Applications of Stochastic Modeling and Data Analytics Techniques in Healthcare Decision Making

by

Ozden Onur Dalgic

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2017

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Mehmet A. Begen Associate Professor
Supervisor(s):	Fatih Safa Erenay Assistant Professor Osman Y. Ozaltin Assistant Professor
Internal Member:	Sibel Alumur Alev Assistant Professor
Internal Member:	James H. Bookbinder Professor
Internal-External Member:	Catherine Burns Professor

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

We present approaches utilizing aspects of data analytics and stochastic modeling techniques and applied to various areas in healthcare. In general, the thesis has composed of three major components.

Firtsly, we propose a comparison analysis between two of the very well-known infectious disease modeling techniques to derive effective vaccine allocation strategies. This study, has emerged from the fact that individuals are prioritized based on their risk profiles when allocating limited vaccine stocks during an influenza pandemic. Computationally expensive but realistic agent-based simulations and fast but stylized compartmental models are typically used to derive effective vaccine allocation strategies. A detailed comparison of these two approaches, however, is often omitted. We derive age-specific vaccine allocation strategies to mitigate a pandemic influenza outbreak in Seattle by applying derivative-free optimization to an agent-based simulation and also to a compartmental model. We compare the strategies derived by these two approaches under various infection aggressiveness and vaccine coverage scenarios. We observe that both approaches primarily vaccinate school children, however they may allocate the remaining vaccines in different ways. The vaccine allocation strategies derived by using the agent-based simulation are associated with up to 70% decrease in total cost and 34% reduction in the number of infections compared to the strategies derived by the compartmental model. Nevertheless, the latter approach may still be competitive for very low and/or very high infection aggressiveness. Our results provide insights about the possible differences between the vaccine allocation strategies derived by using agent-based simulations and those derived by using compartmental models.

Secondly, we introduce a novel and holistic scheme to capture the gradual amyotrophic lateral sclerosis progression based on the critical events referred as tollgates. Amyotrophic lateral sclerosis is neuro-degenerative and terminal disease. Patients with amyotrophic lateral sclerosis lose control of voluntary movements over time due to continuous degeneration of motor neurons. Using a comprehensive longitudinal dataset from Mayo Clinic's ALS Clinic in Rochester, MN, we characterize the progression through tollgates at the body segment (e.g., arm, leg, speech, swallowing, breathing) and patient levels over time. We describe how the progression based on the followed tollgate pathways varies among patients

and ultimately, how this type of progression characterization may be utilized for further studies. Kaplan-Meier analysis are conducted to derive the probability of passing each tollgate over time. We observe that, in each body segment, the majority of the patients have their abilities affected or worse (Level1) at the first visit. Especially, the proportion of patients at higher tollgate levels is larger for arm and leg segments compared to others. For each segment, we derive the over-time progression pathways of patients in terms of the reached tollgates. Tollgates towards later visits show a great diversity among patients who were at the same tollgate level at the first clinic visit. The proposed tollgate mechanism well captures the variability among patients and the history plays a role on when patients reach tollgates. We suggest that further and comprehensive studies should be conducted to observe the whole effect of the history in the future progression.

Thirdly, based on the fact that many available databases may not have detailed medical records to derive the necessary data, we propose a classification-based approach to estimate the tollgate data using ALSFRS-R scores which are available in most databases. We observed that tollgates are significantly associated with the ALSFRS-R scores. Multi-class classification techniques are commonly used in such problem; however, traditional classification techniques are not applicable to the problem of finding the tollgates due to the constraint of that a patients' tollgates under a specific segment for multiple visit should be non-decreasing over time. Therefore, we propose two approaches to achieve a multi-class estimation in a non-decreasing manner given a classification method. While the first approach fixes the class estimates of observation in a sequential manner, the second approach utilizes a mixed integer programming model to estimate all the classes of a patients' observations. We used five different multi-class classification techniques to be employed by both of the above implementations. Thus, we investigate the performance of classification model employed under both approaches for each body segment.

Acknowledgements

Firstly, I owe my deepest gratitude to my supervisors, Dr. Fatih Safa Erenay who was always there for me throughout my doctoral educations and Dr. Osman Ozaltin, for all his encouragement and comprehensive advice.

The courses I have taken at the University of Waterloo have been instrumental in my development and in shaping this thesis. I would like to thank the instructors of these courses for making hard concepts easy to understand and introducing me to new topics and ideas: Samir Elhedhli, Hossein Abouee Mehrizi, Qi-ming he, Bon Koo, Levent Tuncel.

I also would like to thank my group-mates who helped in challenging my ideas and generating new ones. The long and frequent talks with Tagi Khaniyev and Abdelhalim Hiassat have sharpened my understanding and enhanced my knowledge. I am thankful for the joyful times and the anxiety about future plans our meetings have brought.

I am also very thankful to my family, especially to my wife Gizem Sultan Nemutlu. Being colleagues at university and partners at home are meant to be different and she achieved both of them in the most pleasant way.

Dedication

Didn't even know that so much love existed... to my handsome boy and beautiful wife.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ALS	Amyothropich lateral sclerosis
ALSFRS-R	Revised ALS Functional Rating Scale
BiPAP	Bilevel Positive Airway Pressure
FDA	Foods and Drug Administration
MADS	Mesh adaptive direct search
MDP	Markov decision process
MLR	Multinomial logistic regression
NOMAD	Non-linear optimization with MADS
OR	Operations Research
PRO-ACT	Pooled Resource Open-Access ALS Clinical Trials
RF	Random forests
RFECV	Recursive feature elimination with cross validation
SVC	Support vector classifier

List of Symbols

Chapter 2

λ_{ij}	The incidence rate of new infections in age group i caused by infectious individuals in age group j
N	Total population size
N_j	The population size of age group j
Φ_{ij}	The contact rate from age group i to j
β_j^u (β_j^v)	denotes the transmission rate of unvaccinated (vaccinated) infectious individuals in age group j
S	Susceptible compartment
I	Infectious compartment
E	Exposed compartment
R	Recovered compartment
D	Dead compartment
p_i	The vaccinated proportion of age group i
ϵ_i^h	The reduction in the likelihood of becoming infected after vaccination
ω_i^h	The proportion of exposed individuals who eventually become symptomatic
τ_i^h	Exposed-to-Infected transition rate
γ_i^h	Exposed-to-Recovered transition rate
ξ_i^h	Infected-to-Recovered transition rate
χ_i^h	Infected-to-Death transition rate
V	Total number of vaccines
c_i^b	Vaccination cost
c_i^u (c_i^v)	Infection cost of unvaccinated (vaccinated) individuals
c_i^d	Mortality cost

Chapter 4

OA	Overall accuracy
MSE	Mean squared error
SS_{tot}	Total some of squares of classes
SS_{est}	Total some of squares of residuals
X	Features
y	Classes
D_{tl}	The probability of the class being l, given the features at visit t
T	Total number of observations
L	Total number of classes
R	Recovered compartment
u_{tl}	1, if $l \in \{1, 2, \ldots, L\}$ is the selected class of observation $t \in \{1, 2, \ldots, T\}$; 0, otherwise
$\beta_{i,l}$	Coefficient parameters

Chapter 1

Introduction

This thesis focuses on the use of stochastic modeling and data analytics techniques in several healthcare applications. This chapter discusses a brief review of these techniques and their studied applications in the current healthcare literature. We also present backgrounds on the healthcare problems we have studied in detail. Finally, the the outline of thesis is presented.

1.1 Preface

Healthcare is a perpetually expanding industry and operations research (operations research (OR)) has been used for providing effective solutions for healthcare issues for many years (Dobrzykowski et al., 2014). Due to growing health expenditures and increase in demand for quality of life, healthcare field arouses many OR problems with distinctive characteristics (Eveborn et al., 2006; Wang, 2009). Although problem structures are similar to those in other industries, searching for high-quality-of-life resolutions differentiates OR problems in healthcare from other cost-oriented OR problems (Kaplan and Bush, 1982). Rais and Viana (2011) provided a detailed classification of OR problems in healthcare:

- healthcare planning (demand forecasting, location selection, capacity planning),

- healthcare management and logistics (patient scheduling, resource scheduling, logistics),
- healthcare practice (disease diagnosis, treatment planning),
- specialized and preventive healthcare (organ donation and transplant, prevention of diseases).

Our studies in this thesis can be categorized under "disease diagnosis", "treatment planning", and "prevention of diseases".

Many OR tools are employed in modeling and solving challenging problems in healthcare (Brandeau et al., 2004). Simplifying assumptions enable deterministic models applicable for OR problems in healthcare despite their stochastic nature (Harper et al., 2005). Deterministic models provide great insights and derive effective solutions. However, methodologies with stochastic capabilities become popular due to their accuracy in representing the real life and thanks to ever-increasing computational power. In particular, queuing models have been employed many times in healthcare literature (Preater, 2002; Fomundam and Herrmann, 2007; Lakshmi and Iyer, 2013). They usually require little data and are easy to use. On the other hand, they require analytic approaches; therefore, are suitable for systems that have relatively low or medium complexity. Simulation is another very prevalent tool in healthcare applications (Jacobson et al., 2006; Robinson et al., 2012). Its capability for modeling highly complex systems grants the opportunity to design real-life-like models. Thus, researchers benefit from easy policy evaluation in a realistic environment (Robinson, 2002). However, due to vast amount of computation needs, simulations are usually not preferable in optimization applications. Moreover, they usually require large amounts of data (Caro et al., 2010). Reviews of simulation literature in healthcare field are presented in Fone et al. (2003); Brailsford et al. (2009); Günal and Pidd (2010); Katsaliaki and Mustafee (2011); and Sobolev et al. (2011).

On the decision analysis side, Markov decision process (MDP) models draw a growing attention (Ayer et al., 2012; Alagoz et al., 2015; Yaesoubi and Cohen, 2011). Schaefer et al. (2005) presented a broad review of MDP models in healthcare. MDP models are very advantageous for healthcare problems which require dynamic decisions under uncertainty.

Treatment planning, organ donation and transplant, and dynamic resource allocation are major healthcare problems employing MDP models. As it is the case in many other fields, data analytics applications are also frequently highlighted in healthcare. Simpao et al. (2014); Raghupathi and Raghupathi (2014); Srinivasan and Arunasalam (2013) reported the necessity, importance, and opportunities of data analytics in medical fields. Due to the trend in digitization of health records, more and more health data becomes available for healthcare researchers (Raghupathi and Raghupathi, 2014). Potential benefits including early detection of the disease, accurate disease progression patterns, effective health demand forecast may improve patients' life quality and reduce health expenditures substantially (Raghupathi and Raghupathi, 2014).

1.2 Influenza, infectious disease modeling, and vaccination

Influenza is a highly contagious viral disease. Each year a large portion of the worlds' population is infected with influenza resulting in 3-5 million severe cases and 250,000-500,000 deaths (CDC, 2010). Together with pneumonia, influenza is the seventh leading cause of death in the U.S. (CDC, 2014). Considering all the direct and indirect effects such as the cost of medical treatments, working day losses etc., Molinari et al. (2007) estimated the annual burden of influenza epidemics to the U.S. economy as $87.1 billion (95% CI: $47.2-$149.5) in 2003.

An emerging virus that spreads globally may lead to an influenza pandemic (Parvin et al., 1986). Unlike seasonal epidemics, pandemics occur abruptly and cause horrendous death tolls, e.g., the 1918 Spanish influenza pandemic infected around 500 million and killed approximately 50 million people worldwide Taubenberger and Morens (2006). Among pandemic mitigation interventions (e.g., social distancing, public health measures, antiviral prophylaxis), vaccination provides the most efficient and durable response (Chao et al., 2010; Talbot et al., 2013). However, the amount of influenza vaccines during a pandemic is limited due to production restrictions. Thus, the population is prioritized based on risk-factors related to influenza exposure and transmissibility when distributing the available

doses Nichol and Treanor (2006).

To study infection dynamics and evaluate effects of different public health interventions, several studies in the literature proposed compartmental models (Hethcote, 2000; Zaric and Brandeau, 2001; Dasbach et al., 2006; Huang and Li, 2009; Eames et al., 2012), agent-based models (Das et al., 2008; Ventresca and Aleman, 2013; Andradóttir et al., 2014), and discrete-time generation models. We review studies that are closely related to our paper and refer the reader to Anderson et al. (1991) and Jackson et al. (2014) for more general reviews and to Anderson et al. (1991) for a review of earlier studies. In addition, several issues related to control of influenza pandemics/epidemics such as vaccine composition selection Wu et al. (2005); Özaltin et al. (2011), vaccine supply chain design Chick et al. (2008); Deo and Corbett (2009), and optimal allocation of influenza intervention resources are studied in the OR literature.

Medlock and Galvani (2009) considered an influenza pandemic in the U.S., and proposed a deterministic transition model with Susceptible, Exposed, Infectious and Recovered compartments (SEIR model). They found the optimal age-specific vaccine allocation with respect to five performance measures: number of infections, number of deaths, years of life lost, contingent valuation, and total cost. They parametrized their model with survey-based contact data and mortality data from 1918 A (H1N1) and 1957 A (H2N2) influenza pandemics. Medlock and Galvani (2009) suggested prioritizing 5-19 and 30-39 age groups. as 5-19 year-old school children can transmit the disease to their 30-39 year-old parents serving as a bridge to the rest of the population. Medlock et al. (2009) extended the analysis of Medlock and Galvani (2009) by incorporating the vaccine availability schedule. Araz et al. (2012) used a compartmental model of an influenza pandemic to find effective strategies for allocating vaccines to different communities characterized by geographic region and age group. They evaluated four allocation strategies where vaccines become available according to particular schedules. Their analyses showed that when vaccines are delayed until later during a pandemic, communities that are expected to be worst affected by the latest waves of the outbreak should be prioritized. However, allocating the vaccines among different communities proportional to their population size yields minimum waiting time for vaccination. Das et al. (2008) developed a network-based simulation model to evaluate the effect of vaccination, prophylaxis, hospitalization, and social distancing for an

influenza pandemic in a synthetic population with heterogeneous mixing groups and daily schedules for inhabitants.

Uribe-Sánchez et al. (2011) proposed a simulation-based predictive method for dynamic mitigation of influenza pandemics affecting multiple regions. Their approach progressively allocates a limited budget to procure vaccines, antivirals, administration capacity and resources for enforcing social distancing. They considered a synthetic outbreak involving over four million people, and presented a sensitivity analysis to estimate the impact of changes in the budget. Their numerical experiments showed that the marginal effect of vaccine efficacy in a pandemic with a low transmission rate is less significant than in a pandemic with high transmission rate. Patel et al. (2005) built an agent-based simulation of an influenza pandemic, and used a meta-heuristic to find effective age-specific vaccine allocation strategies with respect to the number of infections and deaths. Their numerical experiments showed that the proposed vaccine allocation strategies are up to 84% more effective than randomly vaccinating the population.

In a closely related study, Yaesoubi and Cohen (2011) formulated a Markov decision process (MDP) model, and derived dynamic intervention policies to mitigate an influenza pandemic in a small and closed population. Parvin et al. (2012) also formulated an MDP model to control an infection, and derived asymptotically optimal solutions based on fluid approximation. Finally, Yarmand et al. (2014) formulated a two-stage stochastic linear programming model to minimize the expected cost of vaccination for containing the infection. In the first stage, the model determines the proportion of the vaccines allocated to different regions before the influenza season begins; while the second stage determines how the remaining vaccines are allocated after the onset of the influenza season. They calibrated the model for a hundred counties in North Carolina, and showed that their proposed two-stage vaccination policy potentially results in lower attack rates as well as lower vaccine procurement and administration costs. Rahmandad and Sterman (2008) compared compartmental models and agent-based simulation models under different network structures and heterogeneity conditions. They analyzed differences in disease propagation patterns generated by these two types of models.

1.3 Natural history of amyotrophic lateral sclerosis

Amyothropich lateral sclerosis (ALS) is a disease which causes the degeneration of motor neurons reaching from the brain to the spinal cord (upper motor neurons) and from the spinal cord to the muscles (lower motor neurons) (Rowland and Shneider, 2001). The perpetual degeneration leads to the loss of the motor neurons; thereby, the loss of brain function on voluntary muscle movements (National Institute of Neurological Disorders and Stroke, 2015). As a result, muscles controlling the voluntary movement weaken and atrophy. Patients may suffer from losing their ability to speak, eat, move and breathe. In 90–95 percent of all cases in the U.S., ASL is sporadic; while the remaining 5–10 percent is familial (inherited) ALS (Kiernan et al., 2011). Families carrying the disease have 50 percent chance that each offspring will inherit the gene developing the disease. There is no known cure or treatment for its diminution or termination (National Institute of Neurological Disorders and Stroke, 2015). Patients have mean survival duration of 2–4 years after the onset of the disease (Hobson and McDermott, 2016). An FDA approved medication called *Riluzole* may extend patients survival duration by several months (Miller et al., 2012). Other major interventions, including *endoscopic gastrostomy*, *ventilation support*, *recombinant human insulin-like growth factor 1 treatment*, are known to modestly increase the survival duration and improve the quality of life (Miller et al., 2012). Although the world wide incidence rate has not been estimated yet the disease affects about two people per 100 000 each year in Europe and the United States (Cronin et al., 2007; Kiernan et al., 2011). In the U.S., around 5 600 people are diagnosed with ALS each year and approximately 30 000 patients have the disease in the population (National Institute of Neurological Disorders and Stroke, 2015). Due to patients' constant need for caregivers, expensive interventions, and hospital care, the financial burden of ALS disease extends to $1.5 million per patient in the total disease duration (Obermann and Lyon, 2015).

ALS is a progressive disease and eventually affects all voluntary movements. However, disease progression may differ among patients. A salient distinction among patients is the disease onset site. The disease is mainly categorized as either bulbar onset or limb onset (Magnus et al., 2002). Bulbar-onset ALS starts with degenerating the motor neurons in the brain stem. Primarily, muscles used in speech, swallowing, and/or chewing become

weak by the progress of the degeneration (Kiernan et al., 2011). On the other hand, limb-onset ALS affects arms and legs of patients at early stages of the disease (Kiernan et al., 2011). Muscles controlling the limbs become impaired. Arms and/or legs may display poor performance resulting in having difficulties in daily tasks such as walking, buttoning a shirt, or reaching a shelf (Boillée et al., 2006).

Although patients' level of impairment cannot be truly quantified, an instrument called *ALS functional rating scale (ALSFRS)* provides a detailed assessment for the impacts on patients (Brooks et al., 1996). ALSFRS is a questionnaire-based scale that monitors the physical condition in performing 10 functionality items of daily tasks. Each functionality item is evaluated with a five-level scale (0: not able, 4: normal) making a total score between 0 (severe impairment) and 40 (normal capability). Cedarbaum et al. (1999) referred to a disproportionate weighting to limbs and bulbar functions compared to respiratory function in ALSFRS and proposed *revised ALS functional rating scale (ALSFRS-R)*. ALSFRS-R replaces breathing item in ALSFRS with three other items. Table 1.1 shows the functionality items in ALSFRS and ALSFRS-R. Both scales are validated to predict survival, whereas ALSFRS-R is a better survival predictor due to its broader evaluation of respiratory symptoms (Cedarbaum et al., 1999). Simon et al. (2014) provided a systematic comparison for the instruments which quantify the clinical progression of ALS.

As in many fatal diseases, determining prognostic factors plays an important role in ALS patients' anticipation for achieving a decent disease period and death. Therefore, making accurate predictions on the disease progression or survival time imposes a big burden on clinicians. Küffner et al. (2015) remarked on the necessity and the importance of different methodologies used in predicting the progression of the ALS disease. Since the disease progression in ALS displays a huge variability among patients, accuracy of the estimators is crucial to identify the next sites of disease spread and survival time.

1.4 Thesis outline

In this thesis, our main purpose is to combine analytics and stochastic modeling techniques to derive effective predictions and applicable solutions for important healthcare problems.

Table 1.1: **Functionality items in ALSFRS and ALSFRS-R.**

	Item	ALSFRS	ALSFRS-R
1	Speech	√	√
2	Salivation	√	√
3	Swallowing	√	√
4	Handwriting	√	√
5	Cutting food and handling utensils	√	√
6	Dressing and hygiene	√	√
7	Turning in bed and adjusting bed clothes	√	√
8	Walking	√	√
9	Climbing stairs	√	√
10	Breathing	√	–
11	Dyspnea	–	√
12	Orthopnea	–	√
13	Respiratory insufficiency	–	√

Our studies shine a light on the possible improvements of applying analytics and stochastic modeling techniques in the medical field. In particular, we study three problems, *i) deriving effective vaccine allocation strategies for pandemic influenza*, *ii) Tollgate-based progression pathways of ALS patients*, and *iii) Deriving tollgates from ALSFRS-R scores using multi-class classifiers*.

In the first problem, we study age-specific vaccine allocation strategies in a pandemic influenza outbreak. We apply derivative-free optimization to an agent-based simulation and a compartmental model to obtain effective vaccine allocation strategies. This study has been published in the journal, PLOS One (Dalgıç et al., 2017)

In the second problem, we propose a novel holistic mechanism to monitor the progression of patients with amyotrophic lateral sclerosis in terms of the critical events referred to as tollgates. We perform an extensive data analysis on data from the Mayo clinic, Rochester, MN. to illustrate how patients progress based the proposed mechanism estimate. Moreover, we estimate the likelihood of the critical events over time.

In the third problem, considering that deriving the tollgates for the existing databases is not possible due to lack of detailed medical records, we propose a mapping approach from

the ALSFRS-R scores to the tollgates with the help of various classification techniques.

Problem definitions, objectives, literature reviews, solution approaches and related analyses for the first and second problem are briefly discussed in Chapter 2, Chapter 3, and Chapter 4, respectively.

Chapter 2

Deriving Effective Vaccine Allocation Strategies for Pandemic Influenza

In the literature, compartmental models (Hethcote, 2000; Zaric and Brandeau, 2001; Dasbach et al., 2006; Huang and Li, 2009; Medlock and Galvani, 2009; Medlock et al., 2009; Eames et al., 2012; Araz et al., 2012) and agent-based simulations (Patel et al., 2005; Das et al., 2008; Basta et al., 2009; Lee et al., 2010; Uribe-Sánchez et al., 2011; Andradóttir et al., 2014) are frequently employed to make mitigation plans for influenza pandemics and evaluate the effectiveness of various public health interventions (Dhamodharan and Proano, 2012; Feng et al., 2013; Chhatwal and He, 2015). Although it is known that the infection propagation is different in these two approaches (Rahmandad and Sterman, 2008), a detailed comparison of the strategies derived by using them is often omitted. On the one hand, compartmental models represent the number of individuals in each stage (or compartment) of the epidemic (e.g., susceptible, exposed, infected, recovered) by continuous-time state variables, and formulate the transitions among different compartments using differential equations. These models can rapidly evaluate many scenarios and intervention strategies, but they assume that individuals in each compartment mix uniformly and randomly with each other. Moreover, deterministic compartmental models do not consider the uncertainties in disease propagation (e.g., stochasticity in transmission events, incubation, and recovery periods) (Andradóttir et al., 2014). Thus, such models

may not accurately model infection dynamics, especially at the initial and final stages of a pandemic when few infectious individuals exist (Germann et al., 2006). Despite their simplifying assumptions, compartmental models have proved to be predictive (Anderson et al., 1991; Mollison et al., 1994; Bansal et al., 2007), and they have been successfully extended to capture large-scale host heterogeneities. These extensions of the simple compartmental framework include age-specific contact patterns (Medlock and Galvani, 2009) and heterogeneities induced by spatial structure (Lloyd and May, 1996). Agent-based simulations, on the other hand, consider uncertainties about the infection parameters; and they store individual-level information to model contact patterns in a population at the expense of increased computational burden (Ferguson et al., 2006; Merler et al., 2009; Hladish et al., 2012).

Our goal in this paper is to identify possible scenarios under which performances of the effective age-specific vaccine allocation strategies derived by using compartmental models and agent-based simulations may differ significantly in practical settings. For this purpose, we consider an influenza pandemic in Seattle using a custom-built deterministic compartmental model and an agent-based simulation developed by Chao et al. (2010). The compartmental model is calibrated to closely approximate the results of the agent-based simulation under no vaccination. We apply mesh-adaptive optimization to derive effective age-specific vaccine allocation strategies based on four different objective functions. At each iteration of the optimization process, the performances of the newly generated vaccine allocation strategies are evaluated using the agent-based simulation in one set of experiments, and using the compartmental model in the other set of experiments. We perform sensitivity analysis to identify potential distinctions between these two approaches under different vaccine coverage and infection aggressiveness scenarios.

