A multi-state model for pricing critical illness insurance products

by

Jingdan Li

B.Sc., Central University of Finance and Economics, 2017

Project Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Statistics and Actuarial Science
Faculty of Science

SIMON FRASER UNIVERSITY
Summer 2019

Approval

Name:	**Jingdan Li**
Degree:	**Master of Science (Actuarial Science)**
Title:	**A multi-state model for pricing critical illness insurance products**
Examining Committee:	**Chair:** Jinko Graham Professor
	Yi Lu Senior Supervisor Professor
	Gary Parker Supervisor Associate Professor
	Cary Chi-Liang Tsai Internal Examiner Professor
Date Defended:	**August 21, 2019**

Abstract

Due to increasing cases of cancer and other severe illnesses, there is a great demand of critical illness insurance products. This project introduces a Markovian multi-state model based on popular critical illness plans to describe the policyholder's health condition over time, which includes being diagnosed with certain dread diseases such as cancer, stroke and heart attack. Critical illness insurance products with life insurance or other optional riders are considered. Following the idea of Baione and Levantesi (Insurance: Mathematics and Economics, 58: 174-184, 2014), we focus on the method of modelling mortality rates, estimating transition probabilities with Canadian prevalence rates and incidence rates of covered illnesses, and calculating premium rates based on the multi-state model. A comparison of transition intensities under various mortality models and premium rates for critical illness policies under several graduation approaches are also illustrated.

Keywords: Critical illness insurance; Markovian multi-state model; Prevalence rate; Mortality model; Graduation function

Dedication

To my beloved parents.

Acknowledgements

I would like to express my sincere gratitude to all those who have encouraged and supported me during my graduate study. First and foremost, I want to thank my senior supervisor Dr. Yi Lu for her kindness, guidance and encouragement in my research project. I would not be able to finish this project without her great support. I owe my deep gratitude to her for the assistance in graduate study, research and writing in the past two years.

I'm truly grateful to Dr. Gary Parker and Dr. Cary Tsai, for attending my defense, and for the thoroughly reviewing and insightful comments on the project. I would also like to thank Dr. Jinko Graham for being the chair of my examining committee.

My thanks are further extended to all faculty members and staff in our department for their help in study and life. In addition, I would like to express my great appreciation to my fellow students and friends, for your support and advice during the graduate study.

Last, my thanks would go to my family for their invaluable love and encourage all these years.

Table of Contents

Approval ii

Abstract iii

Dedication iv

Acknowledgements v

Table of Contents vi

List of Tables viii

List of Figures ix

1 Introduction 1

1.1 Background . . . 1
1.2 Overview of Previous Studies . . . 2
1.3 Outline . . . 3

2 Assumptions and Models 5

2.1 Critical Illness Insurance Products . . . 5
2.2 Assumptions . . . 5
2.3 Multi-state Models . . . 7
2.3.1 Transition Probabilities . . . 7
2.3.2 Transition Intensities . . . 9
2.4 Pricing Models . . . 16

3 The Data 18

3.1 Prevalence Rate . . . 19
3.2 Incidence Rate . . . 20
3.3 Mortality Rate . . . 22

4 Empirical Analysis 25

4.1 Mortality Models . . . 25

4.2 Transition Intensities 27
4.3 Graduation of Intensities 29
4.4 Comparison of Premium Rates 36
4.5 Impact of the Force of Interest 39

5 Conclusion **43**

Bibliography **46**

Appendix A Graduated incidence rates and transition intensities **48**

List of Tables

Table 2.1 Benefit payouts for different transitions under three types of policies . 7

Table 3.1 Prevalence rates of stroke and heart attack in 2015 19
Table 3.2 Prevalence rates of cancer (per 100,000) 19
Table 3.3 Rates of new cases of primary cancer in each age group (per 100,000) 20
Table 3.4 Population estimates in each age group 21
Table 3.5 Incidence rates of cancer in each age group 21
Table 3.6 Incidence rates of stroke and heart attack in each age group 22
Table 3.7 Mortality rates of total death and by causes 23
Table 3.8 Mortality rates of transportation accidents (per 100,000) 24

Table 4.1 Estimated parameters of mortality models (from Healthy to Death) . 26
Table 4.2 Estimated parameters of mortality models (from Stroke to Death due to Stroke) 27
Table 4.3 Estimated parameters of mortality models (from Heart Attack to Death due to Heart Attack) 28
Table 4.4 Estimated parameters of mortality models (from Cancer to Death due to Cancer) 29
Table 4.5 Transition intensities for males under Gompertz and Weibull models . 29
Table 4.6 Transition intensities for females under Gompertz and Weibull models 31
Table 4.7 NSP rates under three graduation approaches 38
Table 4.8 Male premium rates under different scenarios of the force of interest . 40
Table 4.9 Percent change of male premium rates from “Base” 40
Table 4.10 Female premium rates under different scenarios of the force of interest 41
Table 4.11 Percent change of female premium rates from “Base” 41

List of Figures

Figure 2.1 States and transitions for a Full Accelerated policy 6

Figure 3.1 Mortality rates of transportation accidents 24

Figure 4.1 Estimated mortality models (Gompertz v.s. Weibull) 30
Figure 4.2 Comparison of transition intensities under Gompertz and Weibull models 31
Figure 4.3 Graduated prevalence rates of stroke and heart attack 33
Figure 4.4 Graduated transition intensities of stroke and heart attack 34
Figure 4.5 Transition intensities of stroke and heart attack under different approaches 35
Figure 4.6 NSP rates for males and females under three approaches 39

Chapter 1

Introduction

1.1 Background

According to a report by the Canadian Cancer Society[1], nearly half of Canadians will develop cancer in their lifetime, while one in four will die from cancer. The total number of cancer cases in both sexes has increased in the past 30 years. It also shows that cancer and heart disease, the leading causes of death, are responsible for around 50% of all deaths. Therefore, there is a great demand of critical illness insurance, also known as dread disease insurance, that can help insureds with a lump sum payout after being diagnosed with a particular illness such as cancer.

