0 (intercept)	c_0 = -0.5	-0.62	-0.12	0.17	0.18	0.90
w_1 (binary)	c_1 = 0.1	0.10	0.00	0.15	0.15	0.96
w_2 (continuous)	c_2 = 0.3	0.28	-0.02	0.10	0.10	0.95
μ	b_2 = 0	0.01	0.01	0.87	0.67	0.94
v_0 (intercept)	d_0 = -0.4	-0.43	-0.03	0.24	0.25	0.97
v_1 (binary)	d_1 = -0.16	-0.15	0.01	0.15	0.14	0.97
v_2 (continuous)	d_2 = 0.5	0.47	-0.03	0.20	0.21	0.96

[a]Standard deviation of the point estimates.
[b]Standard error, calculated from the average of squared root of estimated variance for each run.

Additionally, we evaluated the performance of the simulated data (Scenario (i) and (ii)) in a regular Cox PH model. The goal for this additional analysis was to further examine the simulated data and expect to observe the differences in the longitudinal covariate effect on the survival outcome in two scenarios. When the effect of transition rates on the survival outcome is null (i.e., Scenario (ii)), the regular Cox PH model is expected to perform better in Scenario (ii)

than in Scenario (i). Table 3.3 shows the summary of the regular Cox PH model performance in each scenario. Overall, the result in Scenario (i) showed a slightly higher bias on the binary variable x_1, while all the other results remained the same in both scenarios.

Table 3.3. Simulation results of the Cox PH model in two scenarios (N=150 subjects, R=1,000, weeks=10)

Covariates in joint model	Parameter true value	Estimate	Bias	SD[a]	SE[b]	Coverage Probability for 95% CI
Scenario (i). Markov Chain effect on survival outcome. ($b_1 \neq 0$, $b_2 \neq 0$)						
x_1 (binary)	$\beta_1 = 0.3$	0.29	-0.01	0.21	0.21	0.95
x_2 (continuous)	$\beta_2 = -0.2$	-0.20	0.00	0.09	0.09	0.96
Scenario (ii). No Markov Chain effect on survival outcome. ($b_1 = 0$, $b_2 = 0$)						
x_1 (binary)	$\beta_1 = 0.3$	0.30	0.00	0.21	0.21	0.95
x_2 (continuous)	$\beta_2 = -0.2$	-0.20	0.00	0.09	0.09	0.96

[a]Standard deviation of the point estimates.
[b]Standard error, calculated from the average of squared root of estimated variance for each run.

3.4 Application

3.4.1 Study Population and Description of the Joint Approach

We used a longitudinal PRO study (called HN study) with head and neck cancer patients recruited from the Head and Neck Clinic at MD Anderson Cancer Center (MDACC) from February 2006 to August 2007. All the patients were 18 years old or older. Due to the limited data collected on patients with other ethnicities (less than 10% of the originally collected data), we only used non-Hispanic white patients as a cohort in the analysis. The study was approved by the MD Anderson Institutional Review Board, and all the patients had completed the informed consent before the first evaluation of PRO. The PROs were collected by MDASI-HN [32]. Patients completed their MDASI-HN assessment weekly at baseline and the treatment

(radiotherapy (RT) or chemoradiotherapy (CRT)) period. The survival outcome was collected from the MDACC medical record system.

One hundred eighty one non-Hispanic white patients participated in the study. Among them, 21 patients were excluded from the analysis because the length of PRO repeated measurements were not long enough (less than six). Thus, 160 patients were included in the application of the proposed joint approach. Among these patients, about half (51.52%) received radiotherapy (RT) only, and the other half received chemoradiotherapy (CRT). The average age of the selected patients was 58.88 years old, and 21.9% of them were female. Additionally, one fourth (27.81 %) of these patients had late stage (tumor stage III and IV) head and neck cancer. The median progression-free survival time was 270 weeks, and the median length of repeated PRO measurements was 7 observations.

In the literature, several studies had been published using this HN cohort. For example, Shi et al. [18] used GBTM (see Section 3.1) to characterize the subjects into several symptom severity groups with the average of the top five severe symptoms. Specially, they used the modeled severity membership from GBTM as an outcome in logistic regression to find the clinical factors that were associated with the outcome. Moreover, Rosenthal et al. [33] used LMM to evaluate the linear pattern of the symptoms in MDASI-HN between the treatment arms (RT vs. CRT), and differences in symptoms such as fatigue, nausea, and disturbed sleep were found.

3.4.1.1 Markov Chain model

In order to build a prognostic model, we only selected PRO at baseline and PRO during the treatment period in CTMC. Symptom scores (0-10) were dichotomized to a binary variable with a cutpoint at 4 based on previous research [34, 35]. Scores equal to and larger than 4 were

defined as moderate/severe level, and scores less than 4 were identified as none/mild level. The transition from none/mild to moderate/severe was termed the worsening rate (λ), and the transition from moderate/severe to none/mild was called the improving rate (μ). Baseline demographical variables and baseline clinical variables were evaluated in CTMC through model selection. The CTMC model was conducted in PROC NLMIXED on SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

3.4.1.2 Cox PH model

Progression-free survival (PFS) was used as the main outcome of interest in this application. The survival time is calculated from the end of the treatment period to a survival event (progression-free) or to the end of the study (as censored). We modeled the transition rates of a symptom (must be included) and adjusted with baseline covariates in the Cox PH model.

3.4.2 Model Selection

In both CTMC and the Cox PH model, we considered covariates such as demographic variables and clinical variables in our model selection. Demographical variables included age, gender (female vs. male), education (high school and above vs. below high school), and employment status (employed vs. not employed). Clinical variables included treatment (CRT vs. RT), Eastern Cooperative Oncology Group Performance Status (ECOG-PS, 0 vs. 1+), tumor stage, previous induction chemotherapy (yes/no) and radiation dose (standardized).

The model selection of the proposed joint approach was separated into two steps:

Step 1. Covariates in CTMC. We selected covariates in CTMC using a backward approach based on Wald's test statistics.

Step 2. Covariates in the Cox PH model. This step is to identify meaningful clinical covariates to the survival outcome while the transition rates of a symptom were included.

3.4.3 Results of the Proposed Model

Table 3.4. Result of the proposed joint approach of Markov Chain and Cox PH model applied to the HN study on fatigue

Covariates in the joint model	Estimate	SE*	Pr > |t|
Cox PH model			
Gender[a]	-0.29	0.30	0.34
Age[b]	0.14	0.15	0.36
Treatment[c]	-0.73	0.35	0.04
λ (1/10 of a unit)	0.11	0.17	0.52
μ (1/10 of a unit)	-0.51	0.21	0.01
Markov Chain for transition rate λ			
w_0 (intercept)	-1.00	0.11	<0.01
Age[b]	0.14	0.11	0.19
Gender[a]	-0.36	0.23	0.12
Previous induction chemotherapy[d]	-0.64	0.28	0.02
Markov Chain for transition rate μ			
v_0 (intercept)	-1.52	0.25	<0.01
Age[b]	0.10	0.14	0.49
Radiation dose[e]	-0.07	0.12	0.52
Education[f]	0.42	0.26	0.11
Treatment[c]	-0.55	0.25	0.03

*Standard error, [a]female vs. male, [b]per 11.5 years, [c]CRT vs RT, [d]yes/no, [e]per 4.4 gray, [f]high school and above vs. below high school.

Several symptoms were used as an application for the proposed joint approach. However, to better demonstrate a clinical application of our method, we only present one model involving one symptom fatigue which demonstrated strong associations between the transition rates and PFS (see Table 3.4). Age, gender, and treatment were selected to the model for the worsening rate λ, while age, education, radiation dose, and treatment were included for the improving rate μ**.** In the Cox PH model, the worsening rate λ had a positive association with the risk of disease progression. However, the improving rate μ had a negative effect on the risk of disease progression. Other covariates that were adjusted in the Cox PH modeling included gender, age, and treatment.

3.4.3.1 Interpretation of the Transition Rate Effect on the Outcome

Transition rates represent the speed of a process from one state to another. In other words, transition rates can also be interpreted as the time consumption between states. If the worsening rate becomes larger, the time of transition from none/mild to moderate/severe will be shorter, the patients should also have reported more frequently moderate/severe symptoms. In contrast, if a patient had reported more frequently none/mild levels of a symptom, his/her improving rate should be larger.

In Table 3.4, for a one-tenth of a unit increment of worsening rate (λ), a patient's risk of head and neck cancer progression was expected to change by a multiplicity of $\exp(0.11) = 1.12$ when other covariates remained constant. This means the hazard rate of disease progression would increase by 12% when there was a 0.1 unit change in the worsening rate. Similarly, if the improving rate (μ) increased by a one-tenth unit, a patient's risk of cancer progression would decrease by 40% ($\exp(-0.51) = 0.60$). Based on Wald's test statistics, the p-value of improving rate was statistically significant (pvalue=0.01), while the worsening rate did not show a significant result (p-value=0.52). Both transition rates followed proportional hazard assumption.

3.4.3.2 Interpretation of Covariate Effects on the Outcome

The interpretation of covariate effects on the outcome in this joint approach depended on the inclusion of covariates in each submodel. If the covariate only existed in the Cox PH model, the interpretation of that covariate effect was the same as the regular Cox PH model. However, if a covariate existed in both CTMC and the Cox PH model, the interpretation should include the covariate effect through the transition rates in CTMC. For example, for a unit of standard deviation change on age (i.e., 11.5 years), the effect of the progression risk of head and neck

cancer is expected to change by a multiplicity of $\exp(0.13 + 0.11 \times \exp(-1.00 + 0.14 \times 1) - 0.51 \times \exp(-1.52 + 0.10 * 1)) = 1.05$. Therefore, the older the patient, the higher the risk will be.