We observe that age-specific vaccine allocation strategies derived by using computationally expensive but more realistic agent-based simulation and those derived by using fast but more stylized compartmental model are different, although both models are calibrated to generate similar results under no vaccination. We use the agent-based simulation to evaluate the performances of strategies derived by the compartmental model. Our results show that the vaccine allocation strategies derived by the agent-based simulation are associated with up to 70% decrease in total cost and 34% reduction in the number of infections

compared to the strategies derived by the compartmental model. Nevertheless, the latter approach may still be competitive for very low and/or very high infection aggressiveness scenarios.

It is clear that any two infectious disease spread models can differ from each other with respect to the assumptions in their design and parametrization. Furthermore, the degree to which each modeling approach allows for inclusion of heterogeneity and uncertainty varies. Therefore, the empirical comparison presented in this study is valid for the considered agent-based influenza pandemic simulation that is well-known and commonly used in the literature (Chao et al., 2010). Our results, however, still provide important insights into the possible differences between the vaccine allocation strategies derived by agent-based simulations and deterministic compartmental models.

2.1 Materials and methods

We consider different levels of vaccine coverage and infection aggressiveness, and apply mesh-adaptive optimization (Audet et al., 2009) to find effective strategies for allocating limited vaccine doses to different age-groups in the population with respect to four different performance measures: total cost, number of deaths, number of infections, and years of life lost. We evaluate the performances of the trial vaccine allocation strategies at each iteration of the optimization algorithm using an agent-based simulation in one set of experiments, and a compartmental model in the other set of experiments.

2.1.1 Agent-based simulation

We employ FluTe, an open-source and validated agent-based pandemic influenza simulation developed by Chao et al. (2010). FluTe's contact network is composed of census tracts divided into communities of 500-3,000 individuals. Each community consists of randomly generated households of 1-7 individuals in one of the five age-groups: preschool children (0-4), school children (5-18), young adults (19-29), adults (30-64), and seniors (65 and over). Individuals can be members of multiple community-based mixing groups such as

households, household clusters (composed of socially close households), neighborhoods and communities. The simulation has two time epochs for each day: day- and night-time. At night time, individuals can only make contacts within their community-based mixing groups, whereas they may contact other individuals in day-time if they share the same social mixing group, e.g., daycare, school, workplace.

In each time epoch, a contact for potential disease transmission between any two individuals sharing a mixing group is generated. During a contact between a susceptible and an infectious individual, influenza transmission may occur with a probability that depends on vaccine efficacy, virus load and symptoms of the infectious individual. Each infected individual follows a predefined daily viral load profile representing the level of infectiousness on each day of the disease duration. Infected individuals may become symptomatic after an asymptomatic incubation period of one to three days (Carrat et al., 2008). Symptomatic individuals are twice as infectious as the asymptomatic ones. Infected individuals recover and become immune after six days.

The vaccinated individuals have reduced likelihood of getting infected during a contact, becoming symptomatic when infected, and transmitting the disease (Struchiner et al., 2010). The vaccine efficacy reaches its maximum level in two weeks with exponential increments after the vaccination, and the maximum vaccine efficacy varies among the age groups. Due to incremental nature of vaccine efficacy, the timing of vaccine interventions affects success in containing influenza pandemics. FluTe allows administrating vaccines before (pre-vaccination) or after (reactive vaccination) the onset of the pandemic. We refer the reader to Chao et al. (2010) for further details about FluTe.

2.1.2 Compartmental model

We propose a deterministic compartmental model that closely approximates the results of FluTe for Seattle under no vaccination. Similar to FluTe, we divide the population into five age groups, $AG = \{$preschool children (0-4), school children (5-18), young adults (19-29), adults (30-64), seniors (65+)$\}$. Each age group $i \in AG$ includes vaccinated and unvaccinated individuals in five compartments: susceptible (S), exposed (E), infected (I), recovered (R), and dead (D). We denote the susceptible individuals in age group $i \in AG$

with vaccination status $h \in H = \{$(u)nvaccinated, (v)accinated$\}$ by S_i^h, for example. The exposed compartment (E) corresponds to the asymptomatic individuals in FluTe. In addition, the infected compartment (I) corresponds to the symptomatic individuals in FluTe. A proportion of asymptomatic individuals never develop symptoms in FluTe, therefore, we split the exposed compartment into two sub-compartments: those who eventually show disease symptoms ($E_{\rightarrow}I$), and those who recover without showing symptoms ($E_{\rightarrow}R$). The symptomatic individuals are twice as infectious as asymptomatic ones in both FluTe and the SEIR model.

The incidence rate of new infections in age group i caused by infectious individuals in age group j, denoted by λ_{ij}, is given by:

$$\lambda_{ij} = \frac{\Phi_{ij}(\beta_j^u(I_j^u + E_j^u/2) + \beta_j^v(I_j^v + E_j^v/2))}{N_j}. \tag{2.1}$$

In Eq (2.1), Φ_{ij} is the contact rate from age group i to j. Parameter β_j^u (β_j^v) denotes the transmission rate of unvaccinated (vaccinated) infectious individuals in age group j given a single contact with a susceptible individual. Variables I_j^u (I_j^v) and E_j^u (E_j^v) represent the number of unvaccinated (vaccinated) infected and exposed individuals in age group j, respectively. Note that $E_j^u = (E_{\rightarrow}I)_j^u + (E_{\rightarrow}R)_j^u$ and $E_j^v = (E_{\rightarrow}I)_j^v + (E_{\rightarrow}R)_j^v$. Finally, N_j is the size of age group j, and $\sum_{j \in AG} N_j = N$ where N denotes the total population size. The overall infection rate of individuals in age group i is equal to $\lambda_i = \sum_{j \in AG} \lambda_{ij}$. Note that vaccinated ($v$) and unvaccinated ($u$) compartments are interdependent because the infection rate λ_i depends on the number of vaccinated and unvaccinated infectious individuals. Figure 2.1 depicts the transitions among the compartments, and the model

equations are given by:

$$\frac{dS_i^h}{dt} = -(1-\epsilon_i^h)\lambda_i S_i^h \qquad h \in H, i \in AG \tag{2.2a}$$

$$\frac{d(E_{\to}I)_i^h}{dt} = (1-\epsilon_i^h)\lambda_i \omega_i^h S_i^h - \tau_i^h (E_{\to}I)_i^h \qquad h \in H, i \in AG \tag{2.2b}$$

$$\frac{d(E_{\to}R)_i^h}{dt} = (1-\epsilon_i^h)\lambda_i (1-\omega_i^h) S_i^h - \gamma_i^h (E_{\to}R)_i^h \qquad h \in H, i \in AG \tag{2.2c}$$

$$\frac{dI_i^h}{dt} = \tau_i^h (E_{\to}I)_i^h - (\xi_i^h + \chi_i^h) I_i^h \qquad h \in H, i \in AG \tag{2.2d}$$

$$\frac{dR_i^h}{dt} = \xi_i^h I_i^h + \gamma_i^h (E_{\to}R)_i^h \qquad h \in H, i \in AG \tag{2.2e}$$

$$\frac{dD_i}{dt} = \sum_{h \in H} \chi_i^h I_i^h \qquad i \in AG \tag{2.2f}$$

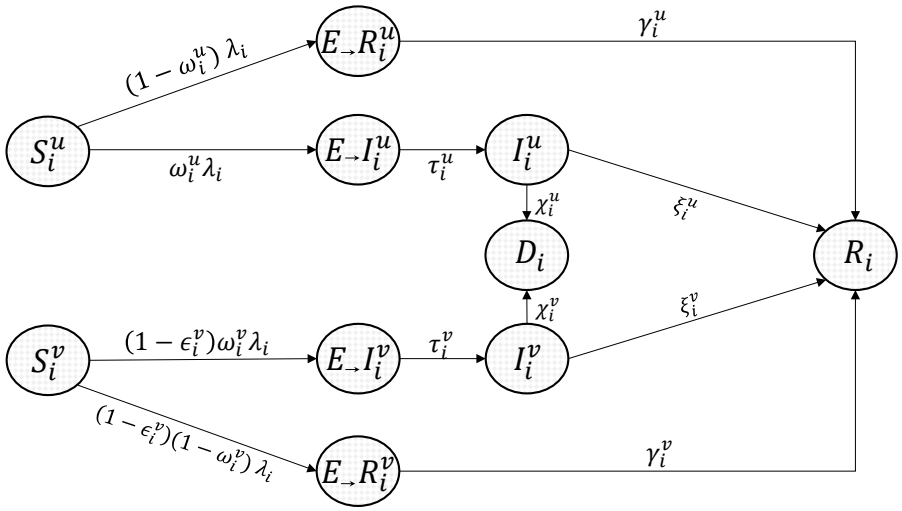


Figure 2.1: **Transition rates between compartments.**

Parameter $\epsilon_i^h \in [0,1]$ denotes the reduction in the likelihood of becoming infected after vaccination, naturally $\epsilon_i^u = 0$ for unvaccinated individuals. Parameter $\omega_i^h \in [0,1]$ denotes the proportion of exposed individuals who eventually become symptomatic. Parameters

τ_i^h, γ_i^h, ξ_i^h, and χ_i^h denote the Exposed-to-Infected, Exposed-to-Recovered, Infected-to-Recovered, and Infected-to-Death transition rates, in that order. We set $\gamma_i^h = \dfrac{1}{1/\xi_i^h + 1/\tau_i^h}$ to ensure that those exposed individuals who recover without showing symptoms stay asymptomatic during the course of the disease. The D compartment is included in the model for tracking the number of influenza-related deaths.

Let p_i denote the vaccinated proportion of age group i. Moreover, let b_i denote the initial number of infecteds in age group i. The boundary conditions $\forall i \in AG$ are given by $S_i^v(0) = p_i(N_i - b_i)$, $S_i^u(0) = (1 - p_i)(N_i - b_i)$, $I_i^v(0) = p_i b_i$, $I_i^u(0) = (1 - p_i)b_i$, $E_i^v(0) = E_i^u(0) = R_i^v(0) = R_i^u(0) = D_i(0) = 0$. We solve the system of differential equations in the SEIR model numerically using the fourth-order Runge-Kutta method (Atkinson, 1989).

2.1.3 Formulating the optimization problem

The optimization aims to find an effective age-specific allocation of a given vaccine supply V such that $\sum_{i \in AG} p_i N_i \leq V$. We consider four different performance measures:

- *Total expected cost* is equal to the sum of vaccination, infection and mortality costs. Vaccination cost (c_i^b) includes the vaccine price, work time lost, and the cost of potential side effects. Infection cost refers to the sum of medication, outpatient visits, and hospitalization expenses, which is different for unvaccinated (c_i^u) and vaccinated (c_i^v) individuals. Mortality cost (c_i^d) stands for the terminal care expenses.

- *Total number of infections* is equal to the number of individuals affected by the pandemic.

- *Total number of deaths* is equal to the number of influenza-related deaths.

- *Total years of life lost (YLL)* weighs each death with the expected remaining life time based on the U.S. life tables and the age distributions (Arias, 2014; United States Census Bureau, 2014).

In particular, at any time during the course of the pandemic, the four performance measures are calculated as follows:

$$\begin{aligned}
\text{Total cost } (TC) &= \sum_{i \in AG} (c^b N_i p_i + c_i^u N_{I_i^u} + c_i^v N_{I_i^v} + c_i^d N_{D_i}), \\
\text{Total number of infections } (TI) &= \sum_{i \in AG} (N_{I_i^u} + N_{I_i^v}), \\
\text{Total number of deaths } (TD) &= \sum_{i \in AG} N_{D_i}, \\
\text{Total YLL } (TY) &= \sum_{i \in AG} Y_i N_{D_i},
\end{aligned}$$

where Y_i denotes the YLL value of age group $i \in AG$. Moreover, $N_{I_i^u}$ ($N_{I_i^v}$), and N_{D_i} represent the total number of infections among unvaccinated (vaccinated) individuals and influenza-related deaths in age group $i \in AG$, respectively.

2.1.4 Calibration

We run FluTe using the population file for Seattle (around 560,000 residents) that is included in the software distribution package. Contact rates within mixing groups and infectious disease parameters of influenza are set based on the values estimated in Chao et al. (2010) so that attack rates are consistent with the 1957 Asian A(H2N2) and 2009 A(H1N1) influenza pandemics without vaccination (Chao et al., 2010). We seed the model with 10 randomly generated infected people. The vaccinated individuals have 40% reduced probability of becoming infected, 40% reduced probability of becoming symptomatic given infection, and 67% reduced probability of transmitting infection. These values represent the effectiveness of a well-matched seasonal influenza vaccine (Basta et al., 2008). The vaccine is only 60% as effective in seniors as everyone else, since older people with weaker immune systems often have a lower immune response to influenza vaccine (CDC, 2016).

The homogeneous mixing assumption of the compartmental model (referred to as SEIR model hereafter) results in faster and more diverse disease spread, whereas the infection

follows a more tranquil pattern in FluTe, i.e., individuals can transmit the disease only to those in their contact list. For a fair comparison of the age-specific vaccine allocation strategies derived by using the SEIR model and FluTe, we calibrate the SEIR model so that the number of new infections for each day of the pandemic closely matches to the corresponding average outcome from FluTe over 100 replications under no vaccination. In particular, we calibrate the SEIR model by varying the contact rates (Φ_{ij}), transmission rates (β_i^h), and initial number of infections (b_i). As a goodness of fit measure, we use Pearson's chi-square statistic (χ^2 measure), and employ the numerical optimization algorithm described in the following section to find parameters of the SEIR model that minimize the maximum χ^2 measure over all age groups.

We first perform the calibration for $R_0 = 1.2$. We then repeat the process for each R_0 value considered in our numerical analyses using the previous calibration results as the initial solution. We keep the contact rates the same as those found for $R_0 = 1.2$ (see Figure A.1) because initial tests show that further calibration of the contact rates for different R_0 values is not necessary to obtain good matches between the results of FluTe and the SEIR model. Figure 2.2(a)a depicts the cumulative number of infections in each day after the calibration process.

The cumulative number of infections in different age groups are presented in Figures 2.2b-2.2f. Furthermore, Figure A.1 shows the similarity of the age-specific attack rates from FluTe and from the SEIR model. The parameters of FluTe and the calibrated SEIR model are provided in Table 2.1. We do not present the parameters of FluTe related to network structure and virus load profile and refer the readers to Chao et al. (2010) for more details.

2.1.5 Solution approach

The pandemic propagation is nonlinear because the incidence of new infections depends on the current number of infectious and susceptible individuals. Moreover, the size and duration of outbreaks are uncertain in agent-based simulations like FluTe. All of these factors render traditional gradient-based optimization methods inapplicable. We therefore use a derivative-free approach, in particular, the mesh-adaptive direct search (MADS) algorithm

Table 2.1: **Input parameters of FluTe and the SEIR model.**

	SEIR	FluTe	Source
Latent period duration (LPD)	1.9 days	1.9 days‡	Chao et al. (2010)
Total disease duration (TDD)	6 days	6 days	Chao et al. (2010)
Infectious period duration (IPD)	4.1 days	4.1 days	TDD−LPD
Transition rate			
Exposed-to-Infectious(τ^u)	0.52632	–	$(\text{LPD})^{-1}$
Exposed-to-Recovered (γ)	0.1667	–	$(\text{TDD})^{-1}$
Infectious-to-Recovered (ξ)	0.2439	–	$(\text{IPD})^{-1}$
Likelihood			
of showing the symptoms (ω^u)†	0.67	0.67	Chao et al. (2010)

	Transmission rate/probability (β^u)*		Initial number of infections (b)		
	SEIR	FluTe	SEIR	FluTe	
1.2	0.0044	0.2057	19.48	10	
1.4	0.0053	0.2435	22.07	10	
1.6	0.0063	0.2812	7.88	10	Calibration
Basic reproduction number (R_0) 1.8	0.0071	0.3190	5.01	10	and Chao et al. (2010)
2.0	0.0077	0.3567	5.01	10	
2.2	0.0083	0.3944	5.01	10	
2.4	0.0089	0.4322	5.01	10	

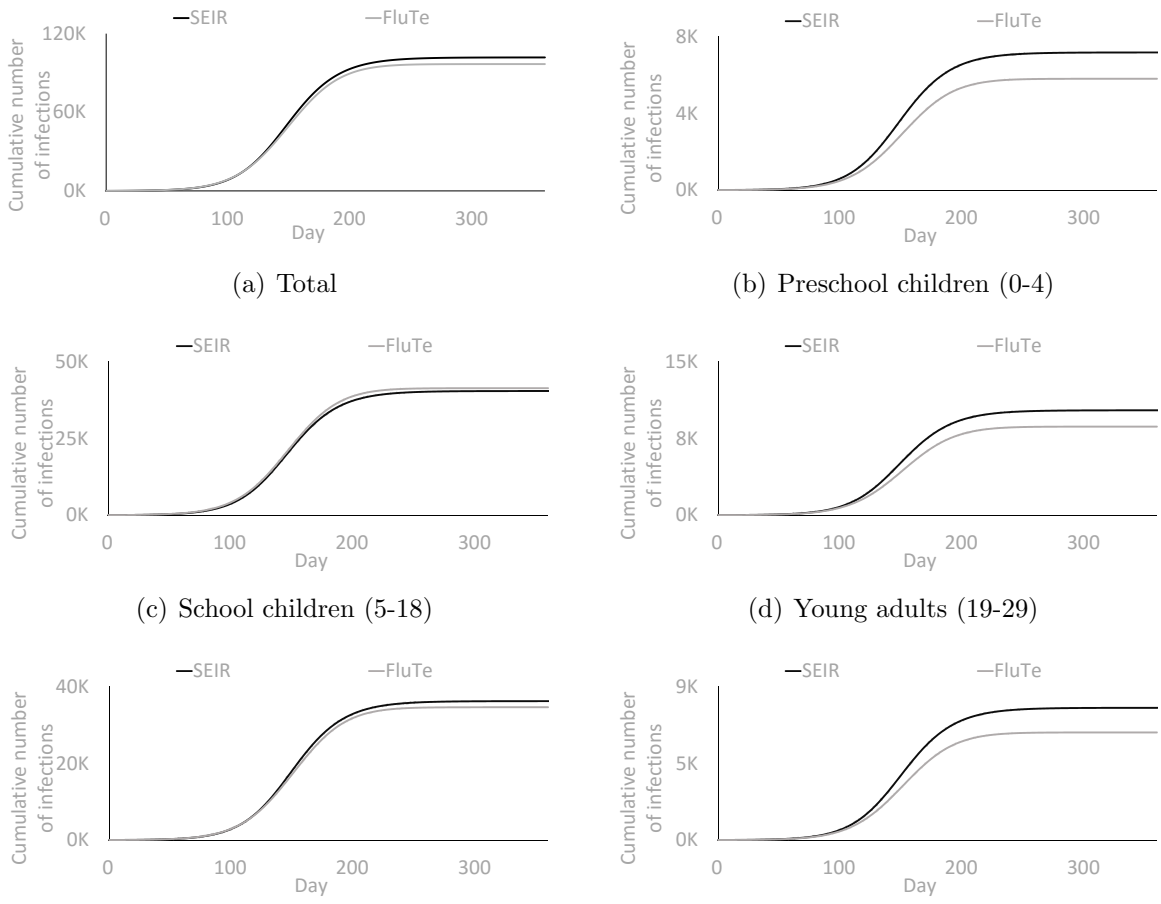
Age group specific disease parameters

	0-4	5-18	19-29	30-64	65+	
Population size	36,722	124,787	71,601	265,047	65,303	Chao et al. (2010)
Death rate (unvaccinated-χ^u)	0.0049	0.002	0.0056	0.0038	0.001	Medlock and Galvani (2009)
Death rate (vaccinated-χ^v)	0.0012	0.0005	0.0017	0.0011	0.0004	Medlock and Galvani (2009)
Vaccine efficacy in						
becoming infected	40%	40%	40%	40%	24%	Chao et al. (2010)
transmitting the disease	40%	40%	40%	40%	24%	Chao et al. (2010)
showing the symptoms	67%	67%	67%	67%	40%	Chao et al. (2010)

Performance related parameters

	0-4	5-18	19-29	30-64	65+	
Infection cost (unvaccinated)($)	275.3	275.3	328.98	328.98	492.56	Medlock and Galvani (2009)
Infection cost (vaccinated)($)	231.58	231.58	264.71	264.71	404.54	Medlock and Galvani (2009)
Terminal care cost($)	3,435	3,435	7,605	7,605	8,309	Medlock and Galvani (2009)
Vaccination cost ($)	37.26	37.26	37.26	37.26	37.26	Medlock and Galvani (2009)
YLL (years)	79.42	70.71	57.92	36.45	13.37	Arias (2014); United States Census Bureau (2014)

‡ : Mean of a discrete distribution of 1, 2, or 3 days with probabilities 0.3, 0.5, and 0.2.
† : These numbers are for the unvaccinated individuals. For the vaccinated individuals, they should be multiplied by (1- vaccine efficacy in showing the symptoms).
* : These numbers are for the unvaccinated individuals. For the vaccinated individuals, they should be multiplied by (1- vaccine efficacy in transmitting the disease).



(a) Total

(b) Preschool children (0-4)

(c) School children (5-18)

(d) Young adults (19-29)

(e) Adults (30-64)

(f) Seniors (65+)

Figure 2.2: **Cumulative number of infections in each age group of FluTe and the SEIR model after the calibration process for $R_0 = 1.2$ without vaccination.**

as implemented in open-source software NOMAD (Abramson et al., 2009). Starting from an initial solution, the MADS algorithm iteratively tries to improve the current best solution by generating trial points on a mesh, which is a discretization of the variable space. Each iteration is composed of two main steps: the search and the poll steps. The search step evaluates a number of trial mesh points. If an improved mesh point is found, then the next iteration is initiated with the new incumbent solution using a larger mesh size. Whenever the search step fails to generate an improved mesh point, then the poll step is

invoked. The poll step explores the variable space near the current incumbent solution. If the poll step also fails to improve the current best solution, then the mesh size and poll size parameters are reduced in order to increase the search resolution. The MADS algorithm stops after a given number of iterations or when the mesh size reaches a precision limit. We refer the reader to Le Digabel (2011) for more information about the MADS algorithm and NOMAD (Le Digabel, 2011).

All numerical experiments are conducted using a PC with 48 cores (2.85 Ghz and 128 GB memory). We run both FluTe and the SEIR model up to one year. We generate 10 trial solutions in each iteration of the MADS algorithm. In one set of experiments, we use FluTe to evaluate the performance of trial solutions (FluTe+MADS), and in another set of experiments we use the SEIR model (SEIR+MADS). When using FluTe, we perform 24 replications to estimate the average performance of each trial solution. Based on our initial experiments, this sample size is sufficient to reduce the effects of sample variance on the results. We terminate the MADS algorithm after 1,000 trial solutions or when the mesh size is less than or equal to 10^{-13}. The SEIR model evaluates trial solutions much faster than FluTe. Therefore, when using the SEIR model, we terminate the MADS algorithm after 100,000 trial solutions or when the mesh size is less than or equal to 10^{-13}. We select $p_i = 0.5,\ \forall i \in AG$ as the initial solution in the MADS algorithm, i.e., vaccinate 50% of the population in each age group. Although this solution may be infeasible under some vaccine coverage scenarios, it is still a proper initial solution because the MADS algorithm allows constraint violations in the intermediate iterations to diversify the search.

2.2 Results

We present the age-specific vaccine allocation strategies derived by FluTe+MADS and SEIR+MADS. We highlight the age groups prioritized by each approach and evaluate the relative performance of the proposed strategies under various scenarios for multiple objective functions.