A typical critical illness insurance plan provides a lump sum from $25,000 to $2 million to the policyholder, if the policyholder is diagnosed with one of the covered conditions that include more than 30 types of critical illnesses such as cancer (life threatening), coma, blindness or deafness. For Canadian critical illness products, the issuing age is usually from 18 to 65 years old and the term of the plan (i.e., a period of time for which a policy is in effect) can be flexible; insureds normally have options at issue for a fixed term such as a 10-year term or a term to a specified age such as a term to age 75. Besides, there are some optional coverages such as a partial benefit payout worth 20% of the benefit amount (face amount) for diagnosis of an early stage cancer or 25% for the loss of independence or a non-life threatening illness. Premium payments for critical illness plans are also flexible. For example, a policyholder can choose to pay a single premium or pay a level premium monthly, quarterly, semi-annually or annually.

Similarly, critical illness products issued in the United States typically provide lump-sum benefits to insureds on a diagnostic of cancer, advanced Alzheimer's disease or other critical illnesses. According to Assurity Life Insurance Company[2], 66% of all U.S. bankruptcies are tied to medical expenses, which indicates the importance of critical illness insurance in the United States. According to PartnerRe (2015), critical illness products are also popular in

[1] Canadian Cancer Society, Canadian Cancer Statistics, 2017

[2] https://www.assurity.com/products/critical-illness

Asia and show a rise in demand and innovations[3]. The coverage of critical illness plans in Asia are similar to that in Canada. In addition, multi-pay critical illness insurance products are commonly provided by Asian life insurance companies; the policy pays the benefit amounts to the policyholder multiple times for claims of different types of coverage.

In this project, we study a pricing model for critical illness insurance with Canadian health data. A Markovian multiple state model (also called multi-state model) is introduced to describe health conditions of an insured being Healthy, Cancer, Stroke, Heart Attack and Dead over time. Canadian mortality rates, prevalence rates and incidence rates data are used to estimate the transition probabilities from Healthy status to status of Cancer, Stroke, Heart Attack or Death. Definitions of the prevalence rate and incidence rate as well as details of the data we use are given in Section 2.3.2 and Chapter 3, respectively. When calculating transition intensities relating to death, both Weibull and Gompertz models are considered as continuous-time mortality models and comparisons of transition intensities under these two mortality models are presented. Following Baione and Levantesi (2014), we discuss a method using prevalence rates to obtain transition intensities and compare with using incidence rates directly as the transition intensities. Then, we propose graduation models on prevalence rates and transition intensities, such as power function or cubic spline interpolant, and examine the impact of these graduation models on transition intensities. A comparison of premium rates under different approaches of graduated transition intensities by age is provided. Finally, we present a sensitivity study of the effect of forces of interest on the premium rates of our pricing model.

1.2 Overview of Previous Studies

Markovian multiple state model is commonly used to describe the transitions of health conditions from healthy to some illnesses or death. For example, Haberman (1984) introduces a multi-state model for the states of healthy, an illness Z and death. Pitacco (1995) uses Markovian multi-state models for permanent health insurance and disability insurance. Haberman and Pitacco (1998) also mention some semi-Markov models and particular calculation procedures used in disability insurance. Dash and Grimshaw (1993) propose a method for calculating transition probabilities from the incidence rate of a dread disease at age x, defined as "a measure of the probability of being diagnosed with a Dread Disease in the year of age x to $x+1$". Meira-Machado et al. (2009) study multi-state models and some non-Markov models for health insurance with conditions of healthy, death and stages of illness. Christiansen (2012) illustrates Markov and semi-Markov models for health insurance policies and provides examples of Markov models for disability insurance, critical illness insurance, long term care insurance, and German private health insurance. Dickson et al.

[3]https://partnerre.com/opinions_research/growth-potential-of-critical-illness-insurance-in-asia/

(2013) explains the assumptions, multiple state models and premiums for the permanent disability insurance and disability income insurance in detail.

Based on the Markovian multiple state models, Baione and Levantesi (2014) present an idea of finding transition probabilities of the multi-state model using the prevalence rate (or called period prevalence rate), which is defined as the number of ill persons during a period over the total cohort size of the same period. Since it is hard to find the incidence rate of dread diseases in Italian insurance market, transition intensities of Markov models are estimated from prevalence rates and mortality rates by specified causes. Then, they apply this method into the pricing of critical illness insurance with Italian health data. Haberman (1984) defines different types of prevalence rates such as point prevalence rates and accumulative prevalence rates, and provides relationships between prevalence rates and transition intensities. Extra mortality is introduced to calculate the transition intensities from a dread disease to death due to other reasons (i.e., the occurrence of death is not a result of the particular dread disease). Haberman (1983) gives the explanation and calculation of extra mortality, defined as the difference between actual deaths and expected deaths from the number of persons exposed to risk divided by expected deaths. Dash and Grimshaw (1993) specifically explain the extra mortality and provide the extra mortality of some specified conditions after a dread disease.

Recently, Baione and Levantesi (2018) compare Weibull and Gompertz models as the mortality models with applications to the pricing of critical illness insurance. Juckett and Rosenberg (1993) also examine Weibull and Gompertz models for distributions of mortality rates, while Wilson (1994) fits the mortality rate data to Logistic functions. The Lee-Carter model (Lee and Carter, 1992) is also well known and widely used as a non-parametric mortality model. Booth and Tickle (2008) review and summarize the methods of mortality modelling and forecasting where the Lee-Carter model and generalized linear model (GLM) are used for forecasting mortality rates.

Health data, or more specifically disease or death related data, collected by government or not-for-profit organizations are typically presented in certain age groups. Graduation methods can be used to estimate rates (e.g., incidence rates) for each age in order to improve the smoothness of pricing. By Klugman et al. (2008), the objective of graduation is "to fit a smooth curve through a set of data according to some specified criteria". Haberman and Pitacco (1998) consider exponential functions as the graduation function for estimating transition intensities. Baione and Levantesi (2014) show the advantages of cubic spline as the graduation function.