3.4.3.3 Interpretation of the Covariates to the Transition Rates

Some baseline covariates were adjusted in the transition rates of CTMC, and their effect on the transition rates was easy to interpret. For example, treatment was selected only in the improving rate at CTMC. The improving rate was expected to change by a multiplicity of $\exp(-0.55) = 0.58$ when comparing CRT to RT. The patients who received CRT tended to report less none/mild state of fatigue as compared to those receiving RT. This result is consistent with the previous research by Rosenthal et al. [33], who found that CRT patients reported more symptom burden than RT patients.

3.5 Discussion

In this paper, we developed a joint approach of CTMC and the Cox PH model that models an observed longitudinal process as covariates to evaluate the effect of the covariates on survival. We validated the two-stage joint approach through simulations and found the model performed well in terms of accuracy and precision for two different scenarios. We also applied the proposed joint approach to a PRO study. The results showed our model was adaptable in clinical research, and the interpretation of the transition rates provide a comprehensive way to interpret longitudinal PRO when distinct trajectories are observed for each subject.

It is noteworthy to discuss the differences among our method, the time-dependent covariates in a Cox PH model [29, 36], and the joint modeling that comprised a longitudinal model and a survival model [11, 37]. Time-dependent covariates approach in the Cox PH model

evaluates the association between a longitudinal covariate and the survival outcome in different periods, and the longitudinal process is followed until the survival event occurs. Similarly, in the joint modeling approach, previous studies [11, 37] have built different joint models from a longitudinal model (e.g., LMM) and a Cox PH model. This joint modeling approach also measures the longitudinal covariate process along with the survival time. The survival outcome may happen at the same time as the longitudinal process modeled by LMM. In contrast, we used the longitudinal covariate information during the treatment period that would only happen before the survival event. The longitudinal process at the treatment period is considered as the information collected before the measurement of the survival time, like other baseline covariates, and then this information is used to predict the post-treatment survival events, which makes our proposed method a prognostic model.

Moreover, our proposed method could reduce the dimension of data information in two ways. First, the original scale is reduced from a continuous variable to a binary variable. When the PRO score originally with a 0-10 scale was dichotomized to a binary variable, the interpretation of PRO in research becomes more practical in clinical practice. For example, patients who reported score 3 and score 4 were not considered as the same severity level because a score 4 would generate further clinical action per the clinical guidance even though there was just a one unit difference in scores [38]. With our proposed method, even though we lost some information due to the transformation from a continuous variable to a binary variable, this loss is acceptable because the binary variable could still provide information on the differences of symptom severity levels. Second, the repeated PRO measurements are reduced to a set of transition rates as summary statistics, and then incorporated to the Cox PH model. This makes the proposed joint approach more acceptable to clinical researchers because they can summarize

longitudinal PRO as baseline covariates in the Cox PH model as the approach they commonly conducted.

There are other advantages in applying the proposed joint approach. First, compared to GBTM which transformed longitudinal data into membership of grouping, the proposed joint approach still presents the change of a longitudinal variable, which makes it suitable to consider the longitudinal variable as covariates in the survival model. Second, CTMC is adequate to the random missingness in clinical data collection. It is true that most studies have pre-specified timeline/schedule to collect clinical/PRO data, but a patient may miss a visit/response to the study. When data are missing at random in a longitudinal covariate, CTMC can model the time between states. Third, the proposed joint approach can be easily applied to common statistical software such as SAS or R through the two-stage procedure.

Nevertheless, this joint approach is subject to several limitations. First of all, lack of enough repeated observations in a longitudinal covariate and a small number of covariates adjusted in transition rates may lead to an identifiability problem [39, 40] because some patients may have the same profile in covariates and the same longitudinal pattern in CTMC. Moreover, the baseline covariates may diminish the effect of transition rates. In Section 3.4, we opted to keep transition rates in the model to maintain the relationship between the longitudinal variable and the survival outcome. However, users can apply different model selection approaches based on their research question.

To conclude, we developed a joint approach of the CTMC and Cox PH model and demonstrated that the estimation procedure is valid. We also applied this approach to a study using longitudinal PRO fatigue at baseline and during treatment course on patients with head and neck cancer. The application results indicate the trajectory of fatigue during the treatment period

is associated with post-treatment PFS. Therefore, when distinct trajectories are observed among subjects in a longitudinal process, our proposed joint approach is expected to provide an alternative to evaluate the effect of a longitudinal process on survival outcomes.

Reference

1. Mercieca-Bebber, R., et al., *The importance of patient-reported outcomes in clinical trials and strategies for future optimization.* Patient related outcome measures, 2018. **9**: p. 353.
2. Scoggins, J.F. and D.L. Patrick, *The use of patient-reported outcomes instruments in registered clinical trials: evidence from ClinicalTrials. gov.* Contemporary clinical trials, 2009. **30**(4): p. 289-292.
3. Vodicka, E., et al., *Inclusion of patient-reported outcome measures in registered clinical trials: evidence from ClinicalTrials. gov (2007–2013).* Contemporary clinical trials, 2015. **43**: p. 1-9.
4. (EMA), E.M.A. *Integrating patients' views in clinical studies of anticancer medicines*. 2016 [cited 2018 11/12]; Available from: https://www.ema.europa.eu/news/integrating-patients-views-clinical-studies-anticancer-medicines.
5. (EMA), E.M.A., *Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies.*, EMA, Editor. 2016.
6. US Department of Health and Human Services, F.a.D.A., *Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance.* Health and Quality of Life Outcomes, 2006. **4**: p. 1-20.
7. US Department of Health and Human Services, F.a.D.A., *Guidance for Industry: Patient-Reported Outcome Measures - Use in Medical Product Development to Support Labeling Claims* 2009.

8. Basch, E., et al., *Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment.* Jama, 2017. **318**(2): p. 197-198.

9. Armstrong, T.S., et al., *Net clinical benefit analysis of radiation therapy oncology group 0525: a phase III trial comparing conventional adjuvant temozolomide with dose-intensive temozolomide in patients with newly diagnosed glioblastoma.* Journal of Clinical Oncology, 2013. **31**(32): p. 4076.

10. Wang, X.S., et al., *Prognostic value of symptom burden for overall survival in patients receiving chemotherapy for advanced nonsmall cell lung cancer.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2010. **116**(1): p. 137-145.

11. Tsiatis, A.A. and M. Davidian, *A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error.* Biometrika, 2001. **88**(2): p. 447-458.

12. Cox, D.R., *Regression models and life-tables (with discussion).* The Journal of the Royal Statistical Society, 1972. **B 34**: p. 182-220.

13. Hassett, M.J., et al., *High-priority topics for cancer quality measure development: Results of the 2012 American Society of Clinical Oncology Collaborative Cancer Measure Summit.* Journal of oncology practice, 2014. **10**(3): p. e160-e166.

14. Elting, L.S., et al., *Patient-reported measurements of oral mucositis in head and neck cancer patients treated with radiotherapy with or without chemotherapy: demonstration of increased frequency, severity, resistance to palliation, and impact on quality of life.* Cancer, 2008. **113**(10): p. 2704-2713.

15. Wang, X.S., et al., *Longitudinal study of the relationship between chemoradiation therapy for non–small-cell lung cancer and patient symptoms.* Journal of clinical oncology, 2006. **24**(27): p. 4485-4491.

16. Lin, E.-Y., et al., *Trajectory of health-related quality of life and its determinants in patients who underwent lumbar spine surgery: a 1-year longitudinal study.* Quality of Life Research, 2018: p. 1-9.

17. Fairclough, D.L., *Design and analysis of quality of life studies in clinical trials*. 2010: Chapman and Hall/CRC.

18. Shi, Q., et al., *Using group-based trajectory modeling to examine heterogeneity of symptom burden in patients with head and neck cancer undergoing aggressive non-surgical therapy.* Quality of Life Research, 2013. **22**(9): p. 2331-2339.

19. Quinten, C., et al., *Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival.* Journal of the National Cancer Institute, 2011. **103**(24): p. 1851-1858.

20. Sukkar, R., et al. *Disease progression modeling using hidden Markov models*. in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. 2012. IEEE.

21. Yesavage, J.A., et al., *Modeling the prevalence and incidence of Alzheimer's disease and mild cognitive impairment.* Journal of psychiatric research, 2002. **36**(5): p. 281-286.

22. Costin, H. and O. Geman, *Parkinson's disease prediction based on multistate markov models.* International Journal of Computers Communications & Control, 2013. **8**(4): p. 525-537.

23. Faglioni, P., et al., *Parkinson's disease affects automatic and spares intentional verbal learning a stochastic approach to explicit learning processes.* Cortex, 1995. **31**(4): p. 597-617.

24. Ho, C.-H., *Logistic regression with Markov chains as covariates*. 2011: The University of Texas School of Public Health.

25. Rubin, M.L., et al., *A joint logistic regression and covariate-adjusted continuous-time Markov chain model.* Statistics in medicine, 2017. **36**(28): p. 4570-4582.

26. Ross, S.M., *Stochastic processes (Vol. 2)*. 1996, John Wiley & Sons New York.

27. Pinsky, M. and S. Karlin, *An introduction to stochastic modeling*. 2010: Academic press.

28. Klein, J.P. and M.L. Moeschberger, *Survival analysis: techniques for censored and truncated data*. 2006: Springer Science & Business Media.

29. Hosmer Jr, D.W. and S. Lemeshow, *Applied survival analysis: regression modelling of time to event data (1999).* Eur Orthodontic Soc, 1999: p. 561-2.

30. Li, Y.-P. and W. Chan, *Analysis of Longitudinal Multinomial Outcome Data.* Biometrical Journal, 2006. **48**(2): p. 319-326.

31. Kleinman, K. and N.J. Horton, *SAS and R: Data management, statistical analysis, and graphics*. 2009: Chapman and Hall/CRC.