The basic reproductive number (R_0) represents the average number of infections generated by a typical infectious person in a completely susceptible population. If $R_0 > 1$, the

infection spreads within the population. Otherwise, the infection eventually dies out without any intervention (Heffernan et al., 2005). Fraser et al. (2009) estimated the R_0 value of the 2009 H1N1 pandemic between 1.4 and 1.6. Medlock and Galvani (2009) reported that 30% vaccine coverage can mitigate a pandemic like the 2009 H1N1 if there is no delay in response time (i.e., vaccination starts on the first day of the pandemic). We therefore use $R_0 = 1.6$, 30% vaccine coverage, and no delay in response time as the base case in our experiments. Note that the capacity of mitigating a pandemic with a certain level of vaccine coverage depends on the vaccine efficacy; that is, a greater coverage is required for mitigation when using a less effective vaccine. Moreover, the whole process of producing a pandemic vaccine for a novel influenza virus takes four to six months (WHO, 2009). Therefore, a base case scenario with no delay in response time may not seem practical. However, we consider no other pandemic interventions such as isolation and quarantine, public health measures, and antivirals, which have substantial effect on delaying an influenza pandemic spread. This delay in pandemic spread may render the no-delay scenario more acceptable as a base case.

Table 2.2 reports the age-specific vaccine allocation strategies derived by FluTe+MADS and SEIR+MADS in the base case. SEIR+MADS vaccinates only school children (5-18) to minimize the total cost (TC) objective. For other objectives, preschool (0-4) and school children are mainly vaccinated, and the remaining vaccines are allocated to young adults (19-29). FluTe+MADS vaccinates school children for the most part, and allocates the remaining vaccines to preschool children and young adults. Observe that SEIR+MADS uses fewer vaccine doses than FluTe+MADS for the TC objective, possibly because the effect of vaccination is more pronounced in the SEIR model as a result of the homogeneous mixing assumption.

In Figure A.1, we compare the overall attack rates of the SEIR model and FluTe for different R_0 values and vaccine coverage levels under the same vaccination policy — the available vaccine stocks are initially allocated to school children, the remaining doses are first allocated to preschool children, and then to young adults. For each coverage level and R_0 value, the SEIR model has resulted in less overall attack rate than Flute, illustrating the more pronounced effect of vaccination.

For comparison, we evaluate performances of the vaccine allocation strategies derived

Table 2.2: **Vaccine allocation strategies obtained by FluTe+MADS and SEIR+MADS under the base-case scenario.**

Performance measure		Vaccination fractions for each age group 0-4	5-18	19-29	30-64	65+
TC	SEIR+MADS	-	87%	-	-	-
	FluTe+MADS	-	99%	34%	-	-
TI	SEIR+MADS	100%	100%	9%	-	-
	FluTe+MADS	89%	100%	9%	1%	1%
TD	SEIR+MADS	100%	100%	9%	-	-
	FluTe+MADS	5%	97%	50%	3%	2%
TY	SEIR+MADS	100%	100%	9%	-	-
	FluTe+MADS	21%	98%	50%	1%	-

$R_0 = 1.6$, 30% vaccine coverage, no delay in response time. **TC:** Total cost, **TI:** Total number of infections, **TD:** Total number of deaths, **TY:** Total YLL

by FluTe+MADS and SEIR+MADS using FluTe with 100 replications. As seen in Table 2.3, the strategy derived by FluTe+MADS is significantly better than the one derived by SEIR+MADS for the TC objective. This is mainly due to the fact that the amount of vaccine used in the strategy derived by SEIR+MADS, while containing the disease effectively in the SEIR model, is insufficient to do so in FluTe.

Table 2.3: **Objective values of recommended vaccine allocation strategies.**

Performance measure	Sample mean SEIR+MADS	FluTe+MADS	Difference	p-value
TC ($M)	23.2	6.5	16.7	< 0.001
TI (infections)	1,170.3	945.2	225.1	0.117
TD (deaths)	10.7	9.9	0.8	0.584
TY (life years lost)	459.8	367.5	92.3	0.094

$R_0 = 1.6$, 30% vaccine coverage, no delay in response time. The performance measures are calculated using FluTe with 100 replications.

2.2.1 Sensitivity analysis on R_0

We vary the basic reproduction number R_0 between 1.2 and 2.4 under 30% vaccine coverage with no delay in response time to analyze the sensitivity of the proposed vaccine allocation strategies to R_0. For the total cost (TC) objective, SEIR+MADS increases the vaccinated proportion of school children from 47% to 100% as R_0 increases from 1.2 to 2.0 (see Figure 2.3a). After covering school children, the remaining vaccines are allocated to preschool children when $R_0 \geq 2.0$. FluTe+MADS vaccinates 45% of school children as well as a small portion of preschool children and young adults when $R_0 = 1.2$ (see Figure 2.3b). Furthermore, FluTe+MADS covers all school children when $R_0 \geq 1.6$. SEIR+MADS allocates 35% of the available vaccine when $R_0 = 1.2$ and all of the available vaccine when $R_0 \geq 2.2$. On the other hand, FluTe+MADS allocates 43% of the available vaccine when $R_0 = 1.2$ and all of the available vaccine when $R_0 \geq 1.8$. Intuitively, both methods use fewer vaccine doses for smaller R_0 values to keep the vaccination cost low.

For the total number of infections (TI) objective, SEIR+MADS recommends a similar strategy for each R_0; vaccinate all preschool and school children as well as small portions of young adults and/or seniors (65+) (see Figure 2.3c). However, vaccine allocation strategies from FluTe+MADS vary by R_0 (see Figure 2.3d). These strategies cover all school children and allocate the remaining vaccines to preschool children and young adults when $R_0 \geq 1.4$. Note that both approaches use all available vaccine stocks, as vaccination cost is no longer a concern. This result is also valid for the total number of deaths (TD) and total YLL (TY) objectives.

For the TD and TY objectives, there is a trade off between reducing the number infections (by vaccinating school children) and reducing the casualties in high-mortality age groups (by vaccinating preschool children and young adults). SEIR+MADS again vaccinates all preschool and school children for all R_0 values (see Figures 2.3e and 2.3g). FluTe+MADS covers all school children in addition to some proportions of preschool children and young adults when $R_0 \leq 2.2$ (see Figures 2.3f and 2.3h). However, when $R_0 = 2.4$, FluTe+MADS vaccinates preschool children and young adults for the most part rather than school children. This unique case is due to high mortality rates of preschool children and young adults (see Table 2.1). The vaccine stocks become insufficient to contain

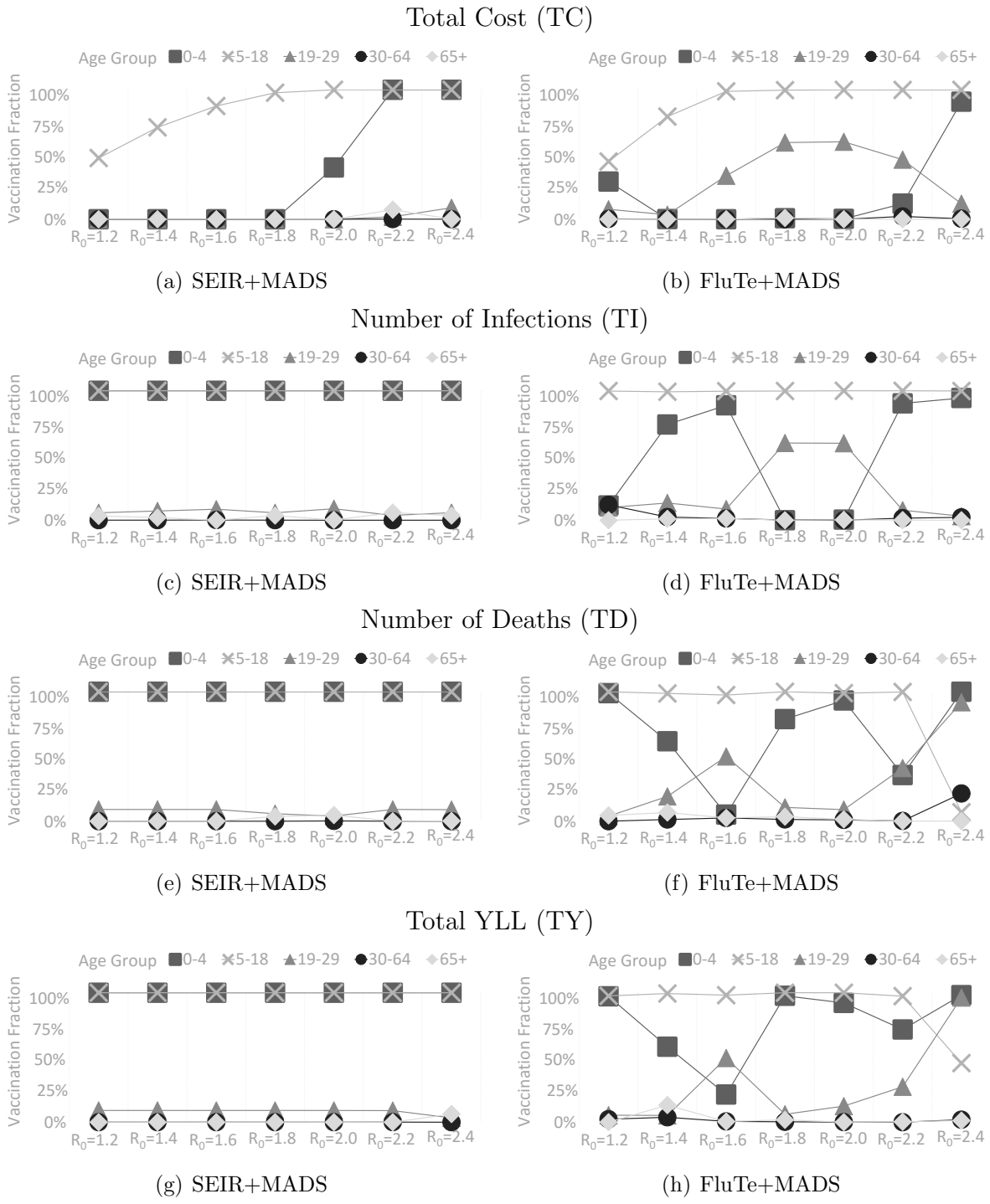


Figure 2.3: **Vaccine allocation strategies derived by FluTe+MADS and SEIR+MADS under all objective functions for various R_0 values (30% vaccine coverage, no delay in response time).**

the pandemic when $R_0 = 2.4$, therefore, the recommended strategy focuses on minimizing casualties in high-mortality age groups.

The vaccine allocation strategies from FluTe+MADS perform better than those from SEIR+MADS when R_0 is between 1.4 and 2.2, i.e., when the effective vaccine allocation becomes critical to control the pandemic (see Figure A.2). However, SEIR+MADS performs better than FluTe+MADS or the difference is not significant when $R_0 < 1.4$ or R_0¿2.2. Recall that we run the MADS algorithm for at most 1,000 trial solutions in FluTe+MADS, while we allow for 100,000 trial solutions in SEIR+MADS. Therefore, it is not unlikely that SEIR+MADS finds better strategies than FluTe+MADS, especially when a small vaccination level is enough to contain the infection (i.e., $R_0 < 1.4$). On the other hand, when the infection is very aggressive, i.e., R_0¿2.2, 30% vaccine coverage in the base case is not enough for containing the pandemic, and therefore, there is limited room for improvement by optimizing the vaccine allocation. One exception to this pattern is for TD and TY objectives where FluTe+MADS performs 7% and 9% better than SEIR+MADS, respectively, by prioritizing high-mortality age groups, i.e., preschool children and young adults when $R_0 = 2.4$.

2.2.2 Sensitivity analysis on vaccine coverage and response time

We set $R_0 = 1.6$ and analyze the sensitivity of the recommended age-specific allocation strategies to vaccine coverage and response time. Table 2.4 presents the strategies from FluTe+MADS and SEIR+MADS for 20%, 30%, and 40% vaccine coverage with no delay in response time (i.e., vaccination starts on the first day of the pandemic). SEIR+MADS vaccinates preschool and school children for the most part, and allocates the remaining vaccines to young adults for all objectives except the total cost (TC) objective for which only 87% of school children is vaccinated under all coverage levels to reduce the vaccination cost. FluTe+MADS vaccinates school children, and allocates the remaining vaccines to preschool children and young adults.

Table 2.5 presents the strategies from FluTe+MADS and SEIR+MADS for three different cases of response time under R_0=1.6 and 30% vaccine coverage. Prevaccination refers to the scenario where the vaccine is administered two weeks before the beginning of the

Table 2.4: **Vaccine allocation strategies for different coverage scenarios.**

		FluTe+MADS					SEIR+MADS				
	Vaccine coverage	**Vaccination fraction for each age group (p_i) 0-4**	**5-18**	**19-29**	**30-64**	**65+**	**0-4**	**5-18**	**19-29**	**30-64**	**65+**
TC	20%	1%	88%	1%	-	-	-	87%	-	-	-
	30%	-	99%	34%	-	-	-	87%	-	-	-
	40%	34%	98%	22%	2%	3%	-	87%	-	-	-
TI	20%	2%	87%	2%	-	1%	-	90%	-	-	-
	30%	89%	100%	9%	1%	1%	100%	100%	9%	-	-
	40%	64%	98%	99%	2%	1%	100%	100%	84%	-	3%
TD	20%	15%	80%	-	-	10%	-	90%	-	-	-
	30%	5%	97%	50%	3%	2%	100%	100%	9%	-	-
	40%	6%	100%	96%	11%	-	100%	100%	80%	2%	-
TY	20%	34%	70%	6%	-	12%	-	90%	-	-	-
	30%	21%	98%	50%	1%	-	100%	100%	9%	-	-
	40%	64%	98%	98%	2%	2%	84%	100%	95%	-	-

$R_0 = 1.6$, no delay in response time

pandemic so that it reaches maximum effectiveness by the time the virus starts spreading. Others refer to the scenarios where the vaccination start d (d =0, 10, 20, 30, 40, 60, 80, or 90) days after the first day of the outbreak. In these delayed scenarios, the vaccine effectiveness will increase gradually reaching its maximum level in two weeks during the course of the pandemic both for the SEIR model and FluTe (Chao et al., 2010).

FluTe+MADS outperforms SEIR+MADS for all objective types when vaccine coverage is 30% see Figure A.3). However, strategies derived by SEIR+MADS for total number of deaths (TD) and total YLL (TY) objectives on average perform better than those derived by FluTe+MADS when the vaccine coverage is 20% or 40% (see Figure A.3). When the vaccine coverage is 20%, there is a small set of effective vaccine allocation strategies that can contain the pandemic. On the other hand, when the vaccine coverage is 40%, the set of effective strategies is large with many local optimums. In both cases, a large number of trial solutions should be evaluated. FluTe+MADS may not find an effective strategy with 1,000 trial solutions when the vaccine coverage is 20%, and gets stuck at a local optimum when the vaccine coverage is 40%. Furthermore, the random noise around the objective values of each FluTe replication may make it difficult for the MADS algorithm to find an effective improvement direction.

Table 2.5: **Vaccine allocation strategies for different response time scenarios.**

		SEIR+MADS					FluTe+MADS				
	Response time (days)	**Vaccination fraction for each age group (p_i)**					**Vaccination fraction for each age group (p_i)**				
		0-4	**5-18**	**19-29**	**30-64**	**65+**	**0-4**	**5-18**	**19-29**	**30-64**	**65+**
TC	P	–	86%	–	–	–	19%	100%	–	–	4%
	0	–	87%	–	–	–	–	99%	34%	–	–
	10	–	89%	–	–	–	77%	100%	15%	–	5%
	20	–	91%	–	–	–	46%	100%	31%	–	4%
	40	–	96%	–	–	–	–	96%	49%	5%	–
	60	–	100%	–	–	–	23%	83%	71%	–	8%
	80	1%	100%	–	–	–	48%	6%	–	54%	–
	90	72%	100%	–	–	–	4%	5%	–	55%	–
TI	P	100%	100%	1%	–	9%	16%	100%	43%	–	–
	0	100%	100%	9%	–	–	89%	100%	9%	1%	1%
	10	100%	100%	9%	–	–	67%	99%	16%	2%	2%
	20	100%	100%	7%	1%	–	–	100%	39%	5%	–
	40	100%	100%	9%	–	–	19%	100%	–	14%	–
	60	100%	100%	9%	–	–	30%	93%	52%	–	5%
	80	100%	100%	1%	2%	–	56%	65%	88%	–	4%
	90	100%	100%	1%	2%	–	59%	40%	–	36%	–
TD	P	100%	100%	9%	–	–	13%	100%	17%	9%	1%
	0	100%	100%	9%	–	–	5%	97%	50%	3%	2%
	10	100%	100%	9%	–	–	–	100%	60%	–	–
	20	100%	100%	9%	–	–	14%	100%	27%	6%	6%
	40	100%	100%	9%	–	–	–	100%	60%	–	–
	60	100%	100%	5%	1%	–	66%	98%	23%	–	8%
	80	100%	100%	9%	–	–	100%	–	100%	22%	–
	90	100%	100%	9%	–	–	87%	–	100%	24%	–
TY	P	100%	100%	9%	–	–	71%	100%	10%	2%	6%
	0	100%	100%	9%	–	–	21%	98%	50%	1%	–
	10	100%	100%	9%	–	–	93%	98%	–	4%	–
	20	100%	100%	9%	–	–	36%	100%	37%	–	4%
	40	100%	100%	9%	–	–	1%	100%	59%	–	–
	60	100%	100%	9%	–	–	100%	100%	9%	–	1%
	80	100%	100%	9%	–	–	100%	47%	100%	–	–
	90	100%	100%	9%	–	–	100%	100%	–	2%	–

$R_0 = 1.6$, 30% vaccine coverage. **P**: Prevaccination.

In Table 2.5, SEIR+MADS vaccinates preschool children and school children for all objectives except for the TC objective and 90-day delay scenario of TC objective. More interestingly, it is very insensitive to response time for TI, TD, and TY objectives. This result is due to the fact that the impact of vaccination is so pronounced in the SEIR model that it can still contain the infection even after a 90-day delay in response time by vaccinating high transmitters. On the other hand, the vaccine allocation strategies from FluTe+MADS significantly vary with response time, although vaccinating the school children is still the main focus in most cases. For total number of deaths (TD) objective, FluTe+MADS mainly increases the vaccinated proportions of age groups who are in high-risk in terms of mortality as the response time is delayed further. When the delay in response time reaches 80 days, the proportion of school children suddenly drops to 0%. These results comply with the literature which suggests prioritizing vaccination of high transmitters (e.g., school children) earlier and prioritizing vaccination of those who have high mortality rate (e.g., young adults, preschool children, adults) later in the pandemic (Matrajt and Longini Jr, 2010; Medlock et al., 2009; Mylius et al., 2008).

For different response time scenarios, the strategies from FluTe+MADS generally outperform those derived by SEIR+MADS (see Figure A.3). However, the performance differences mainly decrease as the response time is delayed further, except TD objective where significant differences are obtained by focusing on high-risk age groups even when late response (60-, 80-, 90-day delay) scenarios.

2.3 Conclusions

Health policy makers commonly use agent-based simulations and compartmental models when evaluating and designing effective vaccine allocation strategies. Our study shines a light on the possible differences between the strategies obtained by these two approaches. In particular, we observe that age-specific vaccine allocation strategies derived using a computationally expensive but realistic agent-based simulation and those derived using a fast but stylized compartmental model may be different from each other. They, however, both recommend vaccinating school children for the most part, which complies with the

literature (Basta et al., 2009; Mylius et al., 2008).

The age-specific vaccine allocation strategies derived using the agent-based simulation significantly outperforms those derived using the compartmental model especially for moderate levels of basic reproduction number (i.e., $1.4 \leq R_0 \leq 2.2$) when vaccine stocks are not very scarce. In other cases, either it is rather easy to control the pandemic (e.g., when $R_0 = 1.2$), so there are several strategies that can effectively control the infection, or vaccination is not sufficient to control the pandemic. In such extreme scenarios, the performance gap between the two approaches is small because there is limited room for improvement.

We also note the following two observations. First, SEIR+MADS tries to limit the number of infected individuals for all R_0 levels using almost the same vaccine allocation strategy for all objective functions under all scenarios. On the other hand, strategies from FluTe+MADS significantly varies for different objective functions and parameter scenarios seeking a balance between the number of infections and influenza-related deaths. Second, the impact of vaccination is very significant in SEIR+MADS possibly due to the homogeneous mixing assumption. Therefore, once evaluated by FluTe, the performance of the strategies from SEIR+MADS appears to be less effective compared those derived by FluTe+MADS. Similar observations are reported by other studies that evaluate the performances of strategies derived by compartmental models using agent-based simulations (Lee et al., 2010).

Our observations summarized above are based on the comparison of a specific agent-based simulation and a compartmental model. However, FluTe is commonly used and well-received in the literature (Cowling et al., 2010; Lazer et al., 2014; Pastor-Satorras et al., 2015). Therefore, although they are not directly generalizable, our observations are likely to hold for other agent-based simulations and compartmental models whose assumptions are similar to those analyzed in this study. The differences between the age-specific vaccine allocation strategies derived by SEIR+MADS and FluTe+MADS are possibly due to the combined effect of considering the heterogeneity in contact patterns and stochasticity in disease progression. However, we left measuring the individual effect of heterogeneity and stochasticity for future research as such an analysis requires a more simplified methodological setting.

There are a few limitations of our analysis. First, we examine age-specific vaccine allocation but do not consider other important risk factors, e.g., chronic medical conditions and pregnancy. Moreover, we use a deterministic numerical optimization algorithm, which does not incorporate the variance in the simulation replications when updating the search direction. Using a faster agent-based simulation, the number of replications can be increased to reduce the variance. Alternatively, ranking and selection methods can be applied to find the proper number of replications for each allocation strategy, which is left for future research.

Chapter 3

Tollgate-based Progression Pathways of ALS patients

3.1 Introduction

Patients with amyotrophic lateral sclerosis (ALS) lose control of voluntary movements over time due to continuous degeneration of motor neurons (Morris, 2015). ALS typically begins in a particular body segment such as arm, leg, and speech and the progression can be in any anatomical direction (rostral, caudal, contralateral) at various paces (Armon, 2008; Mazzini et al., 2008).Although there exist general average progression expectations; from a clinical standpoint, the information is limited to counsel patients about their own specific likely progression pathways (Kiernan et al., 2011). Therefore, there is a need for instruments to better educate providers, patients, and caregivers about what critical events, e.g., weakness starting in a limb, needing feeding tube, or respiratory support, to expect and when to expect them Mitsumoto and Del Bene (2000). Such instruments not only can better prepare the patient and care providers for possible upcoming clinical needs, but also provide a frame of reference to patients when making critical life decisions (e.g., when to retire, remodel a house, consider hospice care, etc.).

Most studies have investigated ALS progression through instruments such as ALSFRS-

R and factors such as site of onset, time between onset and diagnosis, etc. (Brooks et al., 1996; Louwerse et al., 1997; Magnus et al., 2002; Chio et al., 2002; Kimura et al., 2006; Kollewe et al., 2008). Few have analyzed the longitudinal progression for individual ALS patients by tracking changes in ALSFRS-R score (Gomeni and Fava, 2014; Swinnen and Robberecht, 2014; Küffner et al., 2015). Although ALSFRS-R may predict survival well and capture the overall aggressiveness of progression (Kollewe et al., 2008), using ALSFRS-R to determine the timing of specific clinical tollgates or to foresee future needs of ALS patients is challenging due to its design. That is, even though some of the item scores in ALSFRS-R are highly correlated with a few critical events (ALSFRS-R Q3-Swallowing score is 1 = Needs supplemental tube feeding), ALSFRS-R scores do not necessarily characterize the timing of all anticipated tollgate events. Moreover, there is limited research analyzing ALS progression in regards to the time of critical events such as needing respiratory support or wheelchair and studies generally focused on a few critical events for a specific body segment instead of the patient as a whole (Bromberg et al., 2010; Turner et al., 2010; Roche et al., 2012; Balendra et al., 2015).

We introduce a novel and holistic scheme to capture gradual ALS progression based on the critical events referred as tollgates. Using a comprehensive longitudinal dataset from Mayo Clinic's ALS Clinic in Rochester, MN, we characterize the progression through tollgates at the segment and patient levels over time. We describe how the progression based on the followed tollgate pathways varies among patients and ultimately, how this type of progression characterization may be utilized for further studies.