1.3 Outline

The project report is arranged as follows. Chapter 2 provides the assumption of the insurance product, health status and the Markovian multi-state model for pricing. The states of the

Markov model and the choices of the insurance plans are given in detail. The Canadian health data used in this project are described in Chapter 3. In Chapter 4, an empirical analysis is presented based on our multi-state model assumption and two mortality models with Canadian health data. For the assumed insurance plan and the health data, we discuss the impacts of different mortality models on the transition intensities. Later, we compare net single premiums under different graduation functions on transition intensities, and consider the impact of the force of interest on our proposed pricing model. Finally, conclusions are given with discussions on limitations and future extensions in Chapter 5.

Chapter 2

Assumptions and Models

2.1 Critical Illness Insurance Products

Basic critical illness (CI) plans always provide coverage of major types of cancer, while some advanced plans provide a benefit for cancers at specific stages of diagnosis and for some other illnesses besides cancer. For example, Sunlife Financial offers several types of CI plans; a Basic Plan covers cancer only, while an Enhanced Plan covers cancer, stroke and heart attack and a Comfort Plan covers four additional critical health conditions. Aviva has a critical illness plan with 20% of the sum assured for the diagnosis of an early stage of covered cancers or minor illnesses and this benefit payment will not terminate the insurance contract. Aviva also provides a MultiPay CI plan that provides benefits for early stage CI, severe stage CI, first-time CI and re-diagnosed cancer with a limit of 5 claims. Royal Bank of Canada (RBC) offers a critical illness insurance that covers 32 types of critical illness with an optional coverage of scheduled increasing benefits, where the benefit amount is increased by 20% per year for 10 years. There are some plans that accept insureds with existing health conditions such as diabetes or pre-diabetes.

Some insureds may consider to buy life insurance policies along with CI plans. In fact, many CI plans also provide optional riders of term or whole life insurance for all causes of death. A rider of accidental death or transportation accidental death is also widely offered. Similar to some life insurance products, an enhanced CI plan will return a proportion of paid premiums at the end of the term as a cash reward if no claim is made during the policy term.

2.2 Assumptions

In this project, we consider a standard critical illness plan which covers three types of dreaded diseases, Cancer, Heart Attack and Stroke. We assume that a net single premium is charged with this product and the policyholder will receive a lump sum payout once diagnosed with one of the covered illnesses or at the time of death. To show the effectiveness of

our model and approaches to pricing the products, both Stand Alone and Full Accelerated policies are considered. The Stand Alone policy pays a lump sum to the insured at the time he/she is diagnosed with one of the illnesses (cancer, heart attack or stroke), while the Full Accelerated policy pays a lump sum to the policyholder either he/she is diagnosed with one of the illnesses or he/she is dead. Insureds can choose different benefit amounts for Cancer, Stroke, Heart Attack and Death. A Full Accelerated policy can be viewed as a Stand Alone policy with life insurance as a rider, in which the cause of death can be a cause other than the three covered illnesses. Death due to transportation accidents can also be covered as a rider on the Full Accelerated policy, which is included in Dead due to other causes (D) in Figure 2.1. For a Full Accelerated policy, we can illustrate the transitions of an insured's health conditions in the following diagram:

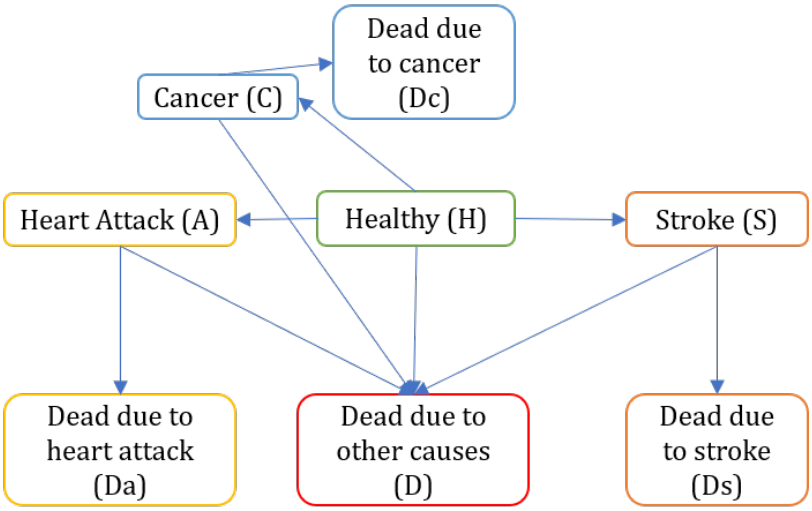

Figure 2.1: States and transitions for a Full Accelerated policy

We use the following abbreviations to represent the states displayed in Figure 2.1:
H = Healthy
D = Dead due to other causes
S = Stroke
Ds = Dead due to stroke
A = Heart attack
Da = Dead due to heart attack
C = Cancer
Dc = Dead due to cancer

It is possible that an insured with cancer dies due to heart attack or stroke, or an insured who survives from heart attack (stroke) dies due to cancer or stroke (heart attack). Since

it is hard to find corresponding reliable mortality data, we ignore these transitions in this project for simplicity.

More specifically, in this project we consider an n-year Stand Alone (SA) insurance plan issued to an insured (male or female) at an age between 35 and 55 while he/she is in Healthy (H) status. The insured will receive a lump sum payment of B_C at the time when the insured transitions from Healthy (H) to Cancer (C), B_S for transitioning to Stroke (S), or B_A for transitioning to Heart Attack (A) during the term. For an n-year Full Accelerated (FA) plan issued to a healthy insured aged x, he/she receives the same lump sum benefit as a Stand Alone policy if he/she transitions from state H to state C/S/A, and of B_D if he/she transitions from state H to state D. With the rider of transportation accidental death (with Rider), state D is divided into two states: state D_1 represents death due to causes other than transportation accidents in state D and state D_2 represents death due to transportation accidents. We assume in this case, the insured receives a lump sum benefit of B_D for transitioning from state H to state D_1, or receive two times of B_D for transitioning to state D_2, and all other benefit payments are the same as those under the Full Accelerated plan. The specific benefit payouts under the three policies are listed in Table 2.1.