32. Rosenthal, D.I., et al., *Measuring head and neck cancer symptom burden: the development and validation of the MD Anderson symptom inventory, head and neck module.* Head & Neck: Journal for the Sciences and Specialties of the Head and Neck, 2007. **29**(10): p. 923-931.

33. Rosenthal, D.I., et al., *Patterns of symptom burden during radiotherapy or concurrent chemoradiotherapy for head and neck cancer: A prospective analysis using the*

University of Texas MD Anderson Cancer Center Symptom Inventory-Head and Neck Module. Cancer, 2014. **120**(13): p. 1975-1984.

34. Berger, A.M., et al., *Cancer-related fatigue, version 2.2015.* Journal of the National Comprehensive Cancer Network, 2015. **13**(8): p. 1012-1039.

35. Wang, X.S., et al., *Prevalence and characteristics of moderate to severe fatigue: a multicenter study in cancer patients and survivors.* Cancer, 2014. **120**(3): p. 425-432.

36. Fisher, L.D. and D.Y. Lin, *Time-dependent covariates in the Cox proportional-hazards regression model.* Annual review of public health, 1999. **20**(1): p. 145-157.

37. Tsiatis, A.A. and M. Davidian, *Joint modeling of longitudinal and time-to-event data: an overview.* Statistica Sinica, 2004: p. 809-834.

38. Shi, Q., T.R. Mendoza, and C.S. Cleeland, *Interpreting Patient-reported Outcome Scores for Clinical Research and Practice: Definition, Determination, and Application of Cutpoints.* Medical care, 2019. **57**: p. S8-S12.

39. Benoit, J.S., et al., *A hidden Markov model approach to analyze longitudinal ternary outcomes when some observed states are possibly misclassified.* Statistics in medicine, 2016. **35**(9): p. 1549-1557.

40. Thom, H.H., et al., *State selection in Markov models for panel data with application to psoriatic arthritis.* Statistics in medicine, 2015. **34**(16): p. 2456-2475.

4. Journal Article 3

A Dimension Reduction on Multiple Continuous-Time Markov Chains with Application to Longitudinal Patient-Reported Outcomes

Journal of Applied Statistics

Keywords: Continuous-Time Markov Chain (CTMC), longitudinal data, factor analysis, dimension reduction, patient-reported outcome (PRO)

4.1 Background

Factor analysis (FA) has been widely applied to patient-reported outcome (PRO) research as a tool when developing PRO instruments to identify unobserved domains. Most PRO instrument validation studies are cross-sectional studies, but PRO studies collecting longitudinal data can also be used to examine the unobserved domains over time [1-3]. Specifically, some studies [4, 5] used Longitudinal Factor Analysis (LFA) [6] to examine the hidden factors on different occasions, but LFA could not address the longitudinal pattern of the original item. To address this issue, in this paper, we propose a novel approach that integrates a two-state Continuous-Time Markov Chain (CTMC) to FA, which not only models longitudinal PRO data as CTMCs at the individual level but also incorporates the generated transition rates into FA. Because the proposed approach applied FA on summary statisticsof longitudinal PRO changes, our method is a more comprehensive approach that interprets the underlying construct of longitudinal variables.

Factor Analysis in PRO research

There are two types of FA, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). The former is used to explore potential unobserved factors from the observed

variables without imposing a preconceived number of factors. On the other hand, CFA is used to test the underlying structure of the observed variables when the unobserved factor structure is usually known [7-9]. In this paper, to discover the underlying construct among longitudinal variables, we will mainly focus on EFA in PRO research and our proposed method.

EFA, in fact, has been extensively used in the development of PRO instruments to test the reliability of items in an instrument. For example, Rosenthal et al. [10] have used EFA to identify the underlying structure of the cancer symptom items in the MD Anderson Symptom Inventory-Head and Neck Module (MDASI-HN). They found the module items could be classified into two factors: one factor included problems with mouth/throat sores, tasting food, constipation, problems with teeth or gum, and skin pain/burning/rash; the other factor included difficulty with voice or speech, choking or coughing, problems with chewing or swallowing, and problems with mucus. Two other studies [11, 12] also applied EFA in a similar strategy to explore the underlying factors among the symptom items in different modules. Furthermore, EFA is also employed in other PRO instrument development or validation studies, such as the Medical Outcomes Study Short Form-36 (SF-36) [13, 14], Functional Assessment of Cancer Therapy-General (FACT-G) [15, 16], and Patient-Reported Outcomes Measurement Information System (PROMIS) [17, 18]. All of these studies performed EFA on a single occasion, even though some [10-12] collected PRO data at more than one event. However, if longitudinal PRO data were analyzed, the underlying structure of the items in the instrument could have provided more useful information.

Longitudinal Factor Analysis

Longitudinal factor analysis (LFA) is a special version of structural equation modeling (SEM), which is a general statistical modeling technique where a latent factor construct is of

interest. In general, LFA can perform factor analysis across time in a single modeling framework by incorporating the correlation among the repeated measurements and simultaneously estimating the unobserved factors over time [6]. The result of LFA could show the underlying factor structure at each occasion as well as the correlation within the observed variables. LFA has been extensively utilized in psychology. For example, Corballis and Traub [19] proposed a model that can simultaneously estimate two constructs of FA for two occasions, and the model considered the relationship between observed variables at two different time points. Then, Joreskog and Sorbom [20], extended the concept and developed a program (LISREL), which can perform LFA for multiple occasions (more than two occasions).

Longitudinal Factor Analysis in PRO Research

However, only few studies have applied LFA in PRO research. So far, the Longitudinal Analysis of Patient-Reported Outcomes Working Group has summarized their discussion related to the methodologies that could be used in future PRO research and commented on the usage of LFA in longitudinal PRO work [21]. For example, LFA can be applied to item response theory (IRT) [22] models to access the variability of latent structure over time, and distinguish response changes due to response shift or true changes. Another study then used LFA to test the stability of the 13-item sense of coherence (SOC) in a longitudinal study on breast cancer patients [4]. In this study, Lindblad et al. applied LFA to examine the change of the underlying constructs on two occasions. They concluded that the SOC scale was stable, so it was suitable to measure life stress in women with breast cancer. On the other hand, LFA requires complete data over time, and the length of occasions should be the same. So, if patients have different lengths of repeated measurements (such as PROs during the treatment period), only those data that overlapped at the same time points can be used in LFA.

Continuous-Time Markov Chain

To accommodate the situation when the lengths of the repeated measurements are different for patients, one solution is to apply Continuous-Time Markov Chain at the individual level. Markov Chain models have been applied to studies such as Alzheimer's disease [23, 24] and Parkinson's disease [25, 26] in describing disease progression. When the disease progression status follows the Markov Property, and can be expressed as changes among different states, a Markov Chain becomes an appropriate model to estimate those changes among the states. For example, Karlsson et al. used the first-order Markov model to describe changes of different sleep levels (lighter sleep vs. deeper sleep) from patients who had insomnia [27], and they found the insomnia treatment (temazepam) improved patients' sleep. In other applications, Markov Chains can also be used to describe the changes of a patient's disease status over time, and then transition rates estimated from these Markov Chains can be incorporated into a second model as covariates. For example, Rubin et al [28] modeled longitudinal intracranial pressure (ICP, high ICP vs. normal/low ICP) with Continuous Time Markov Chain (CTMC) as covariates in a logistic regression which allowed them to predict the future status of patients with traumatic brain injuries.

In this paper, we will apply CTMC to longitudinal PRO data and describe the changes of symptom severity during the treatment period. Specifically, we will apply a two-state CTMC to a longitudinal PRO study that has collected symptom changes in head and neck cancer patients, and then incorporates the estimated transition rates from the two-state CTMC to FA to explore the underlying factors representing structure in the transition rates. These factors may be helpful because they could reflect longitudinal PRO process during the treatment period, and they may indicate whether a symptom change is related to the treatment or to the disease.

Next, in Section 4.2, we will describe the two-state CTMC and its estimation procedure. FA will also be briefly described in the same section. Section 4.3 will show how the proposed approach could be applied to a real-world study and examines the usage of CTMC in FA. A comprehensive discussion of our proposed method will be held in Section 4.4.

4.2 Method

4.2.1 Two-State CTMC

Let $S = 1, 2$ denote the two-state values for a two-state homogeneous CTMC, and let $\boldsymbol{Q}$ be its associated transition rate matrix given by

$$\boldsymbol{Q} = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

where λ is the transition rate from State 1 to State 2, and μ is the transition rate from State 2 to State 1. For detailed specification and methodology of CTMC, see textbooks by Ross [29] and Pinsky and Karlin [30]. In this paper, we will calculate the transition rates of a two-state time-homogeneous CTMC for each longitudinal variable and integrate them to FA. For a CTMC, the transition probabilities can be determined by solving a system of differential equations. When the number of states is equal to 2, the explicit formulas for these probabilities are as follows:

$$P_{11}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

$$P_{12}(t) = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

$$P_{21}(t) = \frac{\mu}{\lambda + \mu} - \frac{\mu}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

$$P_{22}(t) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

where $P_{ij}(t)$ is the transition probability from State i to State j, $i, j = 1, 2$. Given a length T for a longitudinal variable of a subject, the parameters λ and μ can be estimated using the maximum likelihood method. The individual level likelihood function for a longitudinal variable can be written as

$$L = \left(\prod_{i=1}^{2} P[Z(0) = i]^{I_i[Z_n(0)]} \right)$$

$$\left(\prod_{k=2}^{T} \prod_{i=1}^{2} \prod_{j=1}^{2} \left[P_{Z(t_{n,k-1})=i,\, Z(t_{n,k})=j}\left(t_{n,k} - t_{n,k-1}\right) \right]^{I_i[Z(t_{n,k-1})] I_j[Z(t_{n,k})]} \right)$$

$$= \left(P[Z(0) = 1]^{I_1[Z_n(0)]} P[Z(0) = 2]^{I_2[Z_n(0)]} \right) \cdot$$

$$\left(\begin{array}{l} \prod_{k=2}^{T} \left[\frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})} \right]^{I_1[Z(t_{n,k-1})] I_1[Z(t_{n,k})]} \\ \left[\frac{\lambda}{\lambda+\mu} - \frac{\lambda}{q_{12}+\mu} e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})} \right]^{I_1[Z(t_{n,k-1})] I_2[Z(t_{n,k})]} \\ \left[\frac{\mu}{\lambda+\mu} - \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})} \right]^{I_2[Z(t_{n,k-1})] I_1[Z(t_{n,k})]} \\ \left[\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} e^{-(\lambda+\mu)(t_{n,k}-t_{n,k-1})} \right]^{I_2[Z(t_{n,k-1})] I_2[Z(t_{n,k})]} \end{array} \right)$$

where $i = 1, 2$,

$$I_i[Z(t)] = \begin{cases} 1, if\ Z(t) = i \\ 0, otherwise \end{cases}$$

$P[Z(0) = i]$ is the probability when the initial state is equal to i. For each subject, we will fit a two-state CTMC for a longitudinal variable. Therefore, each subject will have a pair of transition rates (λ, μ), representing the changes of the longitudinal variable.