3.2 Methods

3.2.1 Patients and clinical data

In order to analyze the tollgate pathways, we compile a longitudinal dataset containing ALS progression reports for a cohort of 514 ALS patients from Mayo Clinic, Rochester, MN. These patients were referred to the ALS Clinic at Mayo Clinic for their quarterly ALS examinations and to undergo at least one examination. The medical records of the patient

cohort are extracted from Mayo Clinic Amyotrophic Lateral Sclerosis (ALS) Evaluation and Neurologic Examination Forms. The longitudinal data derived from these medical records includes demographics, diagnosis records, ALSFRS-R scores, and assessment reports from neurologists, physiatrists, nurses, speech pathologists, speech therapists, dieticians, and social workers for each examination visit. The assessment reports include a large amount of medical information mostly in free-text form on detailed assessments of the disease progression, medical history of the patients, medication and assistive tools used by the patients. This study was reviewed by the Mayo Clinic Institutional Review Board (IRB) and was deemed as an exempt study.

3.2.2 Tollgate-based staging analysis

For each body segment, we defined a set of critical events which highlights impairment in a segment and any associated assistive technology/equipment needs among ALS patients based on expert opinion. The expert panel, consisting of neurologists, physiatrists, nurses, speech pathologists/therapists, dieticians, and social workers practicing in the ALS Clinic of Mayo Clinic, Rochester, MN, provided their opinions about tollgates related to impairment and disability due to ALS. A set of tollgates are specified for each of five different segments of the body, namely arms, legs, speech, swallowing, and breathing. Table 3.1 summarizes the panel's opinions as a multi-level tollgate scheme for these body segments. For example, the leg tollgate "No weakness" is the initial tollgate when the disease has not affected a patient's legs, whereas Dependent on a wheelchair is the final tollgate when a patient can no longer walk and is always in need of a wheelchair to move around and travel. For each body segment, associated tollgates are labeled with a level value signifying the order of appearance. Considering the irreversible nature of ALS (Orrell et al., 1999), patients reach the tollgates defined for a segment in a non-improving manner in respect to the given level value. That is, once a tollgate was reached by a patient, she cannot return to the tollgates prior to that one in the following visits.

Table 3.1: **Tollgates and their definition for each segment**

Leg

Level	Tollgate	Definition
0	No weakness	No leg weakness.
1	Slight weakness	The patient's leg(s) is weak. However, the weakness does not prevent the patient from performing daily activities such as walking, climbing stairs, or running, etc.
2	Modifying activities	The weakness in the legs requires modification in the patient's activities, such as avoiding long-distance walks, or tripping/falling when walking, but still, no assistance is required.
3	Needing assistance with walking	The patient requires a lightweight assistance tool such as a walker, or cane, but does not use heavyweight assistance tools (e.g., wheelchair, scooter).
4	Needing to use a wheelchair, at least sometimes	The patient starts using a wheelchair-like assistance occasionally but still can walk with a lightweight assistance tool.
5	Dependent on a wheelchair	The patient cannot walk with a lightweight assistance tool and/or has no use (or very little) of legs.

Arm

Level	Tollgate	Definition
0	No weakness	No arm weakness.
1	Slight weakness	ALS has started weakening the arm(s). However, the weakness does not prevent the patient performing daily activities such as dressing, grooming, bathing.
2	Modifying activities	The weakness in the arms requires modification in patient's activities, such as using a button-hook, not being able to raise arms or open lids. Some assistance may be required.
3	Losing useful function of one arm	One arm becomes completely useless. Assistance is required with most of the daily activities.
4	Losing useful function of both arms	Both arms become entirely useless. Constant assistance required with all of the daily activities.

Swallowing

Level	Tollgate	Definition
0	No weakness	ALS has not yet affected the patient's swallowing abilities.
1	Eating/drinking affected	Some difficulty when eating and drinking but still following a regular diet.
2	Modifying what you eat	Eating/drinking is significantly affected. The patient avoids particular foods and/or drinks that are hard to swallow.
3	Needing a feeding tube	The patient cannot swallow consistently; thus, requires a feeding tube.

Breathing

Level	Tollgate	Definition
0	No weakness	ALS has not yet affected the patient's breathing abilities.
1	Limited activity because of SOB	The patient reports shortness of breath (SOB) with exertion or when performing daily activities.
2	Needing BiPAP device at night	The patient requires Bi-level Positive Airway Pressure (BiPAP) device when sleeping. However, BiPAP is not required when awake.
3	Needing BiPAP device during the day	The patient requires Bi-level Positive Airway Pressure (BiPAP) device during the day.
4	Use a ventilator	The patient is using a ventilator.

Speech

Level	Tollgate	Definition
0	No weakness	ALS has not yet affected the patient's speaking ability.
1	Speech affected	The patient's speaking ability is slightly affected resulting in slurred speech. Verbal communication is still possible.
2	Needing a device to assist communication	The patient's speaking abilities are significantly affected. It is difficult to understand the patient. The patient requires a communication device such as IPAD, boogie board, writing to communicate at least some of the time.
3	Losing the ability to talk	The patient is not able to verbally communicate and requires a communication device all the time.

3.2.3 Extracting tollgate data from medical records

The exact time of when a patient reaches a certain tollgate level is not available because medical records usually reflect the patient's condition at regular clinic visits. In addition, a patient may advance several tollgates between two visits for any segment based on the pace of the progression. Therefore, for each body segment, we derived only the tollgate most recently reached at a clinic visit. To do this, we employed natural language processing techniques, using NLTK package in Python (Bird, 2006), as the initial step of extracting and cleaning the data from the free-text clinical notes. These techniques detect specific structures in the free-text data to decompose verbal statements into more quantifiable and standard measures (Cambria and White, 2014).

Firstly, we determined several key-phrases (regularly repeated in the medical records) which might be related to the tollgates reached at each visit. For this purpose, we determined whole sentences in the medical records and excised the commonly-used but not-contributing words known as stopping words (e.g., I/he/she, is, and/or, etc.) from the sentences. Then, two- or three-word phrases (bigrams and trigrams) that frequently repeat in clinical notes were specified. For example, consider the words "right", "arm", "useless" that appear in a clinical note in the respective order. These three words form a trigram possibly indicating the patient has reached Level 3 tollgate of the arm segment, namely losing useful function of one arm. Secondly, because not all bigrams or trigrams were as predictive, we estimated tollgates reached by a classification approach (i.e., classification tree (Loh, 2011)) adopting the bigrams and trigrams as the predictors of tollgates. The classification model was trained by the manually determined tollgates for 10% of patient clinic visits. For the remaining 90%, the tollgates were estimated using the classification model. Note that the described approach was only used for easing the burden of manual abstraction process. Nonetheless, the accuracy and consistency of the tollgate assignments to the visits were verified through a manual review process to correct any false assignment, e.g., it is guaranteed that a patient at a particular tollgate level does not return back to preceding tollgates in subsequent follow-ups.

3.2.4 Handling missing data

Patients are expected to visit the ALS clinic every three months. However, some clinic visits were missed or postponed by patients, resulting in unevenly spaced observations of medical records. In order to simplify the analysis, we considered evenly spaced data points for every three months between the first and the last visits of patients. If any of these evenly spaced data points was not matched with an actual observation (i.e., more than one and a half months away from the closest observation), the reached tollgates and ALSFRS-R scores of that data point were considered to be missing and imputed using available data points. We employed last-observation-carried-forward and linear interpolation techniques for imputing the missing values Chow and Lin (1971); Lane (2008). Considering the performance of the test samples based on cross-validation, we observed that the linear interpolation technique generated more accurate estimates than the last-observation-carried-forward technique (See Figure B.2). Therefore, other analyses were conducted using the dataset where the missing values were imputed by the linear interpolation method. Moreover, ALSFRS-R scores of some patients were missing for some clinic visits. These missing ALSFRS-R scores were imputed based on a classification tree approach using MICE package in R environment (Van Buuren and Groothuis-Oudshoorn, 2011).

3.2.5 Statistical analysis

We conducted Kaplan-Meier analyses to derive the probability of reaching each tollgate at each clinic visit for all body segments considering a 12-month period after the first clinic visit (Kaplan and Meier, 1958). That is, we focused on the ALS progression in the first five consecutive clinic visits (if, of course, a patient did not drop out earlier) including the first visit to the clinic. Patients who dropped out or died by the end of the first year were considered to be censored. Kaplan-Meier curves, including 95% confidence intervals, are generated using Survival package in R environment (Therneau and Lumley, 2015).

3.2.6 Tollgates vs ALSFRS-R scores

ALSFRS-R21 is a questionnaire that monitors the physical conditions of the patients by answering 12 questions (functionality items) related to performing daily tasks (see Table 2). Each item can be answered using a five-level scale (4: normal, 0: unable) making a total score between 0 (severe impairment) and 48 (normal capability). ALSFRS-R is reported as a good predictor of ALS outcomes such as respiratory symptoms failure and survival(Brooks et al., 1996). For many years, ALSFRS-R is used as the standard tool for assessing the conditions of ALS patients and detailed ALSFRS-R score data is available in most databases. Many studies evaluated the progression of the disease based on ALSFRS-R by specifying which factors affects ALSFRS-R scores. We investigate the relationship between the ALSFRS-R items and our tollgates by performing a Spearman's rank correlationanalysis(Sedgwick, 2014).

3.3 Results

3.3.1 Patient characteristics and data

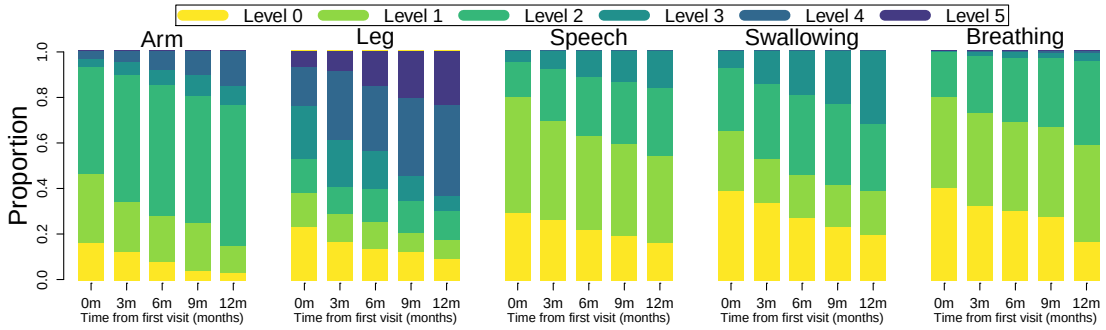
In our patient cohort, there are slightly more males (i.e., 55.3% male), and the mean age is 62.82 (min: 24, max: 93, SD: 11.74). 514 patients visited the ALS clinic at least one time in total with a median of 2 visits per patient. Among patients who had more than one visit, the average time between consecutive visits is around 3.76 months (SD: 2.11). At their first clinic visits, 92.8% of patients displayed some level of limb weakness. Among those who had limb weakness, 17.39% had only arm weakness, 10% had only leg weakness, and both segments were affected in the in the remaining at the first clinic visit. 70.2% of the patients displayed bulbar weakness at the first clinic visit. Among these, respectively, 22.5% and 9.14% had only speech and only swallowing weakness at the first clinic visit. The breathing segment were affected in 59.92% of the population at the first clinic visit. The average ALSFRS-R score in our cohort at their first clinic visit is 36 out of 48. We limited our analysis to the first year after the patients' first clinic visit which makes a total of at most five clinic visits. The number of patients who attended their 1^{st}, 2^{nd}, 3^{rd}, $4^{t}h$,

and 5^th visits are 514 (100%), 346 (67%), 264 (51%), 199 (38%), 164 (32%), respectively.

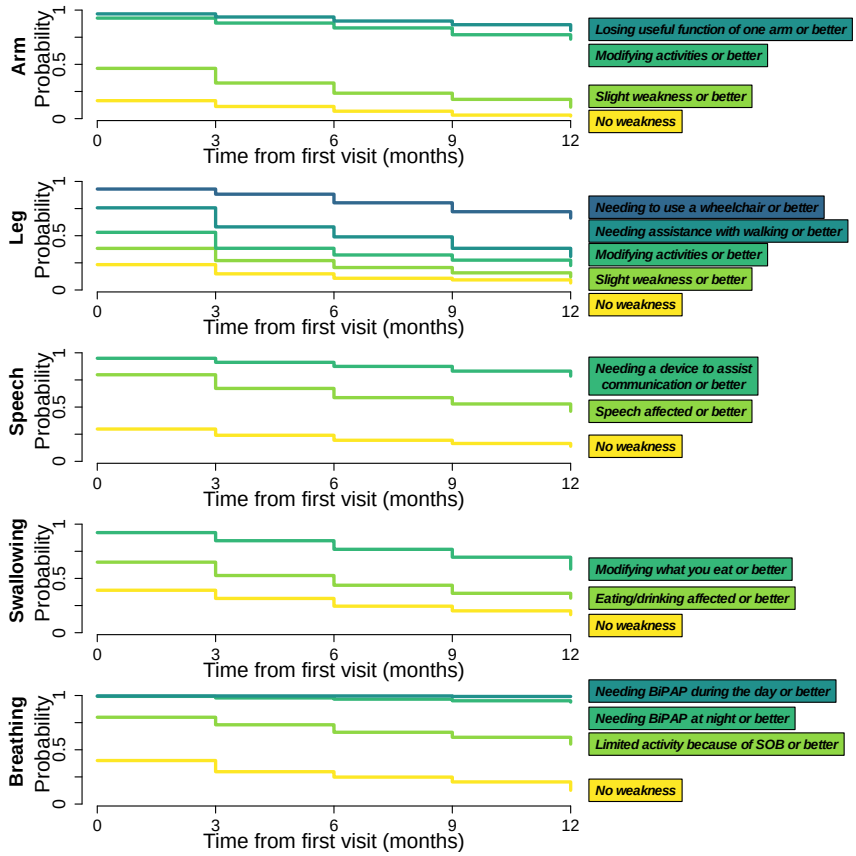
3.3.2 Tollgate profile over time

Figure 3.1(a) illustrates the proportions of patients at each tollgate at each visit (among those who attended the visit) for all segments over a 12–month period. Note that Level 4 is not applicable to speech and swallowing segments, and Level 5 is only applicable to leg segment. See Table 1 for the detailed explanation of tollgate levels. In each segment, the majority of the patients had their abilities affected or worse (Level$\geq$1) at the first visit. Especially, the proportion of patients at higher tollgate levels is larger for arm and leg segments compared to others. Over time, patients moved to higher tollgates due to loss of abilities and had increasing needs for assistive devices. As a result, proportions of patients at Level 3–5 in arm, at 4–5 in leg, 3–4 in other segments increases at each visit. The transition rates to higher tollgate levels are larger in arm, leg, and swallowing than speech and breathing.

The Kaplan-Meier curves in Figure 3.1(b) illustrates the probability of being at a certain tollgate or better at each visit. Thus, for any point in the curve, $(1 - probability)$ indicates probability of reaching beyond the tollgate associated with the curve; i.e., curves with faster decrease over time imply more aggressive ALS progression. In arm segment, the probability of modifying of activities or worse (Arm-Level>2) increases from 0.074 (CI: 0.051–0.096) to 0.268 (CI: 0.212–0.320) in a year; while, probability of losing useful function of at least one arm (Arm-Level>3) increases from 0.035 (CI: 0.019–0.051) to 0.187 (CI: 0.136–0.236). Similarly, the probability of needing a wheelchair or similar equipment drastically increases from 0.243 (CI: 0.205–0.279) to 0.692 (CI: 0.631–0.742) within one year; while, the probability of becoming dependent on the wheelchair (Leg–Level 5) increases up to 0.337 (CI: 0.276–0.393) in one year. Although half of all the comers eventually require an assistive device/tool to communicate or worse (Speech-Level>2); the probability of completely losing ability to speak (Speech–Level=3) is relatively low (0.215, CI: 0.162–0.264, by the end of the year). Finally, the likelihoods of requiring a feeding tube (Swallowing–Level=3) and BiPAP (Breathing-Level>1) increase from 0.078 (CI: 0.101–0.054) to 0.414 (CI: 0.349–0.473) and 0.2 (CI: 0.165–0.234) to 0.448(CI: 0.386–0.503), respectively. Re-



(a) Proportion of patients in each tollgate at each clinic visit over the 12-month period from the initial ALS clinic visit.



(b) Kaplan-Meier curves representing the probability of patients being at a certain tollgate level or better over the 12-month period from the initial ALS clinic visit.

Figure 3.1: **Overall tollgate statistics of the population**.

quiring BiPAP day-and-night or ventilator is not common among the all comers.

3.3.3 Segment-based tollgate pathways

A patient's ALS progression pathway observed until the current point in time may contain clues about the outlook of the pathway. Figure 3.2 illustrates the individual tollgate progression pathways of all patients under each segment. Each patient's pathway is drawn as a line composed of multi-toned sections, each representing the tollgate at a particular visit. These lines are combined in a way that patients sharing the same progression history are grouped together. For instance, patients, who need assistance with walking at the first visit, are represented by the region encapsulated by the red frame under the leg segment column. Of these 81 patients (excluding the drop-outs after the first visit), 39 (45.68%) stayed at the same tollgate, while 38 patients (46.91%) moved to the tollgate for needing a wheelchair at least sometimes (Leg-Level 4), and the remaining 4 (4.94%) patients became dependent on the wheelchair (Leg-Level 5) in their second clinic visit (month 3). In the third visit (month 6), among 38 (46.91%) patients who needed assistance in the second visit, 10 (12.35%) patients dropped-out, 19 (23.46%) patients stayed at the same tollgate, 9 (11.11%) patients moved to Leg-Level 4. By applying the same approach, the proportions of patients following all segment-based pathways can be derived.

From one visit to the next, the majority of the non-dropout patients stay at their previous tollgate level in each segment, e.g., around 71–79%, 81–90%, 84–88%, 81–90%, 86–91% of non-dropout patients stayed in their previous tollgates in arm, leg, speech, swallowing, and breathing segments, respectively, at each visit over a one-year period. Although these proportions are solely high, the proportion of patients who stayed in the same tollgate combination over all segments is much lower (29.48–41.16%) indicating that some progression could be observed in most patients within a 3-month interval. The majority of patients pass at most two tollgates between two consecutive clinic visits, except, in the case of leg segment, a small number of patients have been observed to pass three tollgates in one visit. Figure 2 illustrates that, in arm, speech, and swallowing segments, transition from Level 1 to 2 is relatively more likely (14.95–48.05%, 12.90–55.70%, 10–35.71%, respectively) over all transitions in other tollgate levels, whereas, Level 0 to 1 and

Figure 3.2: **Individual tollgate progression pathways (for each segment separately) over the 12-month period following the first ALS clinic visit.**

White regions represent the drop-outs. Patients having only a single visit are not illustrated for the sake of simplicity.

Level 3 to 4 transitions are relatively more frequent (12.71–55.56%, 10.39–26.35%) over all transitions in breathing and leg segments. For the arm and leg segments, tollgates towards later visits show a great diversity among patients who were at the same level at their first visit.

3.3.4 Tollgate-based ALS progression based on the affected segments at the first clinic visit

To illustrate the effect of the disease history on the future ALS progression in terms of tollgates, the Kaplan-Meier curves in Figure 3.3 characterize the tollgate-specific ALS progression for particular patient groups based on the affected segments in the first clinic visit. As in Figure 3.1(b), $(1 - probability)$ specifies the likelihood of reaching to the next tollgate (or above) at each visit for each patient group. For instance, Figure 3.3g shows that the overall likelihood of not requiring assistance with walking decreases from 0.531 (CI: 0.49–0.576) to 0.225 (CI: 0.183–0.277) in a year, while the same likelihoods of patients in ALSFB (weakness in all segments at the initial visit) and ALB (weakness in arm (A), leg (L), and breathing (B)) groups, decrease from 0.321 (CI: 0.25–0.41) to 0.046 (CI: 0.013–0.16) and from 0.417 (CI: 0.283–0.613) to 0.156 (CI: 0.062–0.394), respectively.

Figure 3 provides some key insights about the tollgate-based ALS progression. First, the Kaplan-Meier curves in Figure 3.3 illustrate the timing of various intervention and assistive device needs within one year from the initial visit for patient groups with different ALS history profiles. For instance, Figure 3.3h shows that patients in groups such as ALSFB and ALB are very likely to need a wheelchair at their third visit around month 9, and up to 45% of those patients will be dependent on a wheelchair by the end of the year. On the other hand, for patients who had weakness only in speech, swallowing, and breathing (SFB), these likelihoods are relatively lower (0 to 0.143 (CI: 0–0.367)).

Second, the Kaplan-Meier curves for the overall patient cohort, i.e., the timing of the transition to the next tollgate for all-comers, may be quite different than those for specific patient groups depending on segment and tollgate type. The differences are especially visible for speech, swallowing, and breathing segments. For groups ALFB (all segments

are affected but speech), ALB, and AL (arm and leg are affected), whose speech and swallowing segments are not affected at the first visit, the likelihood of speech being affected or better stays high (>0.9) within one year, whereas the same probability drops from 0.603 (CI: 0.525–0.693) to 0.226 (CI: 0.135–0.379) for ALSFB patients and from 0.65 (CI: 0.471–0.897) to 0.096 (CI:0.016–0.576) for SFB patients. In addition, these differences are relatively milder for higher tollgates (level≥3) compared to those with lower tollgates. For example, the likelihood of losing the useful function of both arms stays under 0.15 except for patients in SFB.

Third, for a specific patient group, the Kaplan-Meir curves for the tollgates associated with segments that show weakness at the initial visit are lower than those for tollgates of segments that patient does not have any weakness yet. However, even when the curves of the patient's groups for initially affected segments are compared among themselves, significant differences can still be observed. For example, although ALB and SFB patient groups initially had their breathing segment affected, the likelihood (0.95–0.731) for SFB is much higher compared to that (0.611–0.288) for ALB (Figure 3.3q).

With a likelihood of 0.91, patients in SFB display arm weakness at the end of the first year. Losing useful function of both arms is very unlikely for any group within the first year. For the leg segment, SFB curve behaves very distinctly from others for all tollgate levels. Patients in other groups may require wheelchair-like equipment within the first year with likelihood over 0.65. For the speech segment, the probabilities of ALSFB and SFB groups stay below the those of overall, whereas others' likelihoods of staying under a certain tollgate are generally higher than those of the overall. Patients in SFB may lose their ability talk with a probability over 0.5 within the first year. This probability stays below 0.3 for those in ALSFB. Other groups are not likely to lose their ability to talk within the first year. For the swallowing segment, there is a greater chance to modify eating habits within the first year if the patients are in either of the groups ALFB, SFB, or ALSFB. Patients in SFB are very likely to need a feeding tube within the first year. For the breathing segment, all patients in all groups showing breathing weakness (except SFB) generate similar probability curves. Patients in these groups may require a BiPAP device with likelihood over 0.6.