Transition	Policy type		
	SA	FA	with Rider
H $\to$ C	B_C	B_C	B_C
H $\to$ S	B_S	B_S	B_S
H $\to$ A	B_A	B_A	B_A
H $\to$ D_1	0	B_D	B_D
H $\to$ D_2	0	B_D	$2\times B_D$

Table 2.1: Benefit payouts for different transitions under three types of policies

2.3 Multi-state Models

2.3.1 Transition Probabilities

In this section, a Markovian multiple state model is introduced to describe the health conditions of an insured. Multi-state models in continuous-time and discrete-time cases are defined in Haberman and Pitacco (1998) and Dickson et al. (2013) and formulae of transition probabilities are also derived. Following their formulation, we then write down transition probabilities for our proposed multi-state model. Consider an insured aged x at $t = 0$ (policy issuing time) in healthy condition. Define random variable $Y(t)$ at time t, for $t \geqslant 0$, which takes one of the values in $\mathcal{E}$ ={H, C, S, A, D, Dc, Ds, Da}, called state space. That is,

$$Y(t) = i, \quad t \geqslant 0, i \in \mathcal{E},$$

implying that at time t the insured aged $x+t$ is either healthy alive (H), or is diagnosed having one of the diseases (C, S, A), or is dead due to one of the various causes described in our model (Dc, Ds, Da, D).

The set $\{Y(t); t \geqslant 0\}$ is called a continuous-time stochastic process which describes the status of an insured in the multi-state model. We further assume that it satisfies the Markov property so the process $\{Y(t); t \geqslant 0\}$ is a Markov process. That is, intuitively, the probabilities of future events (for the insured being in a particular state) for the process are completely determined by knowing the current state (current status of the insured) of the process. Detailed explanations of the Markov property for this type of multi-state models can be found in Dickson et al. (2013).

The transition probability ${}_tp_x^{ij}$ is defined as the conditional probability of an insured aged x staying in state i at age x and in state j at $x+t$, which can be expressed as

$$ {}_tp_x^{ij} = \Pr\left\{Y(x+t) = j | Y(x) = i\right\}, \quad x \geqslant 0,\, t \geqslant 0,\, i, j \in \mathcal{E}. $$

Then, the transition intensity of an insured aged x from state i to state j is defined as

$$ \mu^{ij}(x) = \lim_{t \to 0_+} \frac{{}_tp_x^{ij}}{t}, \quad i \neq j,\, x \geqslant 0,\, i, j \in \mathcal{E}. $$

For the model described in Figure 2.1, a healthy insured will not transition to state Dc, Ds, Da, directly. Moreover, since the contract will be terminated after the benefit payment is made, the probability of transitioning from one dread disease to another is zero. We also do not consider the recovery of the insured in this insurance plan so that the probability of transitioning from a dread disease to healthy is assumed zero in our project. Clearly, all death states are absorbing states, which means that the probabilities of transitioning from these death states to other states are all 0. Then, the transition probability matrix for an insured from age x to $x+t$ can be expressed as

$$ {}_tP_x = \begin{array}{c} \begin{matrix} \text{H} & \text{C} & \text{S} & \text{A} & \text{D} & \text{Dc} & \text{Ds} & \text{Da} \end{matrix} \\ \begin{bmatrix} {}_tp_x^{HH} & {}_tp_x^{HC} & {}_tp_x^{HS} & {}_tp_x^{HA} & {}_tp_x^{HD} & 0 & 0 & 0 \\ 0 & {}_tp_x^{CC} & 0 & 0 & {}_tp_x^{CD} & {}_tp_x^{CDc} & 0 & 0 \\ 0 & 0 & {}_tp_x^{SS} & 0 & {}_tp_x^{SD} & 0 & {}_tp_x^{SDs} & 0 \\ 0 & 0 & 0 & {}_tp_x^{AA} & {}_tp_x^{AD} & 0 & 0 & {}_tp_x^{ADa} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} $$

The probability that an insured aged x in state i stays in state i throughout the period from age x to age $x+t$ can be expressed as

$$ {}_tp_x^{ii} = \exp\left\{ -\int_0^t \sum_{j\neq i, j\in\mathcal{E}} \mu_{x+s}^{ij} ds \right\}, \quad i \in \mathcal{E}, $$

where μ_{x+s}^{ij} is the transition intensity from state i to state j. Specifically, we have for $x \geqslant 0$, $t \geqslant 0$ that

$$ {}_tp_x^{HH} = e^{-\int_0^t \left[\mu^{HD}(x+u)+\mu^{HC}(x+u)+\mu^{HS}(x+u)+\mu^{HA}(x+u)\right]du}, \tag{2.1} $$

$$ {}_tp_x^{ii} = e^{-\int_0^t \left[\mu^{iDi}(x+u)+\mu^{iD}(x+u)\right]du}, \quad i = \text{C, S, A}. \tag{2.2} $$

Similarly, if the transition intensities are known, we can express the other transition probabilities in the transition matrix, for $x \geqslant 0$, $t \geqslant 0$ as

$$ {}_tp_x^{HD} = \int_0^t {}_up_x^{HH}\mu^{HD}(x+u)du, \tag{2.3} $$

and for i=C, S, A,

$$ {}_tp_x^{Hi} = \int_0^t {}_up_x^{HH}\mu^{Hi}(x+u)\,{}_{t-u}p_{x+u}^{ii}du, \tag{2.4} $$

$$ {}_tp_x^{iDi} = \int_0^t {}_up_x^{ii}\mu^{iDi}(x+u)du. \tag{2.5} $$

2.3.2 Transition Intensities

In this section, we focus on obtaining the transition intensities which are needed to calculate the transition probabilities given in (2.1)-(2.5). Transition intensities can be obtained from mortality rates, incidence rates and prevalence rates, which we discuss below.

Mortality Models

For transition intensities from the healthy or illness states to the death states, also known as the forces of mortality, some analytical laws of mortality similar to that in Baione and Levantesi (2018) are used.