Interpretation of transition rate

By definition, the transition rate has to be larger than 0 because it represents a moving force from one state to another. When a transition rate increases, the time from one state to

another will be shorter. For example, let λ be the transition rate from State 1 to State 2. If λ is higher in one subject, it means time has elapsed $1/\lambda$ from State 1 to State 2.

4.2.2 Factor analysis

Let X be an observed random vector with dimension $p \times 1$. An orthogonal factor model can be expressed as

$$X - \mu_{P\times 1} = L_{p\times m}F_{m\times 1} + \varepsilon_{p\times 1}$$

where $m < p$, and p is the number of observed variables. m is the number of hypothesized factors. μ is the mean vector for observed random variable X. L is the loading matrix for factors. F is a vector with dimension $m \times 1$. ε is a vector for error terms.

An important assumption of factor analysis is that the hypothesized factors are uncorrelated. This purpose can be achieved by using rotation of the loading matrix. A commonly used rotation method, orthogonal varimax rotation, was applied to this study to ensure each variable loads highly on one and only one factor. Such factor structure will result in each factor representing a distinct construct. The number of m factors can be selected by the methods proposed by Bartholomew et al. [31] or Johnson and Wichern [32]. In this paper, the observed variables are the transition rates estimated from the CTMC for each symptom measured in a PRO instrument. Detailed descriptions will be shown in Section 4.3.

4.3 Application

4.3.1 Study Population and Description of the Proposed Approach

We applied the proposed approach, which integrated individual two-state CTMC into FA, to a head and neck cancer study. The study recruited head and neck cancer patients who received radiotherapy (RT) or concurrent chemoradiotherapy (CCRT) at MD Anderson Cancer

Center from February 2006 to August 2007. PRO was followed using MDASI-HN [10] at baseline (before the treatment) and during each week of treatment. In the analysis, we only included non-Hispanic white patients because less than 10% of subjects were of other racial/ethnic background. All patients signed the informed consent before the first PRO evaluation, and the study was approved by the MD Anderson Institutional Review Board.

In summary, 181 non-Hispanic white patients were recruited in the study, but 21 patients were later excluded because only five repeated measurements of PRO were available. So, only 160 patients remained in the final analysis. Among these 160 patients, about 22% of them were female, and the average age was 59 years old. Also, about half (51%) received the RT in the study.

4.3.2 Implementation

Individual Two-State Continuous-Time Markov Chain

In order to fit the longitudinal PRO in a two-state CTMC, we first dichotomized the original scale from 0-10 to a binary variable, according to previous research [33, 34]. The scores 0-3 were defined as the none/mild level, and scores 4-10 were defined as the moderate/severe level. As a result, the transition rate from none/mild to moderate/severe was called the worsening rate (λ), and the transition rate from moderate/severe to none/mild was called the improving rate (μ).

We used the PRO response collected at baseline and weekly measurements during the treatment period to fit the two-state CTMC at the individual level. Following the methods described in Section 4.2.1, we estimated the transition rates for all the symptom items in the MDASI-HN of each patient. That is, for each symptom item, we used two transition rates to represent the original repeated measurements. Note that some patients may not report changing

symptom severity levels (none/mild vs. moderate/severe) during the treatment period. For example, a patient may only report disturbed sleep at none/mild level across the entire treatment period. The CTMC, as a result, cannot estimate the worsening rate because no data were reported at the moderate/severe level. In this case, we set a minimum boundary 0.01 for transition rates which could not be estimated due to missing transitions. Additionally, because FA assumes normality, we took the natural log of each transition rates before running FA.

Factor Analysis

Once the individual transition rates were generated for each symptom item, we examined the proposed approach in two scenarios. First, we ran FA with both improving rates and worsening rates of all symptoms. This is to test whether the transition rates from CTMC can reasonably explain the changes in longitudinal RPO. Second, we ran FA only with worsening rates for all symptoms. Since the PRO was measured through questions regarding patients' symptom severity, the worsening rates probably could better capture the change of symptoms over time for each patient. All analyses were performed by SAS 9.4 (SAS Institute, Cary, NC).

4.3.3 Results

As described in Section 4.3.2, each patient had two transition rates (worsening rate and improving rate) for each longitudinal symptom. Since the distributions of the transition rates are not the interest of this paper, we only list the descriptive statistics of the transition rates of each symptom in Appendix A.

Figure 4.1 then showes the result of FA when both worsening rates and improving rates were considered in the analysis. We selected the first two factors because we expected FA to distinguish the transition rates by their nature. It was clear that the majority of the worsening rates comprised Factor 1, while the majority of the improving rates comprised Factor 2 except

for the symptoms shortness of breath and vomiting. The loadings for shortness of breath were low: 0.16 in Factor 1 and 0.03 in Factor 2, and the loadings for vomiting were also low: 0.12 in Factor 1 and 0.06 in Factor 2. Therefore, FA differentiated the worsening rates and the improving rates into different factors. Detailed loadings on each factors are presented in Appendix B.

Figure 4.2 showed the result of FA when only the worsening rates of symptoms were used. Based on the inflection point in the scree plot, we selected the first two factors. Factor 1 included symptoms such as pain, fatigue, nausea, lack of appetite, drowsy, dry mouth, problems with mucus, difficulty chewing or swallowing, skin pain/burning/rash, problems with tasting food, mouth/throat sores, and problems with teeth or gum. Overall, Factor 1 represented those symptoms that were worse when receiving the treatments (RT or CCRT). In contrast, Factor 2 contained symptoms such as disturbed sleep, distress, shortness of breath, remembering, drowsy, sadness, vomiting, numbness/tingling, choking/coughing, difficulty in voice or speech, skin pain/burning/rash, and constipation. So, unlike Factor 1, Factor 2 represented the local or systemic symptoms that were commonly reported by head and neck cancer patients.