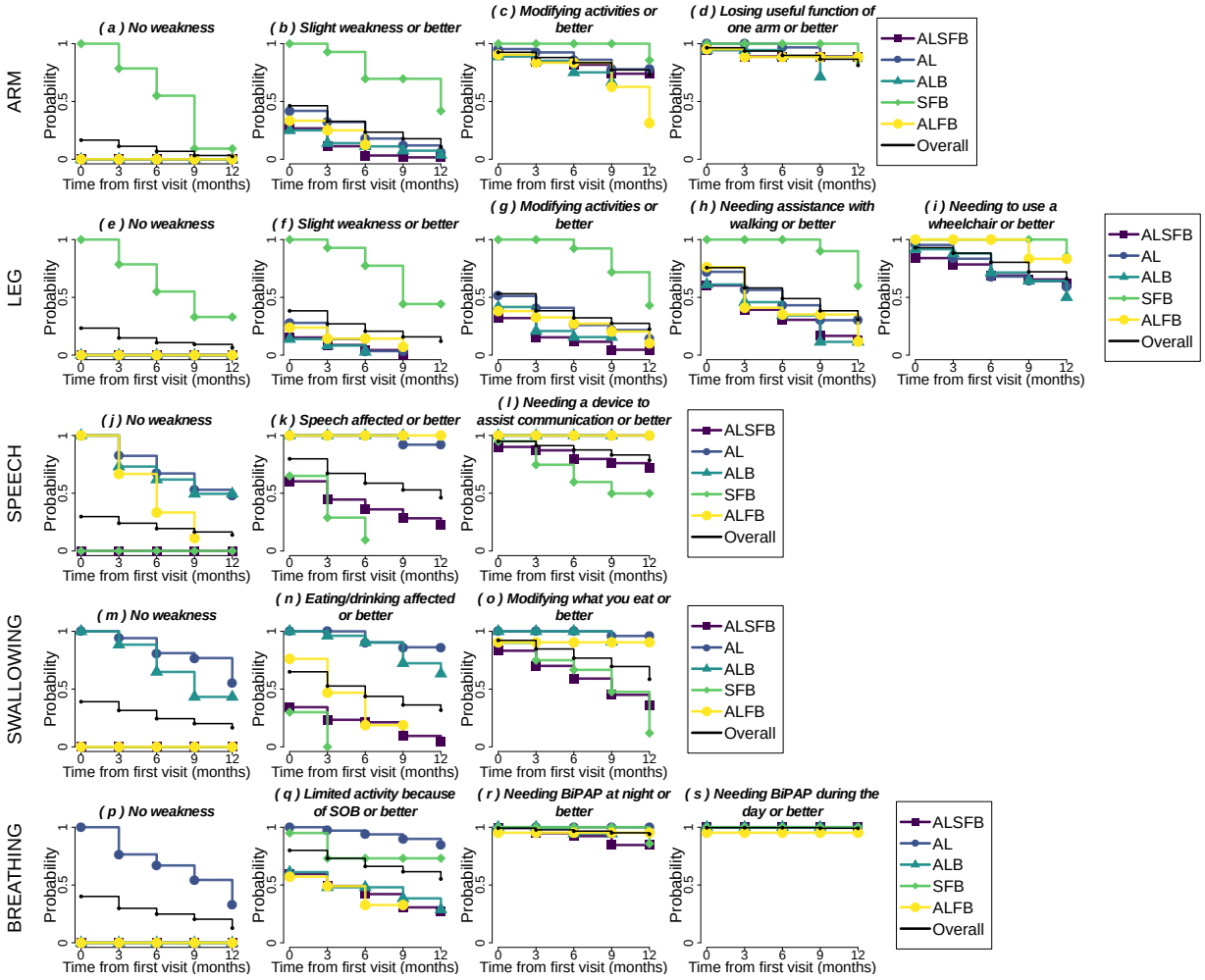


Figure 3.3: **Kaplan-Meier curves representing the probability of being at a certain tollgate level or better for a set of patient groups† over the 12-month period from the initial ALS clinic visit**.

†: Patients are grouped based on the affected segments at the first clinic visit to indicate the effect of the history. Each line in the graphs corresponds to a particular patient group where the initial affected segments are denoted by a combination of the letters: A (Arm), L (Leg), S (Speech), F (Swallowing), B (Breathing). The "Overall" lines illustrate the corresponding probabilities for the all-comers. Note that, among all possible patient groups, only the results for a sample of five are illustrated: ALSFB (131 patients), AL (43 patients), ALB (36 patients), SFB (20 patients), ALFB (21 patients). SOB: shortness of breath, BiPAP: Bi-level Positive Airway Pressure

3.4 Discussion

In this study, we introduced a tollgate mechanism to monitor in which order and when ALS affects different segments of a patient's body. For each segment, tollgates correspond to a set of critical events in a certain order of appearing. The medical records of 514 patients were analyzed in terms of the times when they reach tollgates. We observed that in most patients, the proportions of patients at the high-level tollgates of arms and legs (e.g., losing useful function of an arm or start to use a wheelchair-like equipment) are larger than the those of patients reaching the high-level tollgates of speech, swallowing and breathing segments. This can be explained by the higher prevalence rate of limb-onset patients in the population possibly having relatively milder impairment in the other segments compared to limbs 4. Breathing seems to be the least affected segment, which is expected considering respiratory functions are usually affected later (Gautier et al., 2010). A very small portion of the patients use the BiPAP mask during the day or use a ventilator after the first clinic visit, which agrees with the current literature (Kiernan et al., 2011). The results also show that from one visit to the next, the proportion of patients not showing further impairment (in terms of tollgates) in any of the segments is below 50% within the one year from the first clinic visit. When further impairment was observed, most patients developed symptoms that cause her to reach one upper-level tollgate of a segment within the time between visits.

We showed that when used as a progression monitoring instrument, the proposed tollgate mechanism is capable of capturing the heterogeneity among patients' individual disease courses. A very diverse set of progression pathways is observed among patients for all segments (see Figure 3.2 for details). Using the analysis in Figure 3.2, one may be able to answer what is more likely for a particular patient, given the patient's disease history in terms of when and which of the tollgates are reached. As an example, if a patient uses a cane to assist walking at the first clinic visit, by looking at the existing pathways, we know that there is a high likelihood of needing a wheelchair in the next three months. Moreover, based on the affected regions at the first clinic visit for a specific patient, the likelihood of reaching tollgates over time may vary significantly. For example, the likelihood of requiring a wheelchair for patients not having limb weakness and having all other segments

affected stays under 0.15 during one year after the first clinic visit. Thus, a more conservative recommendation about the acquisition of a wheelchair may be more appropriate for such patients, compared to those having limb weakness at the first clinic visit. Moreover, tollgates can also be employed as a staging system. Noting that the tollgate definitions themselves exhibit an apparent staging scheme within each segment, more comprehensive schemes can be constructed by combing the tollgates from different segments. For example, a staging scheme where the tollgates from all segments are combined, displays the current state (condition) of a patient, i.e., an array with five elements (one value for each segment). The progression of the disease can be monitored in terms of how elements for all segments vary over time. Such schemes are often employed in other disease models. Finally, the proposed mechanism is straightforward to use because the required data can be derived by recording the most recent tollgate level, reasonably by its level value, for each segment at each clinic visit. This is very similar to the way ALSFRS-R scores are recorded, which is the most commonly used instrument to track patients' individual disease course (Marin et al., 2016).

Estimating when and predicting in which order the tollgates are reached could be very useful to better facilitate cooping with the disease burden. The proposed tollgate mechanism could be a useful instrument to track the disease history and predict the future progression. One might question the contribution of the tollgates, considering overlaps between some levels of the tollgates and some items in ALSFRS-R. While likely to be correct at some level, there exist certain discrepancies between them. Firstly, ALSFRS-R scores are usually assessed by patients (or by caregivers), whereas tollgates reached by a patient are proposed to be determined by a clinical team after a comprehensive examination (Montes et al., 2006). Therefore, the assessment of the progression, as captured by the tollgate mechanism, might be a more reliable source of data, considering that the psychological mode is highly associated with the physical impairment of the patients (Böcker et al., 1990). Secondly, ALSFRS-R is designed to capture the current condition of a patient independently of the patient's disease history. As a result, increases in scores of individual ALSFRS-R items are often encountered, implying a betterment in patients' abilities. However, ALS is known to be an irreversible disease. Therefore, a non-improving staging, similar to the one captured by our tollgate mechanism, might be more suitable and infor-

mative in terms of what to expect during the course of the disease. Nevertheless, many studies showed that ALSFRS-R is a very informative measuring tool and ALSFRS-R scores are available in most databases. Therefore, we investigated the associations between the tollgate mechanism and ALSFRS-R questions. We illustrated strong correlations between the segment tollgates and the scores of the questions of the ALSFRS-R (see Figure B.1).

To conclude, the proposed tollgate mechanism can be useful as a clinical progression monitoring system. We observed that it well captures the variability among patients and the history plays a role on when patients reach tollgates. We suggest that further and comprehensive studies should be conducted to observe the whole effect of the history in the future progression. Although the current databases may not have the detailed medical records needed to derive tollgate information to build such models, it might be possible to estimate them via implementing various classification techniques on other available data such ALSFRS-R scores.

Chapter 4

Deriving Tollgates from ALSFRS-R Scores Using Multi-Class Classifiers

Although the aforementioned tollgate mechanism is shown to be informative and easy to interpret, many available databases may not have detailed medical records to derive the necessary data. Therefore, in this chapter, we propose a classification-based approach to estimate the tollgate data using ALSFRS-R scores, which are available in most databases such as Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT) Database (`https://nctu.partners.org/ProACT/`). Recalling Figure B.1, we know that tollgates are significantly associated (correlated) with the ALSFRS-R scores. Here, we want to investigate how this association can be used to estimate the tollgates reached by patients by the time that their ALSFRS-R scores are recorded.

Multi-class classification techniques are commonly used in similar problems to estimate the class given the features of an observation (Aly, 2005). However, traditional classification techniques are not applicable to the problem of deriving tollgate information, due to the constraints requiring a patients' tollgates under a specific segment for multiple visits to be non-decreasing over time (see Section 3.2.3 for details). Therefore, we propose two approaches to achieve a multi-class estimation in a non-decreasing manner given a classification method. The first is the naive approach, which fixes the tollgate estimates of observation in the forward direction in terms of time. The second is the optimization

approach which utilizes a mixed integer programming model to estimate all the tollgates of a patients' observations at the same time. The detailed descriptions of these approaches are presented in Sections 4.3.1 and 4.3.1.

Both of the aforementioned approaches require a multi-class classification model as an input. We use five different multi-class classification techniques. Therefore, in total, 10 different classification models (five models for each approach) are proposed to estimate the tollgates. Note that we investigate the disease progressions in different body segments independently; therefore, the performances of the models are presented for each segment.

4.1 Classification models

Statistical classification is the process of estimating to which of the several predetermined classes/subgroups a new observation pertains, based on a known set of features of the observation. The classification problem where there are only two classes to which observations can be assigned is known as *binary* classification (Crammer and Wagner, 2012). As an example, specifying if a tumor is benign or malignant based on the size of a tumor is a well-known binary classification problem. If the number of classes are greater than two, it becomes a *multi-class* classification (Hsu and Lin, 2002). A well-known example is the classification performed on *Iris flower data set* (Fisher, 1936), which estimates the species (from a set of three) of a flower based on its petal (colorful part) and sepal (green part) size dimensions. In our problem, the classes are the tollgates levels for a specific segment and the features are the ALSFRS-R scores, both of which are available in our data set.

Classification models (also called as classifiers) are studied under the umbrella of supervised learning methods. That is, classification models are trained by a set of observations whose classes are already known, so that associations between the features and classes of the observations can be constructed (Hastie and Tibshirani, 2011).

Let us denote classifier G which estimates the classes (y) of a given observation based its features ($\mathbf{X}$).

$$G{:}\mathbf{X}{\sim}y \tag{4.1}$$

The classifier G maps $\mathbf{X}$ to the class estimation $\hat{y}$ based on the classifier parameters, θ.

$$G(\mathbf{X}|\theta)=\hat{y} \tag{4.2}$$

Although characteristics of the parameters vary significantly among different classification techniques, all classifiers aim to minimize the classification error incurred by the model. The classification error can be measured in different ways. In this study, we use three well-known performance metrics for classification error in a multi-class classification setting.

Performance metrics

Overall accuracy (OA) measures the ratio of the observations whose classes are correctly estimated by the model (Briem et al., 2002):

$$OA=\frac{\sum_{i=1}^{N} A_i}{N} \tag{4.3}$$

where N denotes the total number of observations and A_i=1 is if y_i=$\hat{y}_i$, or A_i=0, otherwise. *Mean squared error* (MSE) measures the average of the squared difference between the estimated class and the actual class of the observations (Kohn and Ansley, 1986):

$$MSE=\frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N}. \tag{4.4}$$

Coefficient of determination (*R-squared*) is defined as the the proportion of the variance in the class that is predictable from the features (Mittlböck and Schemper, 1996). Given $\bar{y}$=$\sum_{i=1}^{N} y_i/N$ is the average class of all observations, the *R-squared* can be expressed as the following:

$$\textit{R-squared}=1 - \frac{SS_{est}}{SS_{tot}} \tag{4.5}$$

where SS_{tot}=$\sum_{i=1}^{N}(y_i - \bar{y})^2$ is the total some of squares of classes and SS_{est}=$\sum_{i=1}^{N}(y_i - \hat{y})^2$ is the total sum of squares of residuals.

Note that both MSE and *R-squared* can be used only if classes have an ordinal structure

so that the difference, $y_i - \hat{y}_i$, has a practical meaning.

4.2 Problem definition and notation

If we assume that the observations of a patient are independent of each other and try to estimate each observation individually, a traditional classifier could be employed to predict the tollgate levels at each visit. The difficulty arises when we try to predict the tollgates of all observations in a non-decreasing fashion over time. This comes from the fact that patients reach the tollgates for a segment in a particular order and the progression is considered to be irreversible. For example, if a patient reaches the leg tollgate *Needing assistance with walking* (Leg-Level 2) at her first clinic visit, we want to make sure that the tollgate level at the second visit (or later visit) is at least the tollgate level at the second clinic visit (leg tollgate level ≥ 3). We should also make sure that this condition is satisfied in the reverse direction. For example, the tollgate level at the first clinic visit can be at most the leg tollgate level at the second visit (leg tollgate level ≤ 3). On the other hand, the ALSFRS-R scores, by definition, is not necessarily restricted to a non-improving assessment of the disease, since it aims to investigate the current condition of a patient independent of her history.

For each segment, we want to estimate the response variable (class) $y_{p,t}$ where t denotes the response (tollgate level) at the t^{th} observation of Patient p. We want to estimate the $y_{p,t}$, using the feature set (scores from a subset of ALSFRS-R items) denoted as $X_{p,t}$. The following is the notation of the classification model, G. For now, we ignore the patient index, p, for brevity of the representation.

$$G{:}(X_1, X_2, X_3, \ldots, X_T){\sim}(y_1, y_2, y_3, \ldots, y_T) \tag{4.6}$$

where T is the total number observations of a given patient. Let $\hat{y}_t; t{=}1, 2, \ldots, T$ be the estimated classes derived using the model defined in (4.6) such that,

$$G(X_1, X_2, X_3, \ldots, X_T|\theta){=}(\hat{y}_1, \hat{y}_2, \hat{y}_3, \ldots, \hat{y}_T). \tag{4.7}$$

The estimated classes must satisfy the following monotonicity constraint.

$$\hat{y}_1 \leq \hat{y}_2 \leq \hat{y}_3 \leq \ldots \leq \hat{y}_T \tag{4.8}$$

4.3 Solution approaches

In this section, we propose two classification approaches where the monoticity constraint (4.8) is guaranteed by the class estimations. Both approaches employ a traditional classifier, assuming observations of patients are independent of each other. Traditional classifiers estimate the corresponding class of an observation using only the features of that observation ($G{:}X_t \sim y_t$). In order to be applicable, a traditional multi-class classifier (e.g., multinomial logistic regression, random forests, etc.) must be able to generate the likelihood of all classes, given futures of the observation t. That is, an estimation of the following probability it required:

$$D_{tl} = \mathbb{P}_G \{y_t = l | \, X_t\}; l = 1, 2, \ldots, L \tag{4.9}$$

where L denotes the total number of classes.

4.3.1 The naive approach

The idea behind this approach is fixing the classes of observations one by one in the direction from the first observations to the last observations. At each step, the class with the highest probability (D_{tl}) is selected among the classes which are greater than or equal to the estimated class of the previous observation. Algorithm 1 presents the procedure of the naive approach.

4.3.2 The optimization approach

The idea behind this approach to determine the classes of all observations together. This is done by a mixed integer programming model that maximizes the sum of probabilities of

Algorithm 1 Naive approach

1: **procedure**
2: **for** $t \in \{1, 2, \ldots, T\}$
3: **if** $t=1$ **then**
4: $\hat{y}_t \leftarrow \max_{l=1,2,\ldots,L} \{D_{1l}\}$
5: **if** $t>1$ **then**
6: $\hat{y}_t \leftarrow \max_{l=\hat{y}_{(t-1)},\ldots,L} \{D_{tl}\}$

the selected classes over all the observations of a patient.

L	:	Total number of classes
T	:	Total number of observations
D_{tl}	:	Probability of the class of observation $t \in \{1, 2, \ldots, T\}$ being $l \in \{1, 2, \ldots, L\}$
u_{tl}	:	1, if $l \in \{1, 2, \ldots, L\}$ is the selected class of observation $t \in \{1, 2, \ldots, T\}$; 0, otherwise

$$\text{Maximize} \qquad z = \sum_{t=1}^{T} \sum_{l=1}^{L} D_{tl} u_{tl} \tag{4.10}$$

$$\text{subject to} \qquad \sum_{l=1}^{L} u_{tl} = 1, \qquad \forall t \in \{1, 2, \ldots, T\} \tag{4.11}$$

$$\sum_{l=1}^{L} l u_{(t-1)l} \leq \sum_{l=1}^{L} l u_{tl}, \qquad \forall t \in \{2, \ldots, T\} \tag{4.12}$$

$$u_{tl} \in \{0, 1\}, \qquad \forall t \in \{1, 2, \ldots, T\}; \forall l \in \{1, 2, \ldots, L\}. \tag{4.13}$$

.

Once the solutions are obtained, the class estimation can be easily derived using the following equation.

$$\hat{y}_t = \sum_{l=1}^{L} l u_{tl} \qquad \forall t \in \{1, 2, \ldots, T\} \tag{4.14}$$

Note that the above MIP model (4.10-4.13) needs to be solved for the observations set of each patient and for each segment, which is done using IBM ILOG Cplex 12.7.1.

Most classifiers such as multinomial logistic regression allow to generate the probability estimates, D_{tl}, after the training processes

4.3.3 Classifiers

Multinomial logistic regression

Multinomial logistic regression classifiers establish logistic regression frameworks to multi-class problems (Kwak and Clayton-Matthews, 2002). The model estimates the probability of each class using the following equation:

$$\mathbb{P}\{y{=}l|\,\mathbf{X}{=}\{x_1,x_2,..,x_n\}\}{=}\frac{exp(\beta_{0,l}+\sum_{i=1}^{n}\beta_{i,l}x_i)}{\sum_{l=1}^{L}exp(\beta_{0,l}+\sum_{i=1}^{n}\beta_{i,l}x_i)};l{=}1,2,\ldots,L \tag{4.15}$$

where x_n is the value of feature i. The model optimizes the coefficient parameters $\beta_{i,l}, i\in\{1,2,..,n\}$ to maximize the likelihood of observed classes.

Support vector classifiers

Support vector classifiers come from a statistical learning theory aiming to determine the decision boundaries on a multi-dimensional space of features, which yields the best separation of classes (Vapnik, 2013). In the simplest case, there are only two classes and a support vector classifier divides two-dimensional feature space into two regions with a linear hyperplane so that the margin is maximized. The margin is denoted by the sum of distances between the hyperplane and the closest data points of the two classes (Vapnik, 2013). These closest data points are called the "support vectors". When there is no hyperplane that can isolate the points from the classes on two sides of the hyperplane, the algorithm then tries to minimize the classification error aside from maximizing the margin (Vapnik, 2013). The algorithm can also be used to handle non-linear separations

with kernel functions (e.g., linear, polynomial, Gaussian, etc.) that aim to alleviate the computational burden (Vapnik, 2013).

Random forest classifier

Random forest classifiers consist of a random mixture of multiple tree classifiers (Breiman, 2001). In each of the single tree classifiers, subgroups (branches) of observations showing low variability among them are created by conditioning on the values from the feature set (Safavian and Landgrebe, 1991). There are many methods for the selection of features to be conditioned on and most approaches utilize a proxy performance measure such as Gini Index (Breiman et al., 1984). Once the branches are created based on training data and the class of a new observation is to be estimated, the observation is matched with one of the branches according to the values of the features; then, the class of the observation is estimated by selecting the most frequent class of a branch (Safavian and Landgrebe, 1991). In the case of the random forest, the classifier first creates many single tree classifiers from random combinations of features. When estimating a class, each tree uses a unit vote on one of the classes as its class estimation, and the highest-voted class becomes the estimation of the random forest (Breiman, 2001).

Classifier settings

We employ all the classifiers mentioned above using the scikit-learn library in a python environment in their default settings (Pedregosa et al., 2011). For the multiple logistic regression, to solve the optimization problem of maximizing the likelihood of observations, the nonlinear conjugate gradient method is used with upto 100 iterations. For support vector classifier, we used three different kernel functions: linear, third-degree polynomial, and Gaussian. For each kernel function, a support vector classifier is introduced. Therefore, a total of three classifiers are investigated under the support vector classification context. For all three classifiers, the parameter C defining the trade-off between the total margin and the classification error is set to the default value, *1*. For the random forest classifier, the total number of trees in the forest is set to the default value, *10*. Moreover, Gini Index is used for the branching operations.

For all of the five classifiers, the balanced sample option is considered to reduce the dominance of the frequent classes. However, this option may possibly reduce the total classification accuracy. For the feature selection, we employ recursive feature elimination with cross-validation (recursive feature elimination with cross validation (RFECV)) algorithm as implemented in scikit-learn package in python environment to select the best set of features (Guyon et al., 2002; Pedregosa et al., 2011). As a prepossessing, we eliminated the ALSFRS-R questions whose correlation are less than 0.4 for all segments. RFECV eliminates features based on their weights assigned using a given classifier. Futures having the smallest weights are removed from set of estimator features. The algorithm repeatedly eliminates features until the scoring metric (e.g., accuracy) no longer improves. Note that, recursive feature elimination is a greedy backward features elimination method (Ellis and Petridis, 2009). Therefore, it does not guarantee to find the optimal set of features.

Note that each classifier has one navie and one optimization approaches. Therefore, a total of 10 classifiers are proposed, the performances of the classifiers based on numerical analysis are presented in the next section.

4.4 Selected futures and numerical results

For all segments and classifiers, the selected features are presented in Table 4.1. On the support vector classifiers with polynomial and Gaussian kernels, RFECV algorithm was not applied because these classifiers do not produce the feature coefficients to score features to be used in the elimination steps (for details see `http://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.\gls{rfecv}.html`). Under the leg segment, out of the proposed features Q5-Q9, RFECV algorithm eliminates Q6 and Q7 for the multinomial logistic regression and linear SVC classifiers, whereas Q4, Q6, and Q7, are eliminated for the random forest classifier. Under the arm segment, out of the same proposed features for the leg segment, RFECV eliminates Q8 and Q9 for all the classifiers and additionally, Q7 is eliminated for the random forest classifier. Under the speech segment, among Q1-Q3, for both the linear support vector and random forest classifiers, Q1 is kept as the single predictor, while Q1 and Q3 reside in the proposed predictor set

for the multinomial logistic regression classifier. Under the swallowing segment, Q1-Q3 are the proposed features and only for the random forest classifier, Q2 is eliminated from the predictor set. Lastly, under the breathing segment, RFECV eliminates only Q11 for the multinomial logistic regression classifier among Q10-Q12.

Table 4.1: **Selected ALSFRS-R questions to predict the tollgate level of a segment**.

			ALSFRS-R question											
		Classifier	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**	**Q11**	**Q12**
Segment	**Leg**	MLR						✓	✓	✓	✓			
		SVC_L						✓	✓	✓	✓			
		SVC_P				✓	✓	✓	✓	✓	✓			
		SVC_G				✓	✓	✓	✓	✓	✓			
		RF					✓			✓	✓			
	Arm	MLR				✓	✓	✓	✓					
		SVC_L						✓						
		SVC_P				✓	✓	✓	✓	✓	✓			
		SVC_G				✓	✓	✓	✓	✓	✓			
		RF				✓	✓	✓						
	Speech	MLR	✓		✓									
		SVC_L	✓											
		SVC_P	✓	✓	✓									
		SVC_G	✓	✓	✓									
		RF	✓											
	Swallowing	MLR	✓	✓	✓									
		SVC_L	✓	✓	✓									
		SVC_P	✓	✓	✓									
		SVC_G	✓	✓	✓									
		RF	✓		✓									
	Breathing	MLR										✓		✓
		SVC_L										✓	✓	✓
		SVC_P										✓	✓	✓
		SVC_G										✓	✓	✓
		RF										✓	✓	✓

4.4.1 Cross validation results

Figure 4.1 illustrates the mean performances of the proposed classifiers based on 10-fold cross validation under aïve (N) and optimization (O) approaches for the leg segment. Based on the accuracy metric, among the optimization approaches, the all classifiers perform similarly. The support vector classifier with linear kernel ($SVC_L^{(O)}$) achieves the highest

accuracy (0.631), while on-average performance of all classifiers with the optimization approach above 0.6. On the other hand, all the naïve approaches perform poorly compared to optimization approaches and their performances are between 0.4 and 0.5. The highest accuracy difference between the optimization and naïve approaches of a classifier is achieved by the support vector classifier with the Gaussian kernel ($SVC_G^{(O \text{ or } N)}$), which incurs an accuracy difference of 0.203 between the optimization and the naïve approaches. Based on the MSE and R-squared metrics, the performances of the classifiers show similar patterns to those observed for the accuracy metric. That is, all optimization approaches of the classifiers perform better than their naïve approach pairs. For all the naïve approaches, the MSE values stay above 1.6 and the R-squared values stay below 0.4. On the other hand, among the optimization approaches, the MSE values stay close to 0.4 while R-Squared values reach above 0.75. The lowest MSE (0.618) and the highest R-Squared (0.772) values are achieved by the optimization approach of the multiple logistic regression classifiers ($MLR^{(O)}$).