In this project, we consider two well-known mortality laws, namely, two-parameter Gompertz model and two-parameter Weibull model. Under the two-parameter Gompertz model, the force of mortality $\mu(x)$ for an insured aged x is given by

$$ \mu(x) = e^{(\beta_1+\beta_2 x)}, \quad x \geqslant 0,\ \beta_2 > 0, $$

and under the two-parameter Weibull model, the force of mortality $\mu(x)$ for an insured aged x is expressed as

$$\mu(x) = \beta_1 x^{\beta_2}, \quad x \geqslant 0,\ \beta_1 > 0,\ \beta_2 > 0.$$

Both models are in continuous time and of simple expressions. They are used in this project for transition intensities from illness S to Ds (death due to S), illness A to Da (death due to A) and H (healthy) to D (death due to other causes). After estimating the parameters based on appropriate and available mortality data, these continuous-time mortality models are used for pricing our CI products.

Now, under Gompertz model, we express these transition intensities, for $x \geqslant 0$, as

$$\begin{cases} \mu^{SDs}(x) = e^{(\beta_1^s + \beta_2^s x)} \\ \mu^{ADa}(x) = e^{(\beta_1^a + \beta_2^a x)} \\ \mu^{CDc}(x) = e^{(\beta_1^c + \beta_2^c x)} \\ \mu^{HD}(x) = e^{(\beta_1^d + \beta_2^d x)} \end{cases}, \tag{2.6}$$

where β_2^s, β_2^a, β_2^c, β_2^d are positive. Under the Weibull model, these intensities are assumed to be

$$\begin{cases} \mu^{SDs}(x) = \beta_1^s x^{\beta_2^s} \\ \mu^{ADa}(x) = \beta_1^a x^{\beta_2^a} \\ \mu^{CDc}(x) = \beta_1^c x^{\beta_2^c} \\ \mu^{HD}(x) = \beta_1^d x^{\beta_2^d} \end{cases}, \tag{2.7}$$

where all the parameters are positive.

According to Centers for Disease Control and Prevention(CDC)[1], the mortality rate is defined as the frequency of occurrence of death in a defined population during a given period of time, which is also known as central death rate. Mortality rates data can normally be obtained from different sources. This project uses Canadian mortality data to estimate the parameters in both models for the four types of transition intensities.

Transition Intensity Data from S to Ds, A to Da and C to Dc

In order to estimate parameters of mortality models, appropriate data are needed. To describe data, we introduce the following notations:

- M_x^{iD}: number of death from state i except due to cancer, stroke or heart attack divided by the population of state i, for, $i = \mathrm{H}, \mathrm{C}, \mathrm{S}, \mathrm{A}$;

[1] https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section3.html

- m_x: total mortality rate;
- m_x^i: mortality rate by state i, for $i = \mathrm{H}, \mathrm{C}, \mathrm{S}, \mathrm{A}$;
- D_x: observed total number of death;
- D_x^{iDi}: observed number of death in illness i due to the same illness, for $i = \mathrm{C}, \mathrm{S}, \mathrm{A}$;
- D_x^{iD}: number of death from state i due to causes except cancer, stroke and heart attack, for $i = \mathrm{H}, \mathrm{C}, \mathrm{S}, \mathrm{A}$;
- L_x: total population;
- L_x^i: population in state i, for $i = \mathrm{H}, \mathrm{C}, \mathrm{S}, \mathrm{A}$;
- f_x^i: prevalence rate of i, for $i = \mathrm{C}, \mathrm{S}, \mathrm{A}$.

The transition intensity from illness i to death due to the same illness, $\mu^{iDi}(x)$, for the insured aged x, is defined as

$$\begin{aligned}\mu^{iDi}(x) &= \frac{\text{Number of death due to the illness } i \text{ at age } x}{\text{Number of people having the illness } i \text{ at age } x}, \\ &= \frac{D_x^{iDi}}{L_x^i}, \quad x \geqslant 0,\ i = \mathrm{C}, \mathrm{S}, \mathrm{A}. \end{aligned} \tag{2.8}$$

Prevalence rate of a particular illness is defined as the number of people who have this particular illness in a population during a given period of time by CDC[2]; here this period of time is from x to $x+1$. This implies that dividing both numerator and denominator in equation (2.8) by the total number of people who are alive at age x (L_x) and using the concept of prevalence rate, we get

$$\mu^{iDi}(x) = \frac{D_x^{iDi}/L_x}{L_x^i/L_x} = \frac{m_x^i}{f_x^i}, \quad x \geqslant 0, \tag{2.9}$$

where m_x^i and f_x^i denote the mortality rate and prevalence rate of illness i, $i = \mathrm{C}, \mathrm{S}, \mathrm{A}$. Hence, from mortality rates and prevalence rates, we get the transition intensities using equation (2.9) which can then be used to estimate the parameters of the mortality models.

Transition Intensity Data from H, C, S, A to D

Generally, it is hard to find data to estimate the transition intensities directly from a particular illness state to state D (death due to other causes). It seems reasonable to introduce an extra positive mortality strength to intensity $\mu^{HD}(x)$ to estimate $\mu^{iD}(x)$ for illness i. The extra mortality is defined by Haberman (1982) as the ratio of difference between the actual number of deaths who are exposed to risk and the expected number of deaths to the

[2]https://www.cdc.gov/csels/dsepd/ss1978/lesson3/section2.html

expected number of deaths during a given period of time. For example, Dash and Grimshaw (1993) assume that the extra mortality strength is the mortality rate of the illness from other causes exceeding that of healthy people. Following this idea, the extra mortality strength of cancer, stroke and heart attack, denoted by γ^c, γ^s and γ^a, respectively, are assumed to be

$$\begin{aligned}\mu^{SD}(x) &= (1+\gamma^s)\cdot\mu^{HD}(x),\\ \mu^{AD}(x) &= (1+\gamma^a)\cdot\mu^{HD}(x),\\ \mu^{CD}(x) &= (1+\gamma^c)\cdot\mu^{HD}(x),\end{aligned}$$

where $x \geqslant 0$, $\gamma^s, \gamma^a, \gamma^c \geqslant 0$.