Figure 4.1. Factor analysis with both worsening rate and improving rate

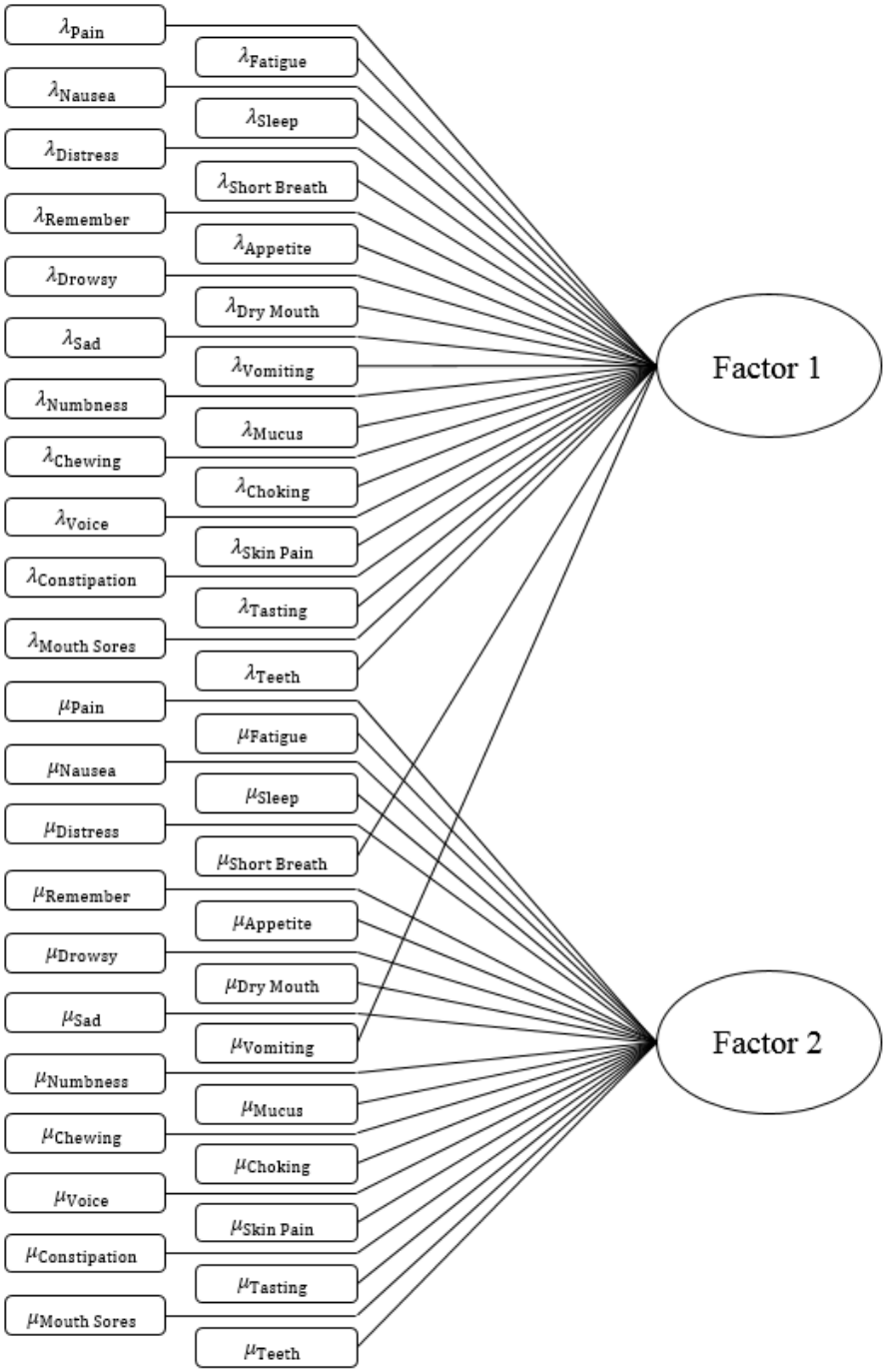
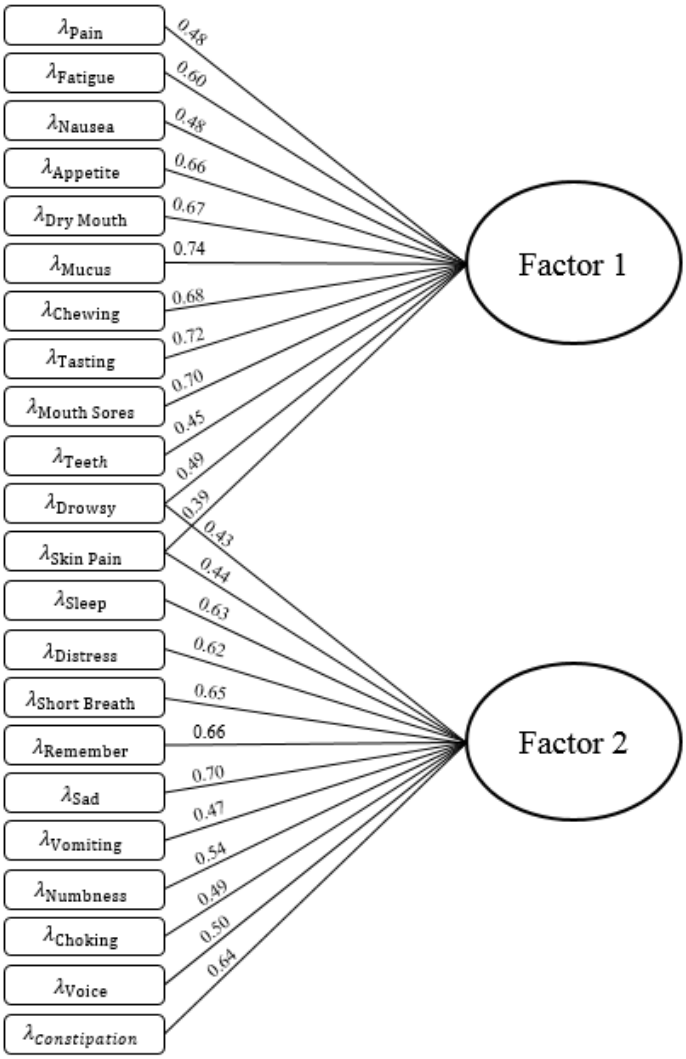


Figure 4.2. Factor analysis with worsening rates only



4.4 Discussion

In the previous sections, we showed how to apply individual two-state CTMC with FA. Using the individual two-state CTMC to summarize the longitudinal PRO symptoms, we were able to decrease the dimension from a longitudinal variable to a set of variables (transition rates) as that in a cross-sectional study. We also inputted these transition rates into FA to explore the performance of transition rates and evaluate their representation in the unobserved factors.

When we included both worsening rates and improving rates in FA, the FA could well identify the nature of the transition rate. Worsening rates and improving rates were separated to different factors. The first factor represented worsening rates, and the second represented improving rates. This result validated the application of using two-state CTMC in longitudinal variables, and it showed that the transition rates could well represent longitudinal patterns. Even though there were two symptoms, shortness of breath and vomiting, that had low loadings on both factors, the overall representation for each factor was still recognizable.

On the other hand, when we only included the worsening rates in FA, the result of FA showed CTMC could sufficiently summarize information from the longitudinal PRO. As indicated in Figure 4.2, Factor 1 represented those worsening rates of symptoms related to the treatment (radiotherapy or chemoradiotherapy), and the remaining worsening rates of symptoms were categorized to Factor 2, which may represent the systemic symptoms of head and neck cancer. Notably, the worsening rate of drowsiness and the worsening rate of skin pain/burning/rash had effects on both factors. In particular, the result from Figure 2 was consistent with the previous literature. For example, Bossi et al. [35, 36] reported pain (called breakthrough cancer pain) were observed during the RT or CCRT among head and neck cancer patients. Also, another study found fatigue was related to the dose of RT [37]. In that study, patients who received higher dose of radiation reported fatigue after the treatment. Other symptoms including mouth sores, dry mouth, difficulty in swallowing, change in tasting food, lack of appetite, rash on skin, and stiff jaw were also common side effects for patients who received RT. Moreover, a previous study [1] that used the same dataset as ours reported the distribution of each MDASI-HN symptom item during the treatment period. Compared with their distributions, the symptom worsening rates in the first factor in our study (Figure 4.2) were

actually the items that had been significantly increased during the treatment period in their study. Patients were asymptomatic at baseline and experienced worsening symptoms during their treatment course. This means those worsened symptoms may have been caused by side effects of the treatment. Furthermore, in Figure 2, we only used the worsening rate in FA because patients who filled MDASI-HN were only asked to report the most severe symptom level in the past 24 hours. Thus, the worsening rate alone could disclose patients' symptom severity during the treatment period and help to simplify the interpretation of results in FA.

Compared to the LFA [6], our method provides a simpler, more comprehensive interpretation when conducting a factor analysis in a longitudinal study. LFA could produce the underlying construct at different occasions by adjusting the correlation from repeated measurements. Thus, the way to interpret of the result from LFA is like interpreting the result from FA multiple times. In contrast, our proposed method summarizes the longitudinal information as a set of variables (transition rates), and then applies these variables to FA as those in a cross-sectional study. Therefore, the interpretation of the results from our method is the same as the interpretation in other studies applying FA in a cross-sectional study. However, our interpretation is based on the worsening rate or improving rate of a symptom.

There are other advantages of our proposed method. First, our method can handle the missing data in the repeated measurements. The CTMC models the transition based on the time between two observed states, and the length of the longitudinal data can be different because we model CTMC at the individual level. Besides, the two-state CTMC can be easily modeled using popular statistical programs such as SAS and R by specifying the likelihood described in Section 4.2.1, and the factor analysis is supported in most statistical programs.

One of the limitations of our proposed approach is information loss through dichotomization of the original scale on PRO. In the two-state CTMC, we dichotomized the original scale (0-10) to a binary variable (none/mild vs. moderate/severe). So, we may lose information on the original scale, but a binary ordinal scale can still provide sufficient information on the severity of patients' symptoms. Second, the cut point selected for the dichotomization of the symptom scales is critical and needs more consideration. In this study, we used score 4 as the cut point because previous studies [33, 34] showed score 4 performed well when differentiating symptoms such as fatigue and pain. However, if a symptom or a PRO measurement has other preferable cut points, users can apply different cut points based on their research interest. As the cutpoint 4 has been recommended by multiple NCCN and ASCO guidelines for symptom management [33, 38-41], the current results are interpretable for further clinical application when symptom monitoring is implemented for patients under active cancer treatments. Third, for patients who only reported symptoms in the same severity level, we set an arbitrary boundary 0.01 for the corresponding transition rates based on the length of the treatment period in the study. As a result, some of transition rate estimates may not directly reflect the true change in the longitudinal process. However, this situation may be improved when PRO is collected over longer periods because, in this scenario, the probability of no transition will be substantiously reduced.

To conclude, we propose a new approach that can comprehensively summarize information from longitudinal PRO studies. Our method uses the two-state CTMC at the individual level to summarize the changes of longitudinal PRO information, and then integrates those estimated transition rates into FA to explore the underlying construct among symptoms. We showed that our approach could provide meaningful clinical information through the

transition rates, and the result of FA proved the utility of CTMC in longitudinal PRO data. It is expected that this approach can help researchers who are developing PRO instruments and needing a solution to explain the underlying construct of instrument items when the longitudinal PRO process is considered.