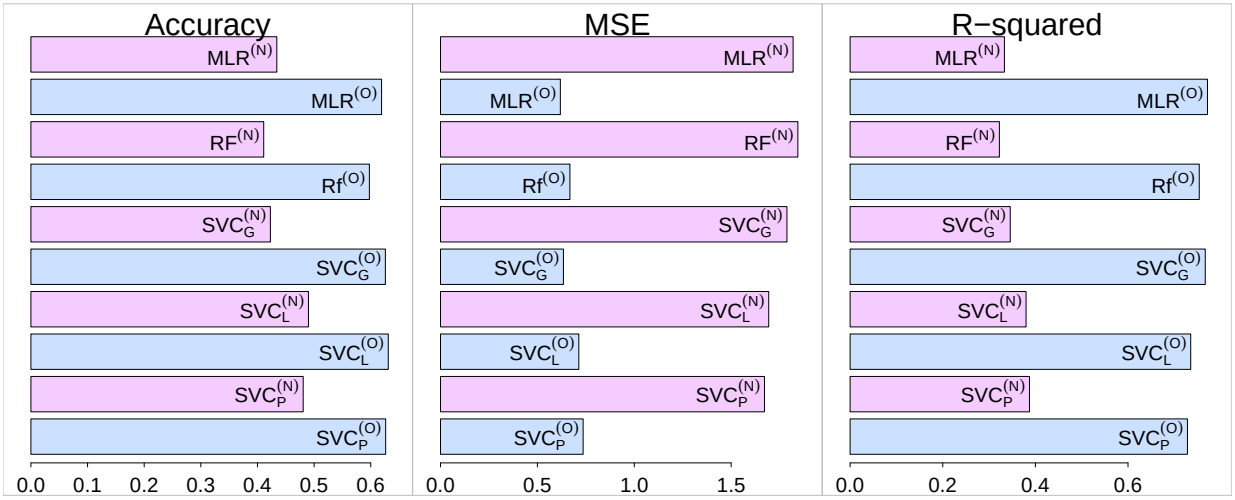


Figure 4.1: **The mean performances of the proposed classifiers based on 10-fold cross validation under the naïve (N) and the optimization (O) approaches for the leg segment.**

Figure 4.2 illustrates the mean performances of the proposed classifiers based on 10-fold cross validation under naïve (N) and optimization (O) approaches for the arm segment.

Based on the accuracy metric, among the optimization approaches, support vector classifiers with polynomial ($SVC_P^{(O)}$) , Gaussian ($SVC_G^{(O)}$), and polynomial ($SVC_P^{(O)}$) kernels perform favorable compared to other classifiers. Both naïve and optimization approaches of the multiple logistic regression classifier ($MLR^{(N)}$ and $MLE^{(O)}$) perform poorly compared to the classifiers within the same approach group. The support vector classifier with Gaussian kernel ($SVC_G^{(O)}$) achieves the highest mean accuracy (0.682) among all classifiers, while mean performance of all classifiers with the optimization approach above 0.6 except $MLR^{(O)}$. On the other hand, all the naïve approaches perform poorly compared to optimization approaches while the performance of $SVC_G^{(N)}$, $SVC_P^{(N)}$, and $SVC_L^{(N)}$ stay above 50%. Based on the MSE and R-squared metrics, the performances of the all classifiers present similar characteristics to those for the accuracy metric on general. However, $MLR^{(O)}$ performs better than $SVC_G^{(N)}$, $SVC_P^{(N)}$, and $SVC_L^{(N)}$ in both MSE and R-squared metrics although the opposite is true for the accuracy metric. For all the naïve approaches, the MSE values stay above .8 and the R-squared values stay below 0.21. Among the the optimization approaches, for $SVC_P^{(O)}$, $SVC_G^{(O)}$, and $SVC_P^{(O)}$, the MSE values stay below 0.6 while R-Squared values stay close to 0.5. The lowest MSE value (0.517) is achieved by $SVC_G^{(O)}$ while the highest R-Squared value (0.495) is by $SVC_L^{(O)}$.

Figure 4.3 illustrates the mean performances of the proposed classifiers based on 10-fold cross validation under naïve (N) and optimization (O) approaches for the speech segment. All the optimization approaches of the classifiers outperform those with the naïve approach. All the optimization approaches of classifiers perform except the support vector classifier with polynomial kernel ($SVC_P^{(O)}$) perform above 0.74. The optimization approach of the support vector classifiers with linear kernel ($SVC_L^{(O)}$) achieves the best accuracy (0.744), MSE (0.269), and R-squared (0.689) values. The naïve approaches of all classifiers achieve accuracy values above 0.5. Expect $SVC_P^{(O)}$, MSE values of the optimization approaches of the classifiers stay between 0.25 and 0.30 while the same metric stays between 0.45 and 0.55 for the naïve approaches of the classifiers. All the optimization approaches incur an R-squared value above 0.6 whereas it stays close to 0.4 for the classifiers with the naïve approaches.

Figure 4.4 illustrates the mean performances of the proposed classifiers based on 10-fold cross validation under naïve (N) and optimization (O) approaches for the swallowing

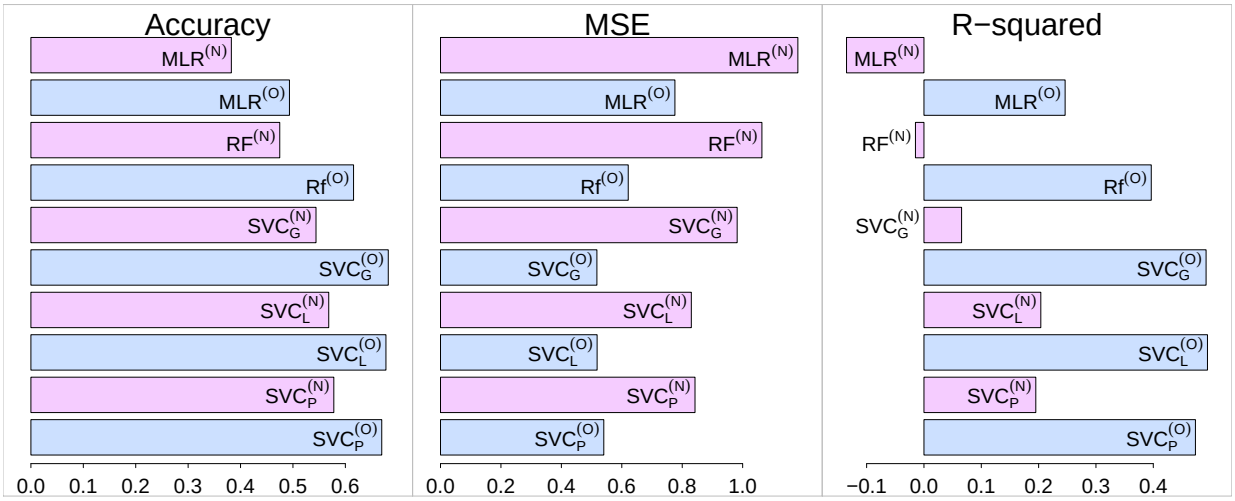


Figure 4.2: **The mean performances of the proposed classifiers based on 10-fold cross validation under the naïve (N) and the optimization (O) approaches for the arm segment.**

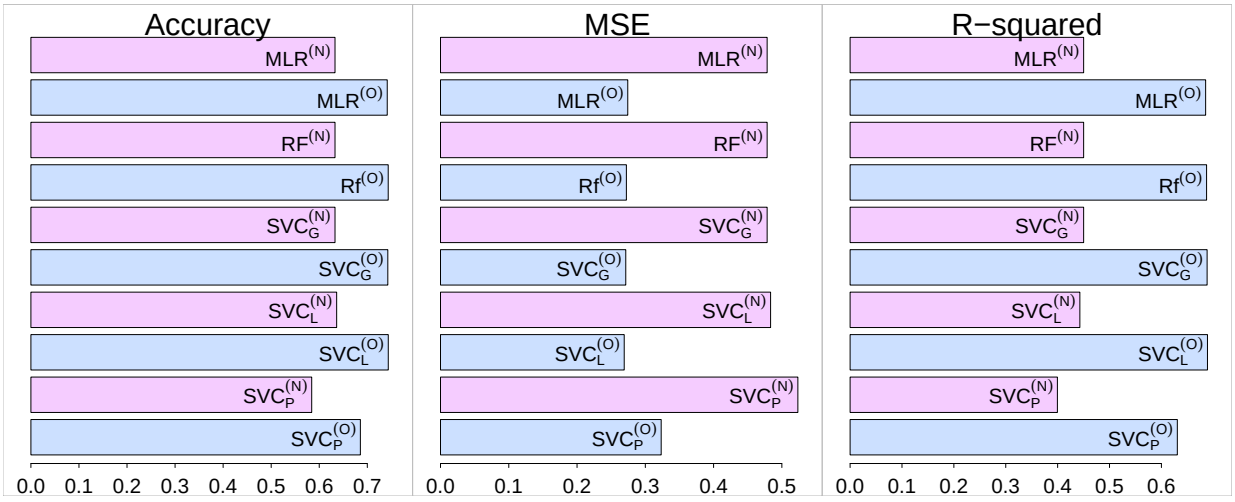


Figure 4.3: **The mean performances of the proposed classifiers based on 10-fold cross validation under the naïve (N) and the optimization (O) approaches for the speech segment.**

segment. Similarly, all the optimization approaches of the classifiers outperform those with the naïve approach. All the optimization approaches of classifiers perform similarly by achieving an accuracy above 0.65. The optimization approach of the support vector classifier with Gaussian kernel $SVC_G^{(O)}$ achieves the highest accuracy (0.696), while optimization approach of the random forest classifier ($RF^{(O)}$) achieves the best MSE (0.460) and R-squared (0.595) values. Similar the speech segment, the naïve approaches of all classifiers achieve accuracy values above 0.5. All the optimization approaches attain an R-squared value between 0.5 and 0.6 whereas it stays below 0.3 for the classifiers with the naïve approaches.

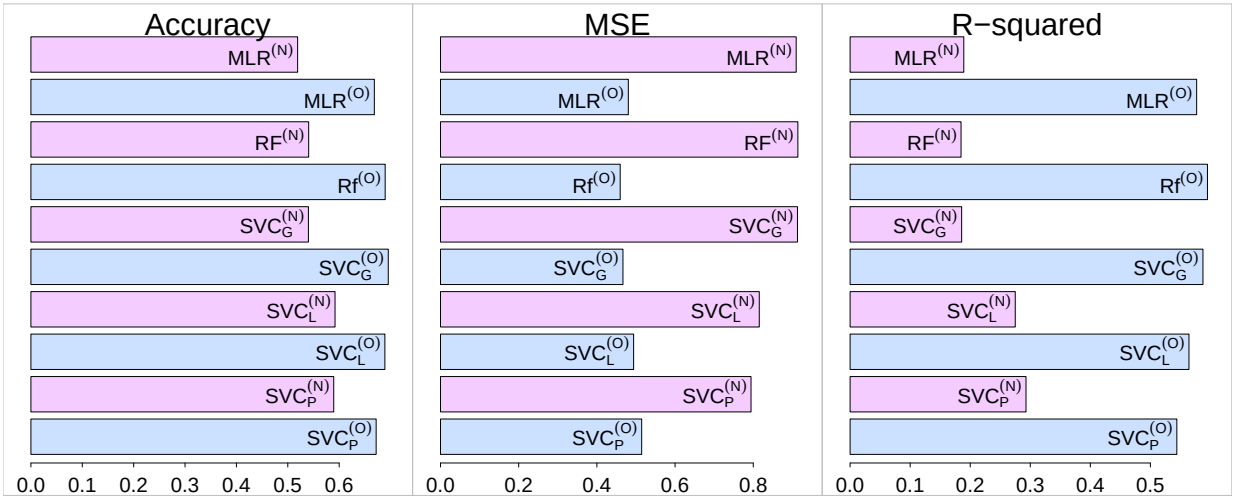


Figure 4.4: **The mean performances of the proposed classifiers based on 10-fold cross validation under the naïve (N) and the optimization (O) approaches for the swallowing segment.**

Figure 4.5 illustrates the mean performances of the proposed classifiers based on 10-fold cross validation under naïve (N) and optimization (O) approaches for the breathing segment. All the optimization approaches of the classifiers perform than those with the naïve approach. Based on the accuracy metric, among the optimization approaches, support vector classifiers with polynomial ($SVC_P^{(O)}$) , Gaussian ($SVC_G^{(O)}$), and polynomial ($SVC_P^{(O)}$) kernels perform more favorable compared to the remaining optimization approaches. The support vector classifier with Gaussian kernel ($SVC_L^{(O)}$) achieves the best mean accuracy

(0.901), MSE (0.109), and R-squared (82.7) among all classifiers. The mean performances of all classifiers stay above 0.70.

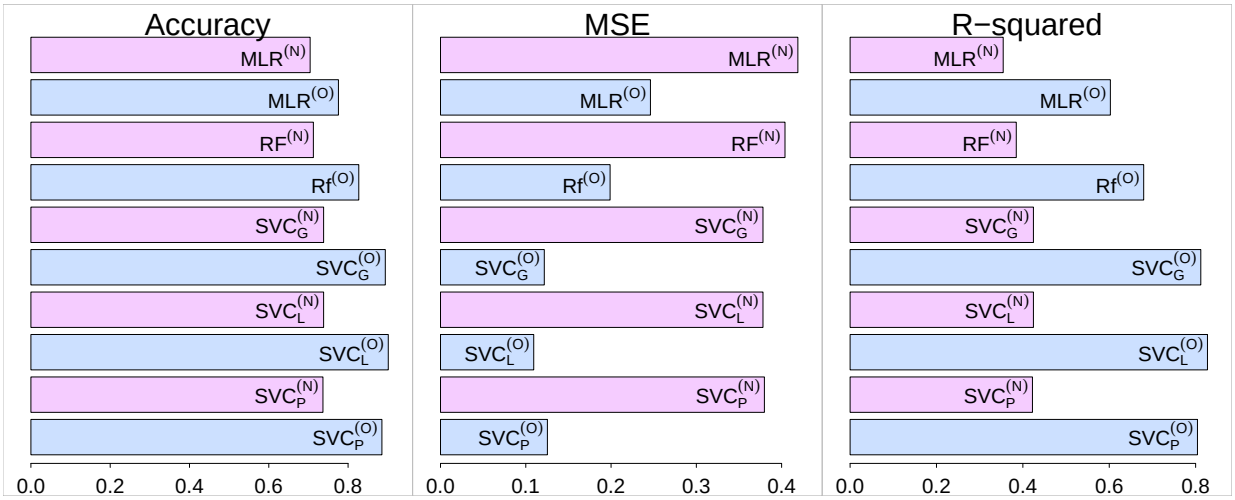
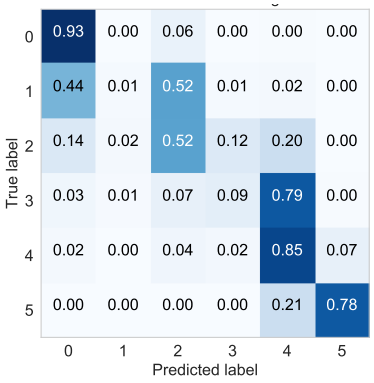


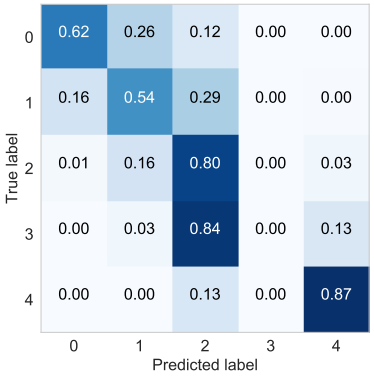
Figure 4.5: **The mean performances of the proposed classifiers based on 10-fold cross validation under the naïve (N) and the optimization (O) approaches for the breathing segment.**

Figure 4.6 illustrates the normalized confusion matrices for the classifiers, achieving the highest accuracy under all segments. The diagonal of confusion matrices denotes the accuracy of a particular tollgate level. For example, among the tollgates levels of the leg segment, the highest accuracy is 93%, achieved for Level 0, while lowest accuracy is 1%, obtained for level 1.

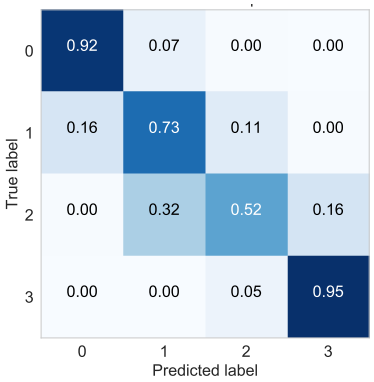
We observe a great variation between the accuracy of different levels. Figure 4.7 illustrates confusion matrices of classifiers $SVC_L^{(O)}$ and $RF^{(O)}$ for the leg tollgates. Although the overall accuracy score of $SVC_L^{(O)}$ is higher than $RF^{(O)}$, the accuracy of the levels resides in a smaller range for $RF^{(O)}$.



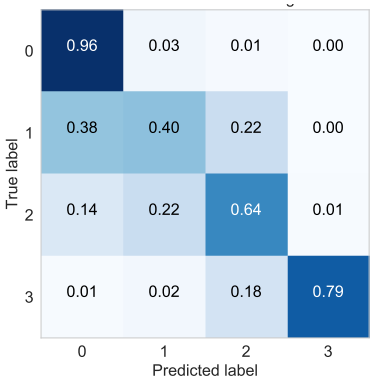
(a) Leg



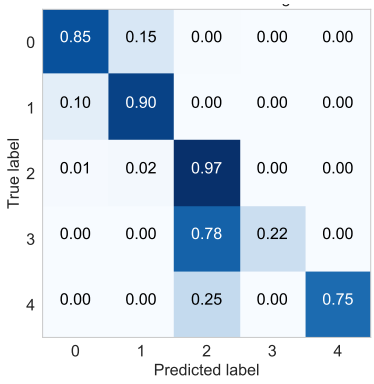
(b) Arm



(c) Speech

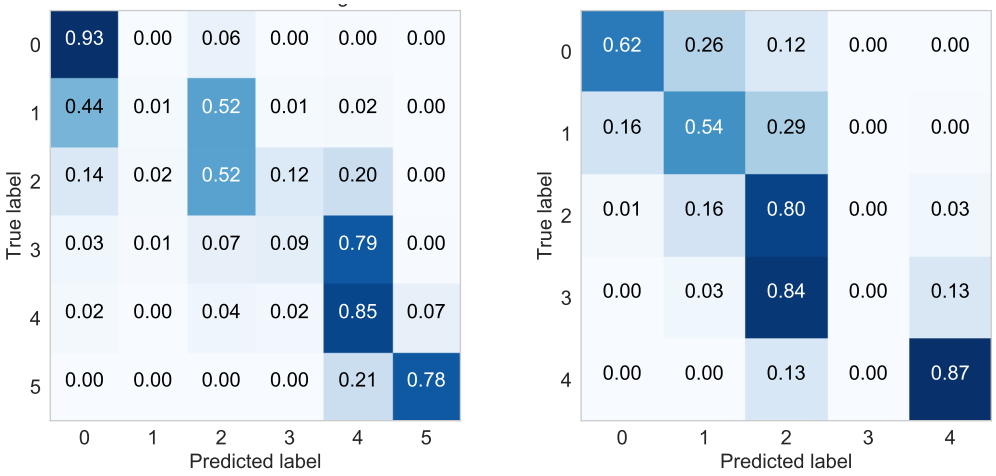


(d) Swallowing



(e) Breathing

Figure 4.6: **The normalized confusion matrices for the classifiers achieving the highest accuracy under all segments.**



(a) $SVC_L^{(O)}$ (b) $RF^{(O)}$

Figure 4.7: **The normalized confusion matrices for the classifiers achieving the highest accuracy under all segments.**

4.4.2 Risk prediction using the estimated tollgates

The estimated tollgates of the patients may allow more comprehensive analysis to be conducted. To illustrate how the estimated tollgate levels can be employed, we build a basic risk prediction model based on the history of patients. Figure 4.4.2 illustrates the probability of reaching a tollgates by the next clinic visit (3 months) for patients having a particular disease history. For example, the probability of staying at Level 1 for the leg segment is 0.5 if the next visit is 10 months after the onset of the disease. This probability steadily increases as the next visit shifts towards later times, indicating a slower pace of progression in the leg segment.

4.5 Conclusion

In this chapter, we tackle the problem of estimating the tollgates reached by patients based on their ALSFRS-R scores. We formulate the estimation problem as a multi-class

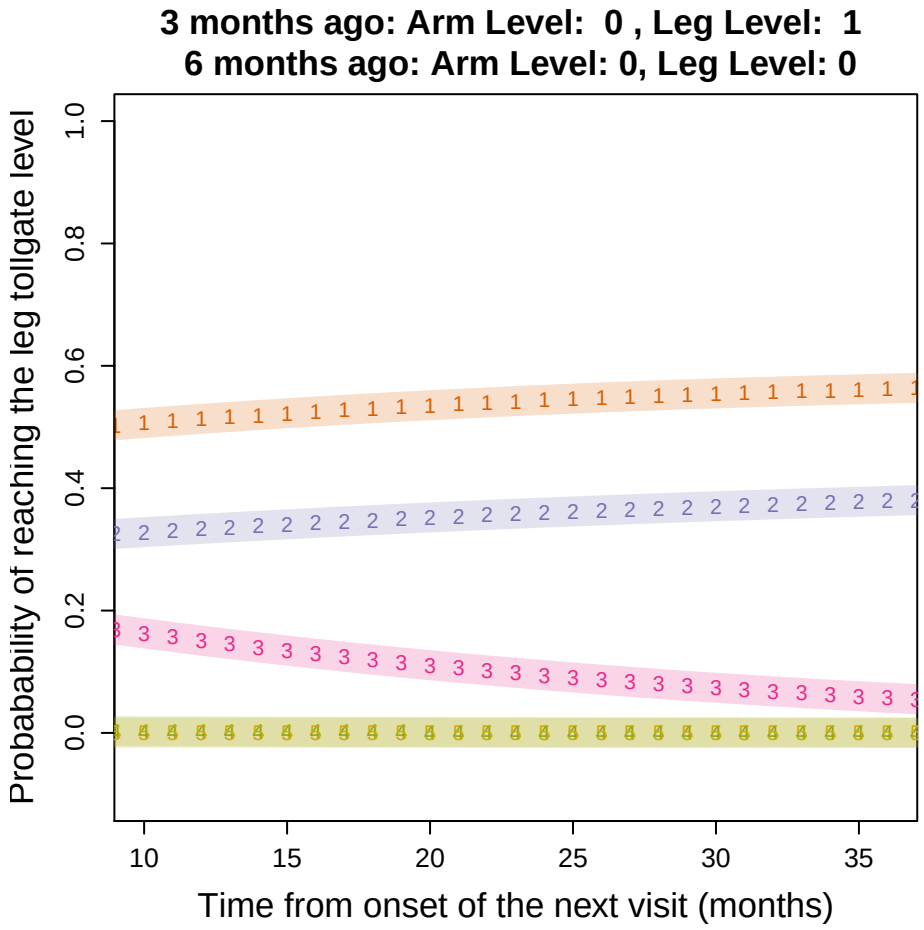


Figure 4.8: **Probabilities of reaching the leg tollgates levels in the next next clinic visit (3 months) over the time of the next clinic visit from the onset time of the disease for patients showing no arm weakness in the last six months, and having no leg weakness six months ago but showing slight leg weakness three months ago.**

classification model subject to a monotonicity constraint. The monotonicity constraint renders the traditional classification techniques inapplicable. Therefore, we propose two alternative approaches (the naïve and optimization) which take the probability estimations of a classification model as the input, and find the tollgates satisfying the monotonicity constraint. As an input for the naïve and optimization approaches, we employ five different classification techniques namely multinomial logistic regression, support vector classifiers with linear, polynomial, and Gaussian kernels, and random forest classifier. All of the classification techniques are used in their default setting, as implemented in scikit-learn package in python environment (Pedregosa et al., 2011).