Then, for $x \geqslant 0$, using the notations introduced before and by our assumptions, for $i = \mathrm{H}, \mathrm{S}, \mathrm{A}, \mathrm{C}$, we have the following relationships:

$$\begin{aligned}
&f_x^i = \frac{L_x^i}{L_x},\\
&m_x = \frac{D_x}{L_x},\\
&m_x^i = \frac{D_x^{iDi}}{L_x},\\
&M_x^{iD} = \frac{D_x^{iD}}{L_x^i}, &(2.10)\\
&M_x^{SD} = (1+\gamma^s)\cdot M_x^{HD},\\
&M_x^{AD} = (1+\gamma^a)\cdot M_x^{HD},\\
&M_x^{CD} = (1+\gamma^c)\cdot M_x^{HD},\\
&\frac{L_x^H}{L_x} = 1 - f_x^C - f_x^A - f_x^S, &(2.11)\\
&D_x = D_x^{CDc} + D_x^{SDs} + D_x^{ADa} + D_x^{HD} + D_x^{CD} + D_x^{SD} + D_x^{AD}. &(2.12)
\end{aligned}$$

Now we derive μ_x^{HD} from the equations above. First, using equation (2.10), equation (2.12) can be expressed as

$$\begin{aligned}D_x = {} & m_x^C L_x + m_x^S L_x + m_x^A L_x + M_x^{HD} L_x^H + (1+\gamma^s) M_x^{HD} L_x^S + (1+\gamma^c) M_x^{HD} L_x^C\\ & + (1+\gamma^a) M_x^{HD} L_x^A.\end{aligned}$$

Dividing both sides by the total population (L_x), we obtain

$$\begin{aligned}m_x = {} & m_x^C + m_x^S + m_x^A + M_x^{HD}\frac{L_x^H}{L_x} + M_x^{HD}(1+\gamma^s) f_x^S + M_x^{HD}(1+\gamma^a) f_x^A\\ & + M_x^{HD}(1+\gamma^c) f_x^C.\end{aligned}$$

With the relationship in equation (2.11), the transition intensity from H to D can be derived as

$$\begin{aligned}\mu^{HD}(x) &= M_x^{HD} \\ &= \frac{m_x - m_x^C - m_x^S - m_x^A}{1 + \gamma^s f_x^S + \gamma^a f_x^A + \gamma^c f_x^C}, \\ &= \frac{m_x - \sum_{i=\text{C, S, A}} m_x^i}{1 + \sum_{i=\text{C, S, A}} \gamma^i f_x^i}, \end{aligned} \tag{2.13}$$

where $x \geqslant 0$, $\gamma^s \geqslant 0$, $\gamma^a \geqslant 0$, $\gamma^c \geqslant 0$. Now, if the data of m_x, m_x^i and f_x^i for $i =$ C, S, A can be obtained, we have the transition intensity data from healthy (H) to death by other causes (D). The latter can be used to estimate the parameters of the mortality models presented above.

Transition Intensities from H to C, S, A

In this project, piece-wise constant transition intensities are assumed for the transitions from healthy (H) to a particular illness following the idea of Jones (1994). Specifically, we assume that for an insured aged x_0 at the time of policy issue, the transition intensities from H to a particular illness i ($i =$C, S, A) are of the form

$$\mu^{Hi}(x) = \begin{cases} 0, & x < x_0 \\ \sigma_k^i, & x_k \leqslant x < x_{k+1} \\ \sigma_n^i, & x \geqslant x_n \end{cases},$$

for $k = 0, 1, ..., n-1$, $x \geqslant 0$ and $i = C, S, A$.

Here we discuss two ways of obtaining these piece-wise intensities. (1) If reliable incidence rates are available, the incidence rates can be directly used as the transition intensities. Given by CDC[2], the definition of incidence rate for a particular disease is the ratio of the number of new cases of this disease to the total population during a given period of time. (2) If the incidence rates are unavailable or unreliable, the prevalence rates can be used to estimate the transition intensities.

Following Haberman (1984), we assume that all the observations are healthy (in state H) at age x, and then the prevalence rate of illness i between x and $x+t$, denoted by f_{x+t}^i, is defined as

$$f_{x+t}^i = \frac{\text{Number of people who have illness } i \text{ at age } x+t \text{ from state H at age } x}{\text{Total cohort population size from age } x \text{ to } x+t},$$

where $i =$ C, S, A, and $x \geqslant 0, t \geqslant 0$.

Since the total cohort population size from age x to $x+t$ is the sum of the number of people having illness i ($i =$ C, S, A) from H and the number of people staying in H,

by dividing the population size at their age x, the prevalence rate defined above can be expressed as the transition probability from H to illness i divided by the sum of the transition probabilities from H to i, $i =$ C, S, A, and staying in H, that is,

$$f^i_{x+t} = \frac{{}_tp^{Hi}_x}{{}_tp^{HH}_x + {}_tp^{HC}_x + {}_tp^{HS}_x + {}_tp^{HA}_x}, \quad x \geqslant 0, t \geqslant 0,\ i = \text{S, A.} \tag{2.14}$$

In the project, the incidence rates of cancer are used as transition intensities from H to C, that is, σ^A_k $(k = 0, 1, \ldots, n)$ are known from the data directly. The transition intensities from H to stroke and heart attack (S and A) are obtained from their prevalence rates data. In the following, we show the steps first to get the transition intensities σ^S_0 and σ^A_0 in age (x_0, x_1) and then to get piece-wise transition intensities σ^S_k and σ^A_k for $k = 1, 2, \ldots, n-1$ iteratively.

Step 1: Express transition probabilities in terms of unknown σ^S_0 and σ^A_0

As stated in Section 2.3.2, two mortality models, Gompertz and Weibull are considered. Using our piece-wise intensity notations and transition probability formulae (2.1) and (2.4), we have the following two sets of formulae under Gompertz model and Weibull model, respectively. For simplicity, we write $t_1 = x_1 - x_0$ so that $x_0 + t_1 = x_1$ in this step.