Reference

1. Rosenthal, D.I., et al., *Patterns of symptom burden during radiotherapy or concurrent chemoradiotherapy for head and neck cancer: A prospective analysis using the University of Texas MD Anderson Cancer Center Symptom Inventory-Head and Neck Module.* Cancer, 2014. **120**(13): p. 1975-1984.
2. Shi, Q., et al., *Using group-based trajectory modeling to examine heterogeneity of symptom burden in patients with head and neck cancer undergoing aggressive non-surgical therapy.* Quality of Life Research, 2013. **22**(9): p. 2331-2339.
3. Wang, X.S., et al., *Longitudinal study of the relationship between chemoradiation therapy for non–small-cell lung cancer and patient symptoms.* Journal of clinical oncology, 2006. **24**(27): p. 4485-4491.
4. Lindblad, C., et al., *Stability of the 13-item sense of coherence (SOC) scale: a longitudinal prospective study in women treated for breast cancer.* Quality of Life Research, 2016. **25**(3): p. 753-760.
5. Schwartz, C.E., et al., *Exploring response shift in longitudinal data.* Psychology & Health, 2004. **19**(1): p. 51-69.
6. TISAK, J. and W. MEREDITH, *Longitudinal factor analysis*, in *Statistical methods in longitudinal research*. 1990, Elsevier. p. 125-149.
7. Brown, T.A., *Confirmatory factor analysis for applied research*. 2014: Guilford Publications.
8. Fabrigar, L.R., et al., *Evaluating the use of exploratory factor analysis in psychological research.* Psychological methods, 1999. **4**(3): p. 272.
9. Thompson, B., *Factor analysis.* The Blackwell Encyclopedia of Sociology, 2007.

10. Rosenthal, D.I., et al., *Measuring head and neck cancer symptom burden: the development and validation of the MD Anderson symptom inventory, head and neck module.* Head & Neck: Journal for the Sciences and Specialties of the Head and Neck, 2007. **29**(10): p. 923-931.

11. Sailors, M.H., et al., *Validating the MD Anderson Symptom Inventory (MDASI) for use in patients with ovarian cancer.* Gynecologic oncology, 2013. **130**(2): p. 323-328.

12. Wang, X.S., et al., *Validation and application of a module of the MD Anderson Symptom Inventory for measuring multiple symptoms in patients with gastrointestinal cancer (the MDASI-GI).* Cancer, 2010. **116**(8): p. 2053-2063.

13. De Vet, H.C., et al., *Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36.* Quality of Life Research, 2005. **14**(5): p. 1203-1218.

14. Ware Jr, J.E., et al., *The factor structure of the SF-36 Health Survey in 10 countries: Results from the IQOLA Project.* Journal of clinical epidemiology, 1998. **51**(11): p. 1159-1165.

15. Smith, A.B., et al., *A Rasch and factor analysis of the Functional Assessment of Cancer Therapy-General (FACT-G).* Health and Quality of Life Outcomes, 2007. **5**(1): p. 19.

16. Winstead-Fry, P. and A. Schultz, *Psychometric analysis of the functional assessment of cancer therapy-general (FACT-G) scale in a rural sample.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 1997. **79**(12): p. 2446-2452.

17. Cella, D., et al., *The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years.* Medical care, 2007. **45**(5 Suppl 1): p. S3.

18. Hays, R.D., et al., *Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items.* Quality of Life Research, 2009. **18**(7): p. 873-880.

19. Corballis, M. and R. Traub, *Longitudinal factor analysis.* Psychometrika, 1970. **35**(1): p. 79-98.

20. Jöreskog, K.G. and D. Sörbom, *LISREL 8: User's reference guide*. 1996: Scientific Software International.

21. Swartz, R.J., et al., *The king's foot of patient-reported outcomes: current practices and new developments for the measurement of change.* Quality of Life Research, 2011. **20**(8): p. 1159-1167.

22. Lord, F.M., *Applications of item response theory to practical testing problems*. 2012: Routledge.

23. Sukkar, R., et al. *Disease progression modeling using hidden Markov models*. in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. 2012. IEEE.

24. Yesavage, J.A., et al., *Modeling the prevalence and incidence of Alzheimer's disease and mild cognitive impairment.* Journal of psychiatric research, 2002. **36**(5): p. 281-286.

25. Costin, H. and O. Geman, *Parkinson's disease prediction based on multistate markov models.* International Journal of Computers Communications & Control, 2013. **8**(4): p. 525-537.

26. Faglioni, P., et al., *Parkinson's disease affects automatic and spares intentional verbal learning a stochastic approach to explicit learning processes.* Cortex, 1995. **31**(4): p. 597-617.

27. Karlsson, M.O., et al., *A pharmacodynamic Markov mixed-effect model for the effect of temazepam on sleep.* Clinical Pharmacology & Therapeutics, 2000. **68**(2): p. 175-188.

28. Rubin, M.L., et al., *A joint logistic regression and covariate-adjusted continuous-time Markov chain model.* Statistics in medicine, 2017. **36**(28): p. 4570-4582.

29. Ross, S.M., *Stochastic processes (Vol. 2)*. 1996, John Wiley & Sons New York.

30. Pinsky, M. and S. Karlin, *An introduction to stochastic modeling*. 2010: Academic press.

31. Bartholomew, D.J., M. Knott, and I. Moustaki, *Latent variable models and factor analysis: A unified approach*. Vol. 904. 2011: John Wiley & Sons.

32. Johnson, R.A. and D.W. Wichern, *Multivariate analysis.* Encyclopedia of Statistical Sciences, 2004. **8**.

33. Berger, A.M., et al., *Cancer-related fatigue, version 2.2015.* Journal of the National Comprehensive Cancer Network, 2015. **13**(8): p. 1012-1039.

34. Wang, X.S., et al., *Prevalence and characteristics of moderate to severe fatigue: a multicenter study in cancer patients and survivors.* Cancer, 2014. **120**(3): p. 425-432.

35. Bossi, P., et al., *The point of pain in head and neck cancer.* Critical reviews in oncology/hematology, 2019.

36. Bossi, P., et al., *The vicious circle of treatment-induced toxicities in locally advanced head and neck cancer and the impact on treatment intensity.* Critical reviews in oncology/hematology, 2017. **116**: p. 82-88.

37. Kamal, M., et al., *Fatigue following radiation therapy in nasopharyngeal cancer survivors: A dosimetric analysis incorporating patient report and observer rating.* Radiotherapy and Oncology, 2019. **133**: p. 35-42.

38. Bower, J.E., et al., *Screening, assessment, and management of fatigue in adult survivors of cancer: an American Society of Clinical oncology clinical practice guideline adaptation.* Journal of clinical oncology, 2014. **32**(17): p. 1840.

39. Holland, J.C. and B.D. Bultz, *The NCCN guideline for distress management: a case for making distress the sixth vital sign.* Journal of the National Comprehensive Cancer Network, 2007. **5**(1): p. 3-7.

40. Paice, J.A., et al., *Management of chronic pain in survivors of adult cancers: American Society of Clinical Oncology Clinical Practice Guideline.* Journal of Clinical Oncology, 2016. **34**(27): p. 3325-3345.

41. Swarm, R.A., et al., *Adult cancer pain.* Journal of the National Comprehensive Cancer Network, 2013. **11**(8): p. 992-1022.

Appendix A

Table 4.1. Distributions of worsening rate and improving rate of symptoms in MDASI-HN

Transition rate	Mean	Std.	Q1	Median	Q3
λ_{pain}	2.00	3.30	0.19	0.38	2.09
μ_{pain}	2.09	3.49	0.01	0.53	2.39
$\lambda_{fatigue}$	2.56	3.68	0.20	0.55	4.40
$\mu_{fatigue}$	1.86	3.34	0.01	0.49	1.83
λ_{nausea}	1.43	2.43	0.01	0.22	1.84
μ_{nausea}	2.72	4.18	0.44	0.56	3.67
λ_{sleep}	1.85	3.17	0.01	0.35	2.48
μ_{sleep}	2.41	3.77	0.34	0.56	3.14
$\lambda_{distress}$	1.30	2.38	0.01	0.01	1.57
$\mu_{distress}$	2.59	4.12	0.52	0.56	2.55
$\lambda_{shortbreath}$	0.72	2.20	0.01	0.01	0.01
$\mu_{shortbreath}$	1.47	2.56	0.53	0.55	0.59
$\lambda_{remember}$	0.76	2.02	0.01	0.01	0.07
$\mu_{remember}$	1.78	3.53	0.53	0.55	0.61
$\lambda_{appetite}$	2.49	3.68	0.08	0.39	3.91
$\mu_{appetite}$	1.92	3.12	0.01	0.54	1.97
λ_{drowsy}	1.75	3.17	0.01	0.32	1.76
μ_{drowsy}	2.10	3.55	0.01	0.54	1.43
$\lambda_{drymouth}$	2.49	4.08	0.23	0.39	4.43
$\mu_{drymouth}$	1.58	3.08	0.01	0.16	1.19
λ_{sad}	1.01	2.16	0.01	0.01	0.60
μ_{sad}	1.90	3.49	0.52	0.55	0.63
$\lambda_{vomiting}$	0.95	2.10	0.01	0.01	0.39
$\mu_{vomiting}$	2.51	4.41	0.53	0.56	0.75
$\lambda_{numbness}$	0.44	1.19	0.01	0.01	0.01

$\mu_{numbness}$	1.62	2.82	0.53	0.56	0.61
λ_{mucus}	2.08	3.64	0.19	0.35	2.63
μ_{mucus}	2.02	3.41	0.01	0.47	1.92
$\lambda_{chewing}$	2.23	3.66	0.19	0.34	2.26
$\mu_{chewing}$	1.84	3.40	0.01	0.53	1.94
$\lambda_{choking}$	1.25	2.35	0.01	0.14	1.67
$\mu_{choking}$	2.66	4.15	0.52	0.56	2.47
λ_{voice}	1.08	2.09	0.01	0.19	0.63
μ_{voice}	2.20	3.92	0.50	0.55	1.20
$\lambda_{skipain}$	1.34	2.23	0.01	0.28	1.66
$\mu_{skipain}$	2.85	4.67	0.01	0.56	2.89
$\lambda_{constipation}$	1.44	2.82	0.01	0.01	1.48
$\mu_{constipation}$	2.57	3.81	0.53	0.57	2.78
$\lambda_{tastefood}$	2.16	3.80	0.23	0.38	2.17
$\mu_{tastefood}$	1.10	2.54	0.01	0.01	0.58
$\lambda_{mouthsore}$	2.03	3.39	0.18	0.33	2.78
$\mu_{mouthsore}$	2.18	3.72	0.01	0.54	2.05
λ_{teeth}	1.10	2.13	0.01	0.15	1.32
μ_{teeth}	2.70	4.15	0.53	0.57	3.83

Appendix B.