In term of the performance of the classifiers, the support vector type classifiers generally produce favorable results for most segments. In addition, the performances of the all classifiers with the optimization approach are very similar in most cases. The exception is the swallowing segment, where the optimization approach of the support vector classifier with linear kernel achieves the best accuracy scores, whereas the optimization approach of the support vector classifier with Gaussian kernel outperforms other classifiers for the swallowing segment.

For the leg, arm, and swallowing segments, the highest mean accuracy achieved by a classifier is between 0.6 and 0.7. The highest accuracy obtained reaches above 0.74 and 0.90 for the speech and breathing segments, respectively. The high accuracy achieved in the breathing segment is as a result of better relevance of the the breathing related ALSFRS-R questions (Q10-Q12) to the breathing tollgate definitions. The overall performances of the classifiers seem low especially for arm, leg, and swallowing segments. When we allow one level deviation from the true tollgates levels, accuracy values more than 0.92 are achieved for all segments.

We also provided the normalized confusion matrices for the classifiers who achieves the highest accuracy under a particular segment. We observe that some classifiers fail to predict certain tollgate levels although their overall accuracy scores is the highest among all others. For the leg segment, other classifiers (e.g, $RF^{(O)}$) may result in a normalized matrix having a more balanced diagonal while producing slightly lower overall accuracy. In such cases, it might be better to chose a classifier with a more balanced diagonal.

Our observations summarized above are based on the computational comparisons of several specific classification techniques under two different approaches. The poor performance of the naïve approach can be explained by the fact that when fixing the tollgate level of an observation, the accuracy of the observation greatly depends on the accuracy of the previous estimations for those patients. That is, if there is a misclassification, in any of the observations, it may cause the later observations to be misclassified.

Although obtaining the true performances of the classification techniques would require further analysis, given that the classification parameters are not tuned, we can still conclude that using off-the-shelf classification techniques may produce reasonable results. By tuning the parameters of classifiers, it might be possible to obtain better performances. Another method to increase the overall accuracy could be the by merging some sequential tollgates levels. Despite the fact that this would reduce the specificity granted by the current definitions of the tollgates levels, the estimations and any other model depending on these estimations would have a higher accuracy. Additionally, considering other demographical factors such as age, disease onset age, and time from onset may also improve the performances of the classifiers.

We also conducted some preliminary study on how the historical tollgate information of a patient can be used in assessing the over-time risk of reaching tollgates that has not yet been reached by the patients. However, this model depends on the assumption that the aforementioned tollgate estimation methodology in this chapter estimates the tollgates with a great accuracy. Although the initial results imply a promising contribution, the validity of the proposed risk prediction model not been should be investigated by further studies.

Chapter 5

Conclusion and Future Work

This chapter gives a summary of the methodologies and discusses the results of the proposed models presented in the thesis. We also briefly describe the future work directions and their potential benefits.

In the first chapter, we study the problem of *deriving effective vaccine allocation strategies for pandemic influenza*. This study compares two commonly-used infectious disease modeling techniques from the literature, based on their performance of deriving effective vaccine allocations in the case of an influenza pandemic. Our results show the more realistic but computationally taxing agent-based simulation model produces more favorable vaccine allocation strategies compared to those from a stylized but easy-to-compute compartmental model on general. However, there are certain scenarios where deriving vaccine allocation strategies from the compartmental model may be still effective.

Regarding the comparison of the strategies derived by both modeling techniques and their resulting performances, we conclude that behaviors of two commonly-used artifacts may fail to agree after a valid optimization process, even though the models behave very similarly before the vaccination. Our study highlights the differences in derived vaccine allocation strategies, as well as the significance and magnitudes of the performance differences. Health policy makers may benefit from applying the optimization techniques using the agent-based simulation and compartmental models together when searching for effective vaccine allocation strategies. Our study shines a light on the differences between the

strategies obtained by these two approaches. The results of this study may advise policy makers to design and employ appropriate modeling techniques to enhance their expectation in real life applications.

There are a few limitations of our analysis. First, we examine age-specific vaccine allocation, but do not consider other important risk factors, e.g., chronic medical conditions and pregnancy. Moreover, we use a deterministic numerical optimization algorithm, which does not incorporate the variance in the simulation replications when updating the search direction. Adjusting the optimization algorithm to consider the stochastic model results is a promising future research direction. Additionally, the employed agent based simulation is not fast enough to test very comprehensive allocation strategies, such as time-dependent dynamic vaccination strategies or strategies considering the composition of social mixing groups. Using a faster model that still captures both the network structure and randomness may enable a more robust analysis, which is also left for future studies. We set the number of replications at each iteration of the optimization algorithm to 24 in order to alleviate the computational burden of the analysis. Using a faster agent-based simulation, the number of replications can be increased to reduce the variation among the outcomes. Alternatively, ranking and selection methods can be applied to find the proper number of replications for each allocation strategy, which is left for future research.

In the second chapter, we study the problem of *tollgate-based progression pathways of ALS patients*. In this problem, we introduce a novel mechanism to monitor the disease progression of patients with amyotrophic lateral sclerosis (ALS) disease. This mechanism proposes critical events, named as tollgates, to track patients throughout the course of their disease. Using medical records of 514 patients from the ALS clinic at the Mayo Clinic, Rochester, MM, we present the progression pathways of the patients over multiple visits in terms of these tollgates.

We observe that the proposed mechanism is powerful enough to present the variability among patients. Moreover, we estimate the probabilities of reaching each tollgate, based on the time from their first clinic visits, to illustrate how the disease can effect the functional abilities of patients over time. The analyses and results might be helpful for clinicians to educate their patients when making critical decisions. Analysis in Section 3.3 shows the effect of history on the future progression, where very little of the disease history (the first

clinic visit) is utilized. More comprehensive studies using patients' histories at multiple points over time and several other factors (e.g., age, onset age, onset region) might be more helpful on the prognosis.

Therefore, more comprehensive studies are needed for better characterization of ALS progression. For examples, a risk prediction model created from the whole history and demographics of patients could help more accurately estimate the likelihood of reaching each tollgate over time (Tripepi et al., 2013). There are certain limitations of our study. First, the sample size is limited. This one of the very common problems in the literature because ALS is a rare disease (Al-Chalabi and Hardiman, 2013). Creating comprehensive models may require more data to yield a strong prediction power level. However, for the existing databases which have ALSFRS-R scores but not having tollgate information, the previously mentioned association between ALSFRS-R and the tollgates can be further investigated to estimate the times of patients reaching tollgates in existing databases.

Second, the fact that the disease onset dates are not available restricts our analysis to be based on the first clinic visits of the patients. While not the best method, analyses based on the first clinic visit are not uncommon in medical studies (Gitau et al., 1999; Swinnen and Robberecht, 2014). Third, drop-outs significantly influenced the scope of the analysis. By the end of one year from the clinic visit, approximately 70% of the patients dropped out. To overcome this, we performed a Kaplan-Meier analysis, which is known to be robust against the negative effects of drop-outs(Rich et al., 2010).

In the third chapter, we study the problem of *deriving tollgates from ALSFRS-R scores using multi-class classifiers*. In this study, we tackle the problem of estimating the tollgates reached by patients using their ALSFRS-R scores. We first formulate the problem as a multi-class classification model subject to a monotonicity constraint. Due to the fact that the monotonicity constraint render the traditional classification techniques inapplicable, we propose two alternative approaches (the naïve and optimization). Both of these take the probability estimations from traditional classification methods as the input, and find the tollgates satisfying the monotonicity constraint. As the traditional classification methods, we employ five different classification techniques namely multinomial logistic regression, support vector classifiers with linear, polynomial, and Gaussian kernels, and random forest classifier. All of the classification techniques are used in their default setting as imple-

mented in scikit-learn package in python environment (Pedregosa et al., 2011). The results illustrates that the optimization approach greatly outperforms the naïve approach. Overall, each of the classification methods perform similarly with the same approach group, while the support vector type classifiers are observed to be slightly more powerful.

The results of this study can be especially useful considering the fact that most available databases may not have the detailed patients record to derive the tollgate information. One might criticize the power of the classification models based on the low accuracy scores produced. Although likely to be correct for some cases, the performance of the classifiers are presented without any parameter tuning. By find the a good set of model parameters, the performance of the classification models can be greatly improved. Moreover, the proposed classifiers estimate the tollgates based solely on the ALSFRS-R scores. Incorporating other factors such as demographics (e.g., age, gender) and disease history may improve the performances of the classifiers. The risk prediction model discussed at the end of Chapter 3 is a very promising potential future direction. However, pursuing such an analysis may require improving the accuracy of the classification methods. This is because the error propagated from the estimated tollgates for each body segment may cause a deviation from the actual outcomes, and misinterpretation of the results.

References

Abramson, M. A., Audet, C., Couture, G., Dennis, Jr., J. E., Le Digabel, S., and Tribes, C. (2009). The NOMAD project. Available at https://www.gerad.ca/nomad/.

Al-Chalabi, A. and Hardiman, O. (2013). The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat. Rev. Neurol.*, 9(11):617–628.

Alagoz, O., Ayvaci, M. U., and Linderoth, J. T. (2015). Optimally solving markov decision processes with total expected discounted reward function: Linear programming revisited. *Computers & Industrial Engineering*, 87:311–316.

Aly, M. (2005). Survey on multiclass classification methods. *Neural Netw*, 19.

Anderson, R. M., May, R. M., and Anderson, B. (1991). *Infectious diseases of humans: dynamics and control*. Oxford University Press, Oxford.

Andradóttir, S., Chiu, W., Goldsman, D., and Lee, M. L. (2014). Simulation of influenza propagation: Model development, parameter estimation, and mitigation strategies. *IIE Transactions on Healthcare Systems Engineering*, 4(1):27–48.

Araz, O. M., Galvani, A., and Meyers, L. A. (2012). Geographic prioritization of distributing pandemic influenza vaccines. *Health Care Management Science*, 15(3):175–187.

Arias, E. (2014). United States life tables, 2010. *National Vital Statistics Reports*, 63(7):1–63.

Armon, C. (2008). From clues to mechanisms: Understanding ALS initiation and spread. *Neurology*, 71(12):872–873.

Atkinson, K. A. (1989). *An introduction to numerical analysis (2nd ed.)*. John Wiley & Sons, New York.

Audet, C., Le Digabel, S., and Tribes, C. (2009). NOMAD user guide.

Ayer, T., Alagoz, O., and Stout, N. K. (2012). Or forum-a pomdp approach to personalize mammography screening decisions. *Operations Research*, 60(5):1019–1034.

Balendra, R., Jones, A., Jivraj, N., Steen, I. N., Young, C. a., Shaw, P. J., Turner, M. R., Leigh, P. N., and Al-Chalabi, A. (2015). Use of clinical staging in amyotrophic lateral sclerosis for phase 3 clinical trials. *J. Neurol. Neurosurg. Psychiatry*, 86(1):45–49.

Bansal, S., Grenfell, B. T., and Meyers, L. A. (2007). When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, 4(16):879–891.

Basta, N. E., Chao, D. L., Halloran, M. E., Matrajt, L., and Longini Jr, I. M. (2009). Strategies for pandemic and seasonal influenza vaccination of schoolchildren in the United States. *American Journal of Epidemiology*, 170(6):679–686.

Basta, N. E., Halloran, M. E., Matrajt, L., and Longini, Jr, I. M. (2008). Estimating influenza vaccine efficacy from challenge and community-based study data. *American Journal of Epidemiology*, 168(12):1343–1352.

Bird, S. (2006). NLTK: The Natural Language Toolkit Steven. In *Proc. COLING/ACL Interact. Present. Sess.*, pages 69–72, Morristown, NJ, USA. Association for Computational Linguistics.

Böcker, F. M., Seibold, I., and Neundörfer, B. (1990). Disability in everyday tasks and subjective status of patients with advanced amyotrophic lateral sclerosis. *Fortschr. Neurol. Psychiatr.*, 58(6):224–236.

Boillée, S., Vande Velde, C., and Cleveland, D. W. W. (2006). ALS: A Disease of Motor Neurons and Their Nonneuronal Neighbors.

Brailsford, S., Harper, P., Patel, B., and Pitt, M. (2009). An analysis of the academic literature on simulation and modelling in health care. *Journal of simulation*, 3(3):130–140.

Brandeau, M. L., Sainfort, F., and Pierskalla, W. P. (2004). *Operations research and health care: a handbook of methods and applications*, volume 70. Springer Science & Business Media.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Briem, G. J., Benediktsson, J. A., and Sveinsson, J. R. (2002). Multiple classifiers applied to multisource remote sensing data. *IEEE Trans. Geosci. Remote Sens.*, 40(10):2291–2299.

Bromberg, M. B., Brownell, A. a., Forshew, D. a., and Swenson, M. (2010). A timeline for predicting durable medical equipment needs and interventions for amyotrophic lateral sclerosis patients. *Amyotroph. Lateral Scler.*, 11(1-2):110–115.

Brooks, B., Sanjak, M., Ringel, S., England, J., Brinkmann, J., Pestronk, A., Florence, J., Mitsumoto, H., Szirony, K., Wittes, J., et al. (1996). The amyotrophic lateral sclerosis functional rating scale-assessment of activities of daily living in patients with amyotrophic lateral sclerosis. *Archives of Neurology*, 53(2):141–147.

Cambria, E. and White, B. (2014). Jumping NLP curves: A review of natural language processing research.

Caro, J. J., Möller, J., and Getsios, D. (2010). Discrete event simulation: the preferred technique for health economic evaluations? *Value in health*, 13(8):1056–1060.

Carrat, F., Vergu, E., Ferguson, N. M., Lemaitre, M., Cauchemez, S., Leach, S., and Valleron, A.-J. (2008). Time lines of infection and disease in human influenza: a review of volunteer challenge studies. *American Journal of Epidemiology*, 167(7):775–785.

CDC (2010). Estimates of deaths associated with seasonal influenza - United States, 1976–2007. *Morbidity and Mortality Weekly Report*, 59(33):1057–1062.

CDC (2014). Leading causes of death.

CDC (2016). Vaccine effectiveness - How well does the flu vaccine work? Available at http://www.cdc.gov/flu/about/qa/vaccineeffect.htm.

Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., Nakanishi, A., study group, B. A., complete listing of the BDNF Study Group, A., et al. (1999). The alsfrs-r: a revised als functional rating scale that incorporates assessments of respiratory function. *Journal of the neurological sciences*, 169(1):13–21.

Chao, D. L., Halloran, M. E., Obenchain, V. J., and Longini Jr, I. M. (2010). Flute, a publicly available stochastic influenza epidemic simulation model. *PLoS Computational Biology*, 6(1):e1000656.

Chhatwal, J. and He, T. (2015). Economic evaluations with agent-based modelling: an introduction. *Pharmacoeconomics*, 33(5):423–433.

Chick, S. E., Mamani, H., and Simchi-Levi, D. (2008). Supply chain coordination and influenza vaccination. *Operations Research*, 56(6):1493–1506.

Chio, A., Mora, G., Leone, M., Mazzini, L., Cocito, D., Giordana, M. T., Bottacchi, E., and Mutani, R. (2002). Early symptom progression rate is related to ALS outcome: A prospective population-based study. *Neurology*, 59(1):99–103.

Chow, G. C. and Lin, A.-L. (1971). Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series By Related Series.

Cowling, B. J., Ng, S., Ma, E. S., et al. (2010). Protective efficacy of seasonal influenza vaccination against seasonal and pandemic influenza virus infection during 2009 in Hong Kong. *Clinical Infectious Diseases*, 51(12):1370–1379.

Crammer, K. and Wagner, T. (2012). Volume Regularization for Binary Classification. In *Neural Inf. Process. Syst.*, pages 341–349.

Cronin, S., Hardiman, O., and Traynor, B. J. (2007). Ethnic variation in the incidence of als a systematic review. *Neurology*, 68(13):1002–1007.

Dalgıç, Ö. O., Özaltın, O. Y., Ciccotelli, W. A., and Erenay, F. S. (2017). Deriving effective vaccine allocation strategies for pandemic influenza: Comparison of an agent-based simulation and a compartmental model. *PloS one*, 12(2):e0172261.

Das, T. K., Savachkin, A. A., and Zhu, Y. (2008). A large scale simulation model of pandemic influenza outbreaks for development of dynamic mitigation strategies. *IIE Transactions*, 40(9):893–905.

Dasbach, E. J., Elbasha, E. H., and Insinga, R. P. (2006). Mathematical models for predicting the epidemiologic and economic impact of vaccination against human papillomavirus infection and disease. *Epidemiologic Reviews*, 28(1):88–100.

Deo, S. and Corbett, C. J. (2009). Cournot competition under yield uncertainty: The case of the us influenza vaccine market. *Manufacturing & Service Operations Management*, 11(4):563–576.

Dhamodharan, A. and Proano, R. A. (2012). Determining the optimal vaccine vial size in developing countries: a monte carlo simulation approach. *Health Care Management Science*, 15(3):188–196.

Dobrzykowski, D., Deilami, V. S., Hong, P., and Kim, S.-C. (2014). A structured analysis of operations and supply chain management research in healthcare (1982–2011). *International Journal of Production Economics*, 147:514–530.

Eames, K. T., Tilston, N. L., Brooks-Pollock, E., and Edmunds, W. J. (2012). Measured dynamic social contact patterns explain the spread of H1N1v influenza. *PLoS Computational Biology*, 8(3):e1002425.

Ellis, R. and Petridis, M. (2009). *Research and Development in Intelligent Systems XXVI: Incorporating Applications and Innovations in Intelligent Systems XVII*. Springer Science & Business Media.

Eveborn, P., Flisberg, P., and Rönnqvist, M. (2006). Laps carean operational system for staff planning of home care. *European Journal of Operational Research*, 171(3):962–976.

Feng, W., Kong, N., and Wan, H. (2013). A simulation study of cadaveric liver allocation with a single-score patient prioritization formula. *Journal of Simulation*, 7(2):109–125.

Ferguson, N. M., Cummings, D. A., Fraser, C., Cajka, J. C., Cooley, P. C., and Burke, D. S. (2006). Strategies for mitigating an influenza pandemic. *Nature*, 442(7101):448–452.

Fisher, R. (1936). Iris Data Set.

Fomundam, S. and Herrmann, J. W. (2007). A survey of queuing theory applications in healthcare.

Fone, D., Hollinghurst, S., Temple, M., Round, A., Lester, N., Weightman, A., Roberts, K., Coyle, E., Bevan, G., and Palmer, S. (2003). Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health*, 25(4):325–335.

Fraser, C., Donnelly, C. A., Cauchemez, S., et al. (2009). Pandemic potential of a strain of influenza A (H1N1): early findings. *Science*, 324(5934):1557–1561.

Gautier, G., Verschueren, A., Monnier, A., Attarian, S., Salort-Campana, E., and Pouget, J. (2010). ALS with respiratory onset: Clinical features and effects of non-invasive ventilation on the prognosis. *Amyotroph. Lateral Scler.*, 11(4):379–382.

Germann, T. C., Kadau, K., Jr, I. M. L., and Macken, C. A. (2006). Mitigation strategies for pandemic influenza in the united states. *Proceedings of the National Academy of Sciences*, 103(15):5935–5640.

Gitau, G. K., Perry, B. D., and McDermott, J. J. (1999). The incidence, calf morbidity and mortality due to Theileria parva infections in smallholder dairy farms in Murang'a District, Kenya. *Prev. Vet. Med.*, 39(1):65–79.

Gomeni, R. and Fava, M. (2014). Amyotrophic lateral sclerosis disease progression model. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 15(1-2):119–129.

Günal, M. M. and Pidd, M. (2010). Discrete event simulation for performance modelling in health care: a review of the literature. *Journal of Simulation*, 4(1):42–51.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.*, 46(1-3):389–422.

Harper, P. R., Shahani, A., Gallagher, J., and Bowie, C. (2005). Planning health services with explicit geographical considerations: a stochastic location–allocation approach. *Omega*, 33(2):141–152.

Hastie, T. and Tibshirani, R. (2011). Statistical learning. *Learning*, 2:08.

Heffernan, J., Smith, R., and Wahl, L. (2005). Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, 2(4):281–293.

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653.

Hladish, T., Melamud, E., Barrera, L. A., Galvani, A., and Meyers, L. A. (2012). Epifire: An open source C++ library and application for contact network epidemiology. *BMC Bioinformatics*, 13(76):1–12.

Hobson, E. V. and McDermott, C. J. (2016). Supportive and symptomatic management of amyotrophic lateral sclerosis. *Nature Reviews Neurology*.

Hsu, C.-W. and Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Networks*, 13(2):415–425.

Huang, G. and Li, L. (2009). A mathematical model of infectious diseases. *Annals of Operations Research*, 168(1):41–80.

Jackson, C., Mangtani, P., Hawker, J., Olowokure, B., and Vynnycky, E. (2014). The effects of school closures on influenza outbreaks and pandemics: systematic review of simulation studies. *PLoS ONE*, 9(5):e97297.

Jacobson, S. H., Hall, S. N., and Swisher, J. R. (2006). Discrete-event simulation of health care systems. In *Patient flow: Reducing delay in healthcare delivery*, pages 211–252. Springer.

Kaplan, E. L. and Meier, P. (1958). Nonparametric Estimation from Incomplete Observations. *Source J. Am. Stat. Assoc.*, 53(282):457–481.

Kaplan, R. M. and Bush, J. W. (1982). Health-related quality of life measurement for evaluation research and policy analysis. *Health psychology*, 1(1):61.

Katsaliaki, K. and Mustafee, N. (2011). Applications of simulation within the healthcare context. *Journal of the Operational Research Society*, 62(8):1431–1451.

Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A., Hardiman, O., Burrell, J. R., and Zoing, M. C. (2011). Amyotrophic lateral sclerosis. *The Lancet*, 377(9769):942–955.

Kimura, F., Fujimura, C., Ishida, S., Nakajima, H., Furutama, D., Uehara, H., Shinoda, K., Sugino, M., and Hanafusa, T. (2006). Progression rate of ALSFRS-R at time of diagnosis predicts survival time in ALS. *Neurology*, 66(2):265–267.

Kohn, R. and Ansley, C. F. (1986). Prediction mean squared error for state space models with estimated parameters. *Biometrika*, 73(2):467–473.

Kollewe, K., Mauss, U., Krampfl, K., Petri, S., Dengler, R., and Mohammadi, B. (2008). Alsfrs-r score and its ratio: a useful predictor for als-progression. *Journal of the neurological sciences*, 275(1):69–73.

Küffner, R., Zach, N., Norel, R., Hawe, J., Schoenfeld, D., Wang, L., Li, G., Fang, L., Mackey, L., Hardiman, O., Cudkowicz, M., Sherman, A., Ertaylan, G., Grosse-Wentrup, M., Hothorn, T., van Ligtenberg, J., Macke, J. H., Meyer, T., Schölkopf, B., Tran, L., Vaughan, R., Stolovitzky, G., and Leitner, M. L. (2015). Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression. *Nat. Biotechnol.*, 33(1):51–7.

Kwak, C. and Clayton-Matthews, A. (2002). Multinomial logistic regression. *Nursing research*, 51(6):404–410.

Lakshmi, C. and Iyer, S. A. (2013). Application of queueing theory in health care: A literature review. *Operations Research for Health Care*, 2(1):25–39.

Lane, P. (2008). Handling drop-out in longitudinal clinical trials: a comparison of the LOCF and MMRM approaches. *Pharm. Stat.*, 7(2):93–106.

Lazer, D., Kennedy, R., King, G., and Vespignani, A. (2014). The parable of Google flu: traps in big data analysis. *Science*, 343(6176):1203–1205.

Le Digabel, S. (2011). Algorithm 909: NOMAD: Nonlinear optimization with the MADS algorithm. *ACM Transactions on Mathematical Software*, 37(4):1–15.