1) If the Gompertz model is used as the mortality models, we get

$$\begin{aligned} {}_{t_1}p^{HH}_{x_0} &= e^{-\int_0^{t_1}\left[\mu^{HD}(x_0+u)+\mu^{HC}(x_0+u)+\mu^{HS}(x_0+u)+\mu^{HA}(x_0+u)\right]du} \\ &= e^{-\int_0^{t_1}\left[e^{\beta^d_1+\beta^d_2(x_0+u)}+\sigma^C_0+\sigma^S_0+\sigma^A_0\right]du} \\ &= \exp\left\{-\frac{e^{\beta^d_1}}{\beta^d_2}\left[e^{\beta^d_2 x_1} - e^{\beta^d_2 x_0}\right] - \left(\sigma^C_0 + \sigma^S_0 + \sigma^A_0\right)t_1\right\}. \end{aligned} \tag{2.15}$$

For $i =$ S, A, C, we get

$$\begin{aligned} {}_{t_1}p^{ii}_{x_0} &= e^{-\int_0^{t_1}\left[\mu^{iD}(x+u)+\mu^{iDi}(x+u)\right]du} \\ &= e^{-\int_0^{t_1}\left[(1+\gamma^i)e^{\beta^d_1+\beta^d_2(x_0+u)}+e^{\beta^i_1+\beta^i_2(x_0+u)}\right]du} \\ &= \exp\left\{-\frac{e^{\beta^d_1}}{\beta^d_2}(1+\gamma^i)\left[e^{\beta^d_2 x_1} - e^{\beta^d_2 x_0}\right] - \frac{e^{\beta^i_1}}{\beta^i_2}\left[e^{\beta^i_2 x_1} - e^{\beta^i_2 x_0}\right]\right\}. \end{aligned} \tag{2.16}$$

Furthermore, using formulae (2.15) and (2.16), for $i =$ S, A, C, we get

$$\begin{aligned} {}_{t_1}p^{Hi}_{x_0} &= \int_0^{t_1} {}_up^{HH}_{x_0}\mu^{Hi}(x_0+u)\,{}_{t_1-u}p^{ii}_{x_0+u}du \\ &= \int_0^{t_1} {}_up^{HH}_{x_0}\sigma^i_0\,{}_{t_1-u}p^{ii}_{x_0+u}du \end{aligned}$$

$$= \exp\left\{-\frac{e^{\beta_1^i}}{\beta_2^i}e^{\beta_2^i x_1} - \frac{e^{\beta_1^d}}{\beta_2^d}\left[e^{\beta_2^d x_1}(1+\gamma^i) - e^{\beta_2^d x_0}\right]\right\}\sigma_0^i$$
$$\times \int_0^{t_1} \exp\left\{-\left(\sigma_0^C + \sigma_0^S + \sigma_0^A\right)u + \frac{e^{\beta_1^d}}{\beta_2^d}\gamma^i e^{\beta_2^d(x_0+u)} + \frac{e^{\beta_1^i}}{\beta_2^i}e^{\beta_2^i(x_0+u)}\right\}du. \tag{2.17}$$

2) If the Weibull function is used as the mortality models, we get that

$$\begin{aligned} {}_{t_1}p_{x_0}^{HH} &= e^{-\int_0^{t_1}\mu^{HD}(x_0+u)+\mu^{HC}(x_0+u)+\mu^{HS}(x_0+u)+\mu^{HA}(x_0+u)du} \\ &= e^{-\int_0^{t_1}\left[\beta_1^d(x_0+u)^{\beta_2^d}+\sigma_0^C+\sigma_0^S+\sigma_0^A\right]du} \\ &= \exp\left\{-\frac{\beta_1^d}{\beta_2^d+1}\left({x_1}^{\beta_2^d+1} - {x_0}^{\beta_2^d+1}\right) - \left(\sigma_0^C+\sigma_0^S+\sigma_0^A\right)t_1\right\}, \end{aligned} \tag{2.18}$$

and for i = C, S, A,

$$\begin{aligned} {}_{t_1}p_{x_0}^{ii} &= e^{-\int_0^{t_1}\mu^{iD}(x_0+u)+\mu^{iDi}(x_0+u)du} \\ &= e^{-\int_0^{t_1}\left[(1+\gamma^i)\beta_1^d(x_0+u)^{\beta_2^d}+\beta_1^i(x_0+u)^{\beta_2^i}\right]du} \\ &= \exp\left\{-\frac{\beta_1^i}{\beta_2^i+1}\left({x_1}^{\beta_2^i+1} - {x_0}^{\beta_2^i+1}\right) - (1+\gamma^i)\frac{\beta_1^d}{\beta_2^d+1}\left({x_1}^{\beta_2^d+1} - {x_0}^{\beta_2^d+1}\right)\right\}, \end{aligned} \tag{2.19}$$

$$\begin{aligned} {}_{t_1}p_{x_0}^{Hi} &= \int_0^{t_1} {}_up_{x_0}^{HH}\mu^{Hi}(x_0+u)\,{}_{t_1-u}p_{x_0+u}^{ii}du \\ &= \int_0^{t_1} {}_up_{x_0}^{HH}\sigma_0^i\,{}_{t_1-u}p_{x_0+u}^{ii}du \\ &= \exp\left\{-\frac{\beta_1^i}{\beta_2^i+1}{x_1}^{\beta_2^i+1} - \frac{\beta_1^d}{\beta_2^d+1}(1+\gamma^i)\left({x_1}^{\beta_2^d+1} - {x_0}^{\beta_2^d+1}\right)\right\}\sigma_0^i \\ &\quad\times\int_0^{t_1}\exp\left\{-\left(\sigma_0^C+\sigma_0^S+\sigma_0^A\right)u + \gamma^i\frac{\beta_1^d}{\beta_2^d+1}(x_0+u)^{\beta_2^d+1} + \frac{\beta_1^i}{\beta_2^i+1}\right. \\ &\quad\left.\times(x_0+u)^{\beta_2^i+1}\right\}du. \end{aligned} \tag{2.20}$$

Step 2: Solve for σ_0^S and σ_0^A from a system of equations

Now putting the expressions of transition probabilities obtained in the first step into equation (2.14), we have the following two equations:

$$f_{x_1}^S = \frac{{}_{t_1}p_{x_0}^{HS}}{{}_{t_1}p_{x_0}^{HH} + {}_{t_1}p_{x_0}^{HC} + {}_{t_1}p_{x_0}^{HS} + {}_{t_1}p_{x_0}^{HA}},$$
$$f_{x_1}^A = \frac{{}_{t_1}p_{x_0}^{HA}}{{}_{t_1}p_{x_0}^{HH} + {}_{t_1}p_{x_0}^{HC} + {}_{t_1}p_{x_0}^{HS} + {}_{t_1}p_{x_0}^{HA}},$$

where $t_1 = x_1 - x_0$. Knowing $f_{x_1}^S$ and $f_{x_1}^A$, this system of two equations can be used to solve for σ_0^S and σ_0^A.