Table 4.2. Rotated Factor Loading in Figure 4.1

Variables	Factor1	Factor2
$log(\lambda_{pain})$	0.56	0.10
$log(\lambda_{fatigue})$	0.58	0.13
$log(\lambda_{nause})$	0.54	-0.13
$log(\lambda_{sleep})$	0.68	-0.11
$log(\lambda_{distress})$	0.64	-0.21
$log(\lambda_{shortbreath})$	0.39	-0.36
$log(\lambda_{remember})$	0.48	-0.28
$log(\lambda_{appetite})$	0.64	-0.02
$log(\lambda_{drowsy})$	0.63	-0.09
$log(\lambda_{drymouth})$	0.64	0.06
$log(\lambda_{sad})$	0.64	-0.31
$log(\lambda_{vomiting})$	0.47	-0.34
$log(\lambda_{numbness})$	0.39	-0.32
$log(\lambda_{mucus})$	0.65	0.13
$log(\lambda_{chewing})$	0.68	0.08
$log(\lambda_{choking})$	0.55	0.00
$log(\lambda_{voice})$	0.49	-0.09
$log(\lambda_{skinpain})$	0.59	-0.05
$log(\lambda_{constip})$	0.53	-0.19
$log(\lambda_{tastefood})$	0.60	0.09
$log(\lambda_{mouthsores})$	0.61	0.09
$log(\lambda_{teeth})$	0.50	-0.14
$log(\mu_{pain})$	-0.13	0.52
$log(\mu_{fatigue})$	-0.28	0.52
$log(\mu_{nause})$	0.02	0.50
$log(\mu_{sleep})$	-0.04	0.44

$\boldsymbol{log(\mu_{distress})}$	-0.01	0.44
$\boldsymbol{log(\mu_{shortbreath})}$	0.16	0.03
$\boldsymbol{log(\mu_{remember})}$	0.12	0.13
$\boldsymbol{log(\mu_{appetite})}$	-0.01	0.45
$\boldsymbol{log(\mu_{drowsy})}$	-0.15	0.48
$\boldsymbol{log(\mu_{drymouth})}$	-0.19	0.53
$\boldsymbol{log(\mu_{sad})}$	0.10	0.39
$\boldsymbol{log(\mu_{vomiting})}$	0.12	0.06
$\boldsymbol{log(\mu_{numbness})}$	0.10	0.21
$\boldsymbol{log(\mu_{mucus})}$	-0.11	0.57
$\boldsymbol{log(\mu_{chewing})}$	-0.07	0.50
$\boldsymbol{log(\mu_{choking})}$	0.08	0.46
$\boldsymbol{log(\mu_{voice})}$	-0.11	0.48
$\boldsymbol{log(\mu_{skinpain})}$	0.09	0.55
$\boldsymbol{log(\mu_{constip})}$	0.16	0.27
$\boldsymbol{log(\mu_{tastefood})}$	-0.13	0.47
$\boldsymbol{log(\mu_{mouthsores})}$	-0.07	0.50
$\boldsymbol{log(\mu_{teeth})}$	0.07	0.37

5. CONCLUSION AND FUTURE RESEARCH

We have developed models using two-state CTMCs to depict longitudinal PRO trajectories and incorporated transition rates in survival modeling and factor analysis. The novelty of our research is that the dynamic longitudinal process has been transformed to cross-sectional transition rates. These rates can be interpreted as the inverse of average durations taken during the transitions from one state to another. Additionally, these transition rates adjusted with baseline covariates will also help to explain the change over time for each individual patient. Specifically, in this research, we built a joint model that combines the longitudinal PRO process in a two-state CTMC with a parametric survival model (Weibull Regression) when the survival time is the outcome of interest. Besides, another semi-parametric modeling involving a two-state CTMC and the Cox PH model was also developed when the risk of a survival event is the primary outcome. Moreover, CTMCs were integrated to factor analysis when the unobserved symptom construct was the main interest such as that in a PRO instrument development process.

In simulation studies, we validated the estimation procedures of our proposed models and demonstrated the application of our models in the setting when a longitudinal process was used to predict time to an event. Specifically, in Section 2, we used a two-state CTMC to describe the longitudinal symptom (cancer fatigue) during the chemotherapy period in lung cancer patients. The transition rates were used as covariates to predict the overall survival in a joint modeling framework. In Section 3, the longitudinal patient-reported fatigue during the radiation/chemoradiation period was then modeled as a CTMC and used to model the risk of head and neck cancer progression in a joint approach. Both applications showed that our proposed models could be applicable in clinical research to address the association between the longitudinal PRO and survival outcomes.

What is more, we further illustrated the usage of CTMC in longitudinal PROs for the method of dimension reduction in Section 3. We built the individual two-state CTMC for each patient to describe their symptoms during treatment and used the generated transition rates for each symptom as summary statistics, which function as the observed variables in factor analysis. The results showed that the transition rates could explain the longitudinal symptom patterns appropriately and provide sufficient information in the factor analysis for the underlying symptom construct.

In this dissertation, the major limitations of our research are discussed. However, some of them can be addressed in future research. First of all, we lost partial information when we dichotomized the longitudinal variable (a continuous or ordinal variable) into a binary variable in the two-state CTMC. Even though we have addressed this issue before (see discussion in Section 2 and Section 3), categorizing more states in the CTMC and turning it to a three-state CTMC or four-state CTMC could mitigate the problem. In fact, if large variation among the states was observed, a three-state CTMC, for example, may be more appropriate to illustrate the longitudinal process. Therefore, a future joint modeling method that combines a three-state CTMC and the survival model may be of interest. Second, we adjusted baseline covariates in the two-state CTMC, but the missing data mechanism of the baseline covariates was not discussed in this research. It is common that a missing clinical variable may occur in real-world data collection; therefore, it is worth exploring further different approaches used for missing data mechanisms and their effect on parameter estimation in the joint modeling. Third, we linked a longitudinal process to a survival outcome in this research; however, a broader joint modeling approach could include two or more longitudinal processes may accommodate the needs of PRO

research when multiple PROs are of interest at the same time. Lastly, we used a two-step backward selection strategy for model selection in the joint model we proposed, but the comprehensive associations among the longitudinal process, baseline covariates, and the survival outcome were not discussed in this research. For example, the PRO (the longitudinal process) may serve as a confounder or serve as a mediator to the survival outcome. Therefore, further research that explicitly studies the association among the longitudinal process, baseline covariates, and survival outcome can help to guide the strategy used in the model selection.

So far, to our knowledge, this is the first study where CTMC was applied to PRO symptom research. Most of the previous studies used a linear mixed model to describe the longitudinal PRO trend, but they were incapable of fitting the data when distinct trajectories were observed. In contrast, CTMC could better depict the longitudinal PRO pattern, and it is also suitable and practical for it to be applied to other statistical models in clinical research.

In summary, we developed statistical methods that model a longitudinal PRO process as a two-state CTMC and integrate it to the survival models or factor analysis. The applications showed that the covariate-adjusted transition rates could summarize longitudinal PROs appropriately and serve as predictors in the survival analysis. Furthermore, we also demonstrated an alternative to reduce the dimension of longitudinal patient-reported symptoms using CTMC. Therefore, with our proposed method, researchers can better understand the unobserved longitudinal symptom construct in a much simpler manner.

6. REFERENCE

1. Rosenthal, D.I., et al., *Patterns of symptom burden during radiotherapy or concurrent chemoradiotherapy for head and neck cancer: A prospective analysis using the University of Texas MD Anderson Cancer Center Symptom Inventory-Head and Neck Module.* Cancer, 2014. **120**(13): p. 1975-1984.
2. Wang, X.S., et al., *Prognostic value of symptom burden for overall survival in patients receiving chemotherapy for advanced nonsmall cell lung cancer.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2010. **116**(1): p. 137-145.
3. Fisher, L.D. and D.Y. Lin, *Time-dependent covariates in the Cox proportional-hazards regression model.* Annual review of public health, 1999. **20**(1): p. 145-157.
4. Hosmer Jr, D.W. and S. Lemeshow, *Applied survival analysis: regression modelling of time to event data (1999).* Eur Orthodontic Soc, 1999: p. 561-2.
5. Tsiatis, A.A. and M. Davidian, *A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error.* Biometrika, 2001. **88**(2): p. 447-458.
6. (EMA), E.M.A., *Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies.*, EMA, Editor. 2016.
7. US Department of Health and Human Services, F.a.D.A., *Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance.* Health and Quality of Life Outcomes, 2006. **4**: p. 1-20.

8. Basch, E. and A.C. Dueck, *Patient-reported outcome measurement in drug discovery: a tool to improve accuracy and completeness of efficacy and safety data.* Expert opinion on drug discovery, 2016. **11**(8): p. 753-758.

9. Aiyegbusi, O.L., et al., *Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol.* BMJ open, 2016. **6**(10): p. e012014.

10. Edbrooke-Childs, J., M. Wolpert, and J. Deighton, *Using patient reported outcome measures to improve service effectiveness (UPROMISE): Training clinicians to use outcome measures in child mental health.* Administration and Policy in Mental Health and Mental Health Services Research, 2016. **43**(3): p. 302-308.

11. Reeve, B.B., et al., *Recommended patient-reported core set of symptoms to measure in adult cancer treatment trials.* JNCI: Journal of the National Cancer Institute, 2014. **106**(7).

12. Williams, L.A., et al., *Measuring symptoms as a critical component of drug development and evaluation in hematological diseases.* Clinical investigation, 2013. **3**(12): p. 1127.