Lee, B. Y., Brown, S. T., Korch, G. W., Cooley, P. C., Zimmerman, R. K., Wheaton, W. D., Zimmer, S. M., Grefenstette, J. J., Bailey, R. R., Assi, T.-M., and Burke, D. S. (2010). A computer simulation of vaccine prioritization, allocation, and rationing during the 2009 H1N1 influenza pandemic. *Vaccine*, 28(31):4875 – 4879.

Lloyd, A. L. and May, R. M. (1996). Spatial heterogeneity in epidemic models. *Journal of theoretical biology*, 179(1):1–11.

Loh, W.-Y. (2011). Classification and regression trees. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, 1(1):14–23.

Louwerse, E., Visser, C., Bossuyt, P., and Weverling, G. (1997). Amyotrophic lateral sclerosis: mortality risk during the course of the disease and prognostic factors. *J. Neurol. Sci.*, 152:S10–S17.

Magnus, T., Beck, M., Giess, R., Puls, I., Naumann, M., and Toyka, K. (2002). Disease progression in amyotrophic lateral sclerosis: predictors of survival. *Muscle & nerve*, 25(5):709–714.

Marin, B., Bianchi, E., Pupillo, E., Lunetta, C., Tremolizzo, L., Logroscino, G., Chiò, A., Preux, P. M., and Beghi, E. (2016). Non-self-sufficiency as a primary outcome measure in ALS trials. *Amyotroph. Lateral Scler. Front. Degener.*, 17(1-2):77–84.

Matrajt, L. and Longini Jr, I. M. (2010). Optimizing vaccine allocation at different points in time during an epidemic. *PloS ONE*, 5(11):e13767.

Mazzini, L., Mareschi, K., Ferrero, I., Vassallo, E., Oliveri, G., Nasuelli, N., Oggioni, G. D., Testa, L., and Fagioli, F. (2008). Stem cell treatment in Amyotrophic Lateral Sclerosis. *J. Neurol. Sci.*, 265(1-2):78–83.

Medlock, J. and Galvani, A. P. (2009). Optimizing influenza vaccine distribution. *Science*, 325(5948):1705–1708.

Medlock, J., Meyers, L. A., and Galvani, A. (2009). Optimizing allocation for a delayed influenza vaccination campaign. *PLoS Currents Influenza*.

Merler, S., Ajelli, M., and Rizzo, C. (2009). Age-prioritized use of antivirals during an influenza pandemic. *BMC Infectious Diseases*, 9:117–117.

Miller, R. G., Mitchell, J., and Moore, D. H. (2012). Riluzole for amyotrophic lateral sclerosis (als)/motor neuron disease (mnd). *The Cochrane Library*.

Mitsumoto, H. and Del Bene, M. (2000). Improving the quality of life for people with ALS: the challenge ahead. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.*, 1(5):329–336.

Mittlböck, M. and Schemper, M. (1996). Explained variation for logistic regression. *Stat. Med.*, 15(19):1987–1997.

Molinari, N.-A. M., Ortega-Sanchez, I. R., Messonnier, M. L., Thompson, W. W., Wortley, P. M., Weintraub, E., and Bridges, C. B. (2007). The annual impact of seasonal influenza in the U.S.: measuring disease burden and costs. *Vaccine*, 25(27):5086–5096.

Mollison, D., Isham, V., and Grenfell, B. (1994). Epidemics: models and data. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 115–149.

Montes, J., Levy, G., Albert, S., Kaufmann, P., Buchsbaum, R., Gordon, P. H., and Mitsumoto, H. (2006). Development and evaluation of a self- administered version of the ALSFRS-R. *Neurology*, 67(7):1294–1296.

Morris, J. (2015). Amyotrophic Lateral Sclerosis (ALS) and Related Motor Neuron Diseases: An Overview. *Neurodiagn. J.*, 55(3):180–194.

Mylius, S. D., Hagenaars, T. J., Lugnér, A. K., and Wallinga, J. (2008). Optimal allocation of pandemic influenza vaccine depends on age, risk and timing. *Vaccine*, 26(29):3742–3749.

National Institute of Neurological Disorders and Stroke (2015). Amyotrophic lateral sclerosis (ALS) fact sheet. http://www.ninds.nih.gov/disorders/amyotrophiclateralsclerosis/detail_ALS.htm. Accessed: 2016-09-20.

Nichol, K. L. and Treanor, J. J. (2006). Vaccines for seasonal and pandemic influenza. *Journal of Infectious Diseases*, 194(Supplement 2):S111–S118.

Obermann, M. and Lyon, M. (2015). Financial cost of amyotrophic lateral sclerosis: A case study. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 16(1-2):54–57.

Orrell, R. W., Habgood, J. J., Malaspina, A., Mitchell, J., Greenwood, J., Lane, R. J., and Debelleroche, J. S. (1999). Clinical characteristics of SOD1 gene mutations in UK families with ALS. *J. Neurol. Sci.*, 169(1):56–60.

Özaltin, O. Y., Prokopyev, O. A., Schaefer, A. J., and Roberts, M. S. (2011). Optimizing the societal benefits of the annual influenza vaccine: A stochastic programming approach. *Operations Research*, 59(5):1131–1143.

Parvin, H., Goel, P., and Gautam, N. (2012). An analytic framework to develop policies for testing, prevention, and treatment of two-stage contagious diseases. *Annals of Operations Research*, 196(1):707–735.

Parvin, J. D., Moscona, A., Pan, W., Leider, J., and Palese, P. (1986). Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. *Journal of Virology*, 59(2):377–383.

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925–979.

Patel, R., Longini Jr, I. M., and Halloran, M. E. (2005). Finding optimal vaccination strategies for pandemic influenza using genetic algorithms. *Journal of Theoretical Biology*, 234(2):201–212.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(Oct):2825–2830.

Preater, J. (2002). Queues in health. *Health Care Management Science*, 5(4):283–283.

Raghupathi, W. and Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, 2(1):1.

Rahmandad, H. and Sterman, J. (2008). Heterogeneity and network structure in the dynamics of diffusion: Comapring agent-based and differential equation models. *Management Science*, 54(5):998–1014.

Rais, A. and Viana, A. (2011). Operations research in healthcare: a survey. *International transactions in operational research*, 18(1):1–31.

Rich, J. T., Neely, J. G., Paniello, R. C., Voelker, C. C. J., Nussenbaum, B., and Wang, E. W. (2010). A practical guide to understanding Kaplan-Meier curves.

Robinson, S. (2002). General concepts of quality for discrete-event simulation. *European Journal of Operational Research*, 138(1):103–117.

Robinson, S., Radnor, Z. J., Burgess, N., and Worthington, C. (2012). Simlean: Utilising simulation in the implementation of lean in healthcare. *European Journal of Operational Research*, 219(1):188–197.

Roche, J. C., Rojas-Garcia, R., Scott, K. M., Scotton, W., Ellis, C. E., Burman, R., Wijesekera, L., Turner, M. R., Leigh, P. N., Shaw, C. E., and Al-Chalabi, A. (2012). A proposed staging system for amyotrophic lateral sclerosis. *Brain*, 135(3):847–852.

Rowland, L. P. and Shneider, N. A. (2001). Amyotrophic lateral sclerosis. *New England Journal of Medicine*, 344(22):1688–1700.

Safavian, S. R. and Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, 21(3):660–674.

Schaefer, A. J., Bailey, M. D., Shechter, S. M., and Roberts, M. S. (2005). Modeling medical treatment using markov decision processes. In *Operations research and health care*, pages 593–612. Springer.

Sedgwick, P. (2014). Spearman's rank correlation coefficient. *BMJ*, 349:g7327.

Simon, N. G., Turner, M. R., Vucic, S., Al-Chalabi, A., Shefner, J., Lomen-Hoerth, C., and Kiernan, M. C. (2014). Quantifying disease progression in amyotrophic lateral sclerosis. *Annals of neurology*, 76(5):643–657.

Simpao, A. F., Ahumada, L. M., Gálvez, J. A., and Rehman, M. A. (2014). A review of analytics and clinical informatics in health care. *Journal of medical systems*, 38(4):1–7.

Sobolev, B. G., Sanchez, V., and Vasilakis, C. (2011). Systematic review of the use of computer simulation modeling of patient flow in surgical care. *Journal of medical systems*, 35(1):1–16.

Srinivasan, U. and Arunasalam, B. (2013). Leveraging big data analytics to reduce healthcare costs. *IT Professional*, 15(6):21–28.

Struchiner, C. J., Longini Jr, I. M., and Halloran, M. E. (2010). *Design and analysis of vaccine studies*. Springer, New York.

Swinnen, B. and Robberecht, W. (2014). The phenotypic variability of amyotrophic lateral sclerosis. *Nat. Rev. Neurol.*, 10(11):661–670.

Talbot, H. K., Zhu, Y., Chen, Q., Williams, J. V., Thompson, M. G., and Griffin, M. R. (2013). Effectiveness of influenza vaccine for preventing laboratory-confirmed influenza hospitalizations in adults, 2011-2012 influenza season. *Clinical Infectious Diseases*, 56(12):1774–1777.

Taubenberger, J. K. and Morens, D. M. (2006). 1918 Influenza: the mother of all pandemics. *Emerging Infectious Diseases*, 12(1):15–22.

Therneau, T. M. and Lumley, T. (2015). Package survival '. *R Top. Doc.*, pages 2–37.

Tripepi, G., Heinze, G., Jager, K. J., Stel, V. S., Dekker, F. W., and Zoccali, C. (2013). Risk prediction models. *Nephrol. Dial. Transplant.*, 28(8):1975–1980.

Turner, M. R., Scaber, J., Goodfellow, J. A., Lord, M. E., Marsden, R., and Talbot, K. (2010). The diagnostic pathway and prognosis in bulbar-onset amyotrophic lateral sclerosis. *J. Neurol. Sci.*, 294(1-2):81–85.

United States Census Bureau (2014). Population estimates. Available at https://www.census.gov/popest/data/datasets.html.

Uribe-Sánchez, A., Savachkin, A., Santana, A., Prieto-Santa, D., and Das, T. K. (2011). A predictive decision-aid methodology for dynamic mitigation of influenza pandemics. *OR Spectrum*, 33(3):751–786.

Van Buuren, S. and Groothuis-Oudshoorn, K. (2011). Multivariate Imputation by Chained Equations. *J. Stat. Softw.*, 45(3):1–67.

Vapnik, V. (2013). *The nature of statistical learning theory*. Springer science & business media.

Ventresca, M. and Aleman, D. (2013). Evaluation of strategies to mitigate contagion spread using social network characteristics. *Social Networks*, 35(1):75–88.

Wang, Z. (2009). The convergence of health care expenditure in the us states. *Health Economics*, 18(1):55–70.

WHO (2009). Pandemic influenza vaccine manufacturing process and timeline.

Wu, J. T., Wein, L. M., and Perelson, A. S. (2005). Optimization of influenza vaccine selection. *Operations Research*, 53(3):456–476.

Yaesoubi, R. and Cohen, T. (2011). Dynamic health policies for controlling the spread of emerging infections: influenza as an example. *PLoS ONE*, 6(9):1–11.

Yarmand, H., Ivy, J. S., Denton, B., and Lloyd, A. L. (2014). Optimal two-phase vaccine allocation to geographically different regions under uncertainty. *European Journal of Operational Research*, 233(1):208–219.

Zaric, G. S. and Brandeau, M. L. (2001). Resource allocation for epidemic control over short time horizons. *Mathematical Biosciences*, 171(1):33–58.

APPENDICES

Appendix A

Supporting information for Chapter 2

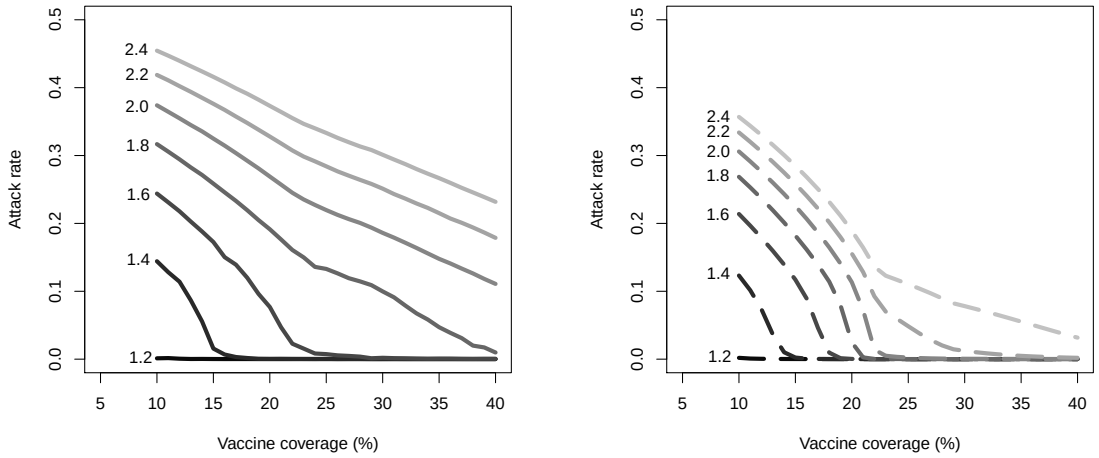
A.1 Age-specific contact rates

Table A.1: **Age-specific contact rates† from age group i to age group j (Φ_{ij}) used in the SEIR model after the calibration process.**

i \ j	0-4	5-18	19-29	30-64	65+
0-4	2.98	32.69	4.71	0.20	0.62
5-18	9.62	35.49	13.36	48.31	8.44
19-29	2.41	23.28	1.26	0.55	0.40
30-64	0.03	22.75	0.15	1.56	2.43
65+	0.35	16.13	0.43	9.85	0.48

†:We assume that the total number of daily contacts between age groups i and j is symmetric. To calculate the daily contact rate from age group i to j, we divided the total number of daily contacts by the population of age group i. So, the age-specific contact rates are asymmetric because the population of age groups i and j are different.

A.2 Overall attack rates for different vaccine coverage levels and R_0 values



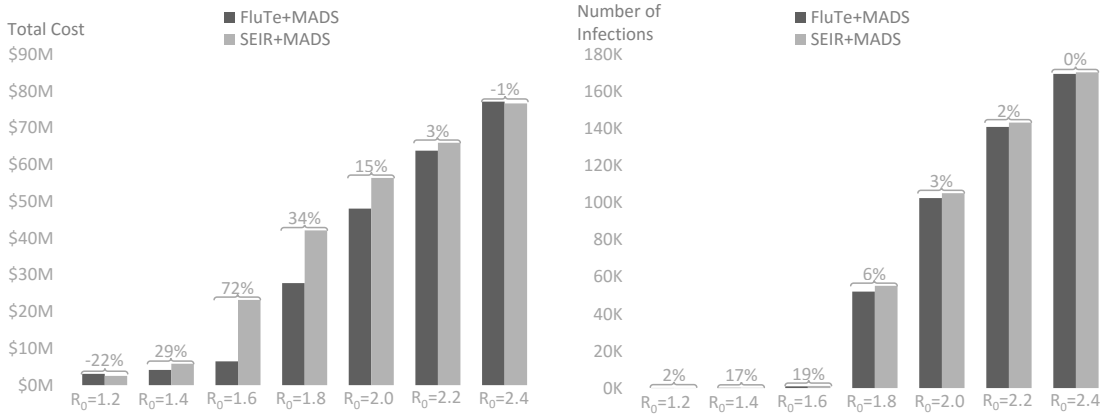
(a) FluTe

(b) SEIR-SEIR model match

Figure A.1: **Overall attack rates for different vaccine coverage levels and R_0 values.**

Both Flute and the SEIR model evaluate the same vaccine allocation policy for all vaccine coverage levels. In this policy, available vaccine stocks are allocated first to school children, then to preschool children, and last to young adults.

A.3 The objective values of vaccine allocation strategies derived by FluTe+MADS and SEIR+MADS under various R_0



(a) Total cost

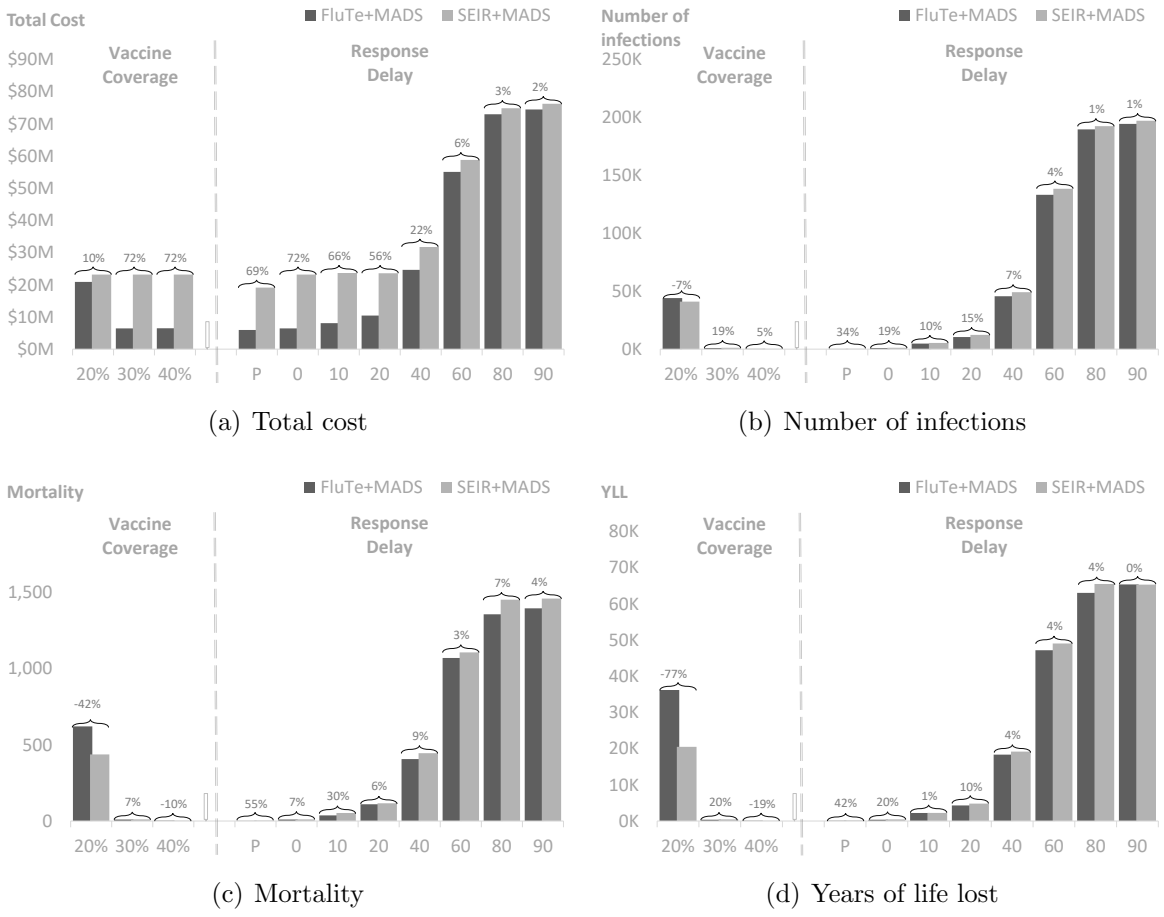
(b) Number of infections

(c) Mortality

(d) Years of life lost

Figure A.2: **Objective values of vaccine allocation strategies derived by FluTe+MADS and SEIR+MADS under various R_0 values (30% vaccine coverage, no delay in response time).**

A.4 Objective values of vaccine allocation strategies for different coverage scenarios (R_0=1.6, no delay in response time) and response time scenarios



(a) Total cost

(b) Number of infections

(c) Mortality

(d) Years of life lost

Figure A.3: **Objective values of vaccine allocation strategies for different coverage scenarios (R_0=1.6, no delay in response time) and response time scenarios (R_0=1.6, 30% vaccine coverage).**

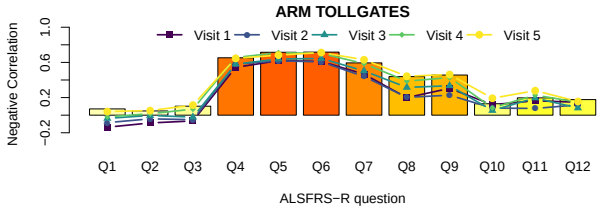
Appendix B

Supporting information for Chapter 3

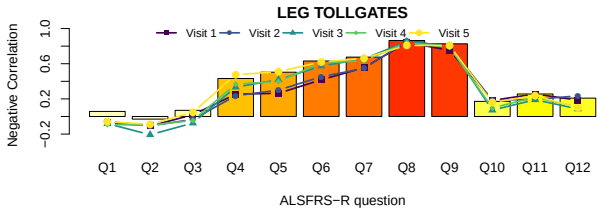
B.1 Correlation between tollgates and ALSFRS-R

We examined the relationship between tollgates and the ALSFRS-R scores by performing a Spearman's rank correlation analysis (Sedgwick, 2014).

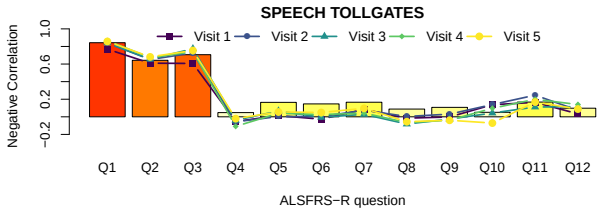
We observed a strong negative correlation between the tollgates and the scores of ALSFRS-R questions and magnitudes of the correlations present a time-invariant. For the arm segment tollgates, most correlated (>0.6) ALSFRS-R questions were Q4-Q6 with negative correlations of 0.56, 0.62, 0.62, respectively. For the leg segment tollgates, most correlated ALSFRS-R questions are Q7-Q9 with respective negative correlations of 0.60, 0.78, 0.73. Q1-Q3 are the most correlated ALSFRS-R questions for the tollgates of both the speech and swallowing segments with the respective negative correlations of 0.79, 0.60, 0.67 for speech and 0.69, 0.55, 0.75 for swallowing. Finally, for the breathing segment tollgates, Q10-Q12 are the most correlated questions with a negative correlation of 0.52, 0.44, 0.71, respectively.



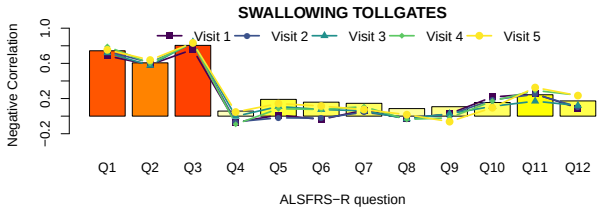
(a) Arm



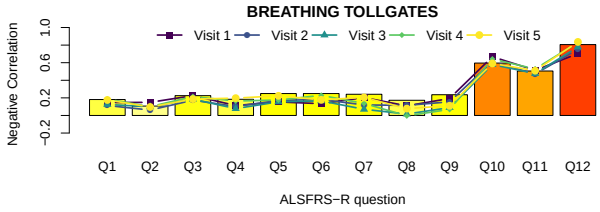
(b) Leg



(c) Speech



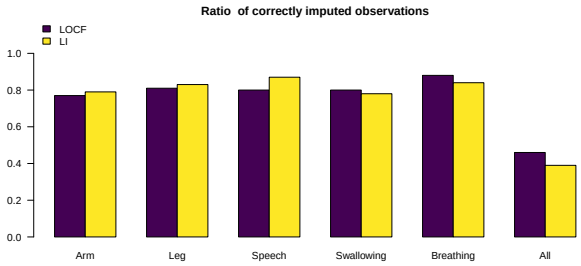
(d) Swallowing



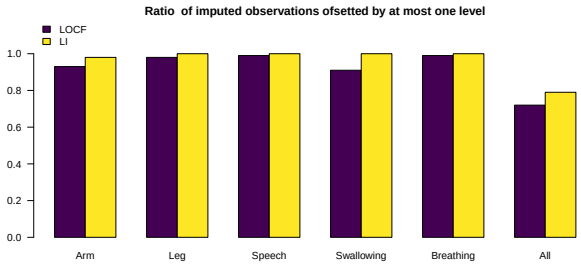
(e) Breathing

Figure B.1: **Correlation analysis between the questions of ALSFRS-R and tollgate levels for each segment.**

B.2 Comparison of missing data imputation methods



(a) Average ratio of correctly imputed values



(b) Average ratio of correctly imputed observations when at most on level offset is allowed

Figure B.2: **Average performances of imputation methods using 100 randomly selected sample sets**.