13. Cleeland, C.S., et al., *Integrating pain metrics into oncology clinical trials.* Clinical Cancer Research, 2011. **17**(21): p. 6646-6650.

14. Atkinson, T.M., et al., *The Brief Pain Inventory and its "pain at its worst in the last 24 hours" item: clinical trial endpoint considerations.* Pain Medicine, 2010. **11**(3): p. 337-346.

15. Raskin, J., et al., *A double-blind, randomized multicenter trial comparing duloxetine with placebo in the management of diabetic peripheral neuropathic pain.* Pain Medicine, 2005. **6**(5): p. 346-356.

16. Turk, D.C., et al., *Analyzing multiple endpoints in clinical trials of pain treatments: IMMPACT recommendations.* Pain, 2008. **139**(3): p. 485-493.

17. Basch, E., et al., *Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment.* Jama, 2017. **318**(2): p. 197-198.

18. (EMA), E.M.A. *Integrating patients' views in clinical studies of anticancer medicines*. 2016 [cited 2018 11/12]; Available from: https://www.ema.europa.eu/news/integrating-patients-views-clinical-studies-anticancer-medicines.

19. Deshpande, P.R., et al., *Patient-reported outcomes: a new era in clinical research.* Perspectives in clinical research, 2011. **2**(4): p. 137.

20. US Department of Health and Human Services, F.a.D.A., *Guidance for Industry: Patient-Reported Outcome Measures - Use in Medical Product Development to Support Labeling Claims* 2009.

21. (EORTC), E.O.f.R.a.T.o.C. *Quality of Life Group Website*. 2018 [cited 2018 Nov. 15]; Available from: http://qol.eortc.org/.

22. Aaronson, N.K., et al., *The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology.* JNCI: Journal of the National Cancer Institute, 1993. **85**(5): p. 365-376.

23. Group, E. *EQ-5D*. 2018 [cited 2018 Nov. 15]; Available from: https://euroqol.org/.

24. Janssen, M., et al., *Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study.* Quality of Life Research, 2013. **22**(7): p. 1717-1727.

25. National Cancer Institute, D.o.C.C.P.S., Health Delivery Research Program. *Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE™)*. 2018 Sep. 14, 2018 [cited 2018 Nov. 15]; Available from: https://healthcaredelivery.cancer.gov/pro-ctcae/.

26. Dueck, A.C., et al., *Validity and reliability of the US National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE).* JAMA oncology, 2015. **1**(8): p. 1051-1059.

27. (FACIT.org), F.A.o.C.I.T.O. *Questionnaires*. 2018 Sep. 19, 2016 [cited 2018 Nov. 15]; Available from: http://www.facit.org/FACITOrg/Questionnaires.

28. Cella, D.F., et al., *The Functional Assessment of Cancer Therapy scale: development and validation of the general measure.* J Clin Oncol, 1993. **11**(3): p. 570-579.

29. Care, R.H. *36-Item Short Form Survey (SF-36)*. 2018 [cited 2018 Nov. 15]; Available from: https://www.rand.org/health/surveys_tools/mos/36-item-short-form.html#section-navigation.

30. Ware Jr, J.E. and C.D. Sherbourne, *The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection.* Medical care, 1992: p. 473-483.

31. Center, T.U.o.T.M.A.C. *The MD Anderson Symptom Inventory*. 2018 [cited 2018 Nov. 15]; Available from: https://www.mdanderson.org/research/departments-labs-institutes/departments-divisions/symptom-research/symptom-assessment-tools/md-anderson-symptom-inventory.html.

32. Cleeland, C.S., et al., *Assessing symptom distress in cancer patients: the MD Anderson Symptom Inventory.* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2000. **89**(7): p. 1634-1646.

33. Scher, H.I., et al., *Trial design and objectives for castration-resistant prostate cancer: updated recommendations from the Prostate Cancer Clinical Trials Working Group 3.* Journal of Clinical Oncology, 2016. **34**(12): p. 1402.

34. Wu, Y.-L., et al., *Does EGFR Mutation Type Influence Patient-Reported Outcomes in Patients with Advanced EGFR Mutation-Positive Non-Small-Cell Lung Cancer? Analysis of Two Large, Phase III Studies Comparing Afatinib with Chemotherapy (LUX-Lung 3 and LUX-Lung 6).* The Patient-Patient-Centered Outcomes Research, 2018. **11**(1): p. 131-141.

35. Donovan, J.L., et al., *Patient-reported outcomes after monitoring, surgery, or radiotherapy for prostate cancer.* New England Journal of Medicine, 2016. **375**(15): p. 1425-1437.

36. Hassett, M.J., et al., *High-priority topics for cancer quality measure development: Results of the 2012 American Society of Clinical Oncology Collaborative Cancer Measure Summit.* Journal of oncology practice, 2014. **10**(3): p. e160-e166.

37. Wang, X.S., et al., *Longitudinal study of the relationship between chemoradiation therapy for non–small-cell lung cancer and patient symptoms.* Journal of clinical oncology, 2006. **24**(27): p. 4485-4491.

38. Elting, L.S., et al., *Patient-reported measurements of oral mucositis in head and neck cancer patients treated with radiotherapy with or without chemotherapy: demonstration of increased frequency, severity, resistance to palliation, and impact on quality of life.* Cancer, 2008. **113**(10): p. 2704-2713.

39. Fairclough, D.L., *Design and analysis of quality of life studies in clinical trials*. 2010: Chapman and Hall/CRC.

40. Lin, E.-Y., et al., *Trajectory of health-related quality of life and its determinants in patients who underwent lumbar spine surgery: a 1-year longitudinal study.* Quality of Life Research, 2018: p. 1-9.

41. Shi, Q., et al., *Using group-based trajectory modeling to examine heterogeneity of symptom burden in patients with head and neck cancer undergoing aggressive non-surgical therapy.* Quality of Life Research, 2013. **22**(9): p. 2331-2339.

42. Quinten, C., et al., *Patient self-reports of symptoms and clinician ratings as predictors of overall cancer survival.* Journal of the National Cancer Institute, 2011. **103**(24): p. 1851-1858.

43. Armstrong, T.S., et al., *Net clinical benefit analysis of radiation therapy oncology group 0525: a phase III trial comparing conventional adjuvant temozolomide with dose-intensive temozolomide in patients with newly diagnosed glioblastoma.* Journal of Clinical Oncology, 2013. **31**(32): p. 4076.

44. Diggle, P., et al., *Analysis of longitudinal data*. 2002: Oxford University Press.

45. Klein, J.P. and M.L. Moeschberger, *Survival analysis: techniques for censored and truncated data*. 2006: Springer Science & Business Media.

46. Molenberghs, G. and G. Verbeke, *Longitudinal data analysis.* Wiley StatsRef: Statistics Reference Online, 2014: p. 1-28.

47. Hickey, G.L., et al., *Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues.* BMC medical research methodology, 2016. **16**(1): p. 117.

48. Tsiatis, A.A. and M. Davidian, *Joint modeling of longitudinal and time-to-event data: an overview.* Statistica Sinica, 2004: p. 809-834.

49. Cox, D.R., *Regression models and life-tables (with discussion).* The Journal of the Royal Statistical Society, 1972. **B 34**: p. 182-220.

50. Wang, C., N. Wang, and S. Wang, *Regression analysis when covariates are regression parameters of a random effects model for observed longitudinal measurements.* Biometrics, 2000. **56**(2): p. 487-495.

51. Wang, C. and Y. Huang, *Functional methods for logistic regression on random-effect-coefficients for longitudinal measurements.* Statistics & probability letters, 2001. **53**(4): p. 347-356.

52. Li, E., D. Zhang, and M. Davidian, *Conditional estimation for generalized linear models when covariates are subject-specific parameters in a mixed model for longitudinal measurements.* Biometrics, 2004. **60**(1): p. 1-7.

53. Wang, J.-L., J.-M. Chiou, and H.-G. Müller, *Functional data analysis.* Annual Review of Statistics and Its Application, 2016. **3**: p. 257-295.

54. Ferraty, F. and P. Vieu, *Nonparametric functional data analysis: theory and practice*. 2006: Springer Science & Business Media.

55. Ramsay, J., *Functional data analysis.* Encyclopedia of Statistics in Behavioral Science, 2005.

56. Ramsay, J.O. and B.W. Silverman, *Applied functional data analysis: methods and case studies*. 2007: Springer.

57. Sukkar, R., et al. *Disease progression modeling using hidden Markov models*. in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. 2012. IEEE.

58. Yesavage, J.A., et al., *Modeling the prevalence and incidence of Alzheimer's disease and mild cognitive impairment.* Journal of psychiatric research, 2002. **36**(5): p. 281-286.

59. Costin, H. and O. Geman, *Parkinson's disease prediction based on multistate markov models.* International Journal of Computers Communications & Control, 2013. **8**(4): p. 525-537.

60. Faglioni, P., et al., *Parkinson's disease affects automatic and spares intentional verbal learning a stochastic approach to explicit learning processes.* Cortex, 1995. **31**(4): p. 597-617.

61. Karlsson, M.O., et al., *A pharmacodynamic Markov mixed-effect model for the effect of temazepam on sleep.* Clinical Pharmacology & Therapeutics, 2000. **68**(2): p. 175-188.

62. Ho, C.-H., *Logistic regression with Markov chains as covariates*. 2011: The University of Texas School of Public Health.

63. Rubin, M.L., et al., *A joint logistic regression and covariate-adjusted continuous-time Markov chain model.* Statistics in medicine, 2017. **36**(28): p. 4570-4582.

64. Basch, E., A.M. Deal, and A.C. Dueck, *Patient-Reported Symptom Monitoring During Chemotherapy—Reply.* Jama, 2017. **318**(19): p. 1935-1936.