Step 3: Obtain σ_1^S and σ_1^A by an iterative approach

Suppose that the insured holds a contract from age x_0 to age x_2. The prevalence rates for illnesses S and A between ages x_1 and x_2 are $f_{x_2}^S$ and $f_{x_2}^A$, respectively, which are assumed known and can be used to find σ_1^S and σ_1^A.

Similar to equations in Step 2, setting $t_1 = x_1 - x_0$ and $t_2 = x_2 - x_1$ we have the following two equations:

$$f_{x_2}^S = \frac{{}_{t_1+t_2}p_{x_0}^{HS}}{{}_{t_1+t_2}p_{x_0}^{HH} + {}_{t_1+t_2}p_{x_0}^{HC} + {}_{t_1+t_2}p_{x_0}^{HS} + {}_{t_1+t_2}p_{x_0}^{HA}}, \tag{2.21}$$

$$f_{x_2}^A = \frac{{}_{t_1+t_2}p_{x_0}^{HA}}{{}_{t_1+t_2}p_{x_0}^{HH} + {}_{t_1+t_2}p_{x_0}^{HC} + {}_{t_1+t_2}p_{x_0}^{HS} + {}_{t_1+t_2}p_{x_0}^{HA}}, \tag{2.22}$$

where

$$\begin{aligned}
{}_{t_1+t_2}p_{x_0}^{HH} &= \left({}_{t_1}p_{x_0}^{HH}\right)\left({}_{t_2}p_{x_0+t_1}^{HH}\right), \\
{}_{t_1+t_2}p_{x_0}^{Hi} &= \left({}_{t_1}p_{x_0}^{Hi}\right)\left({}_{t_2}p_{x_0+t_1}^{ii}\right) + \left({}_{t_1}p_{x_0}^{HH}\right)\left({}_{t_2}p_{x_0+t_1}^{Hi}\right), \quad i = \text{C, S, A.}
\end{aligned}$$

Since ${}_{t_1}p_{x_0}^{HH}$ and ${}_{t_1}p_{x_0}^{Hi}$ for $i =$ C, S, A are known by Steps 1-2, equations (2.21) and (2.22) are now functions of σ_1^S and σ_1^A so they can be solved.

Repeating Step 3, σ_k^S and σ_k^A for $k = 2, 3, \ldots, n-1$ can be obtained iteratively.

2.4 Pricing Models

In this section, we present formulae for net single premiums of the CI products based on the Markovian multi-state models introduced. Suppose that an n-year term CI policy is issued to a healthy male or female aged x. The force interest used for discounting is assumed to be constant δ, implying that the discount function that discounts \$1 from time t to present $(t = 0)$ is given by

$$v_t = e^{-\delta t}, \quad t \geqslant 0.$$

Benefit amounts are assumed to be B_C, B_S, and B_A, payable immediately when the insured is diagnosed with cancer, stroke or heart attack, respectively, or B_D when the insured is dead within n-year term. When these benefit amounts are of one unit (i.e., $B_C = B_S = B_A = B_D = \$1$), as derived in Dickson et al. (2013), the net single premium (NSP) rates, denoted by $\overline{A}_{x:n}^i$ for $i =$ C, S, A and $\overline{A}_{x:n}^D$, respectively, are given by

$$\begin{aligned}
\overline{A}_{x:n}^i &= \int_0^n {}_tp_x^{HH} \cdot \mu^{Hi}(x+t) \cdot e^{-\delta t} dt, \quad i = \text{C, S, A,} \\
\overline{A}_{x:n}^D &= \int_0^n {}_tp_x^{HH} \cdot \mu^{HD}(x+t) \cdot e^{-\delta t} dt.
\end{aligned}$$

Then, the NSP rates for the standard n-year term CI Stand Alone and Full Accelerated policy issued to an insured aged x, denoted by P_x^{SA} and P_x^{FA}, respectively, are given by

$$\begin{aligned} P_x^{SA} &= B_C \times \overline{A}_{x:n}^{C} + B_S \times \overline{A}_{x:n}^{S} + B_A \times \overline{A}_{x:n}^{A} \\ &= \int_0^n {}_tp_x^{HH} \left[B_C \cdot \mu^{HC}(x+t) + B_S \cdot \mu^{HS}(x+t) + B_A \cdot \mu^{HA}(x+t) \right] e^{-\delta t} dt, \end{aligned} \tag{2.23}$$

and

$$\begin{aligned} P_x^{FA} &= B_C \times \overline{A}_{x:n}^{C} + B_S \times \overline{A}_{x:n}^{S} + B_A \times \overline{A}_{x:n}^{A} + B_D \times \overline{A}_{x:n}^{D} \\ &= \int_0^n {}_tp_x^{HH} \left[B_C \cdot \mu^{HC}(x+t) + B_S \cdot \mu^{HS}(x+t) + B_A \cdot \mu^{HA}(x+t) \right. \\ &\qquad \left. + B_D \cdot \mu^{HD}(x+t) \right] e^{-\delta t} dt. \end{aligned} \tag{2.24}$$

If we consider additional transportation accidental death benefit of one unit being paid at the time of accidental death within n-year term for an insured aged x, the NSP for this coverage alone, denoted by $\overline{A}_{x:n}^{Acc}$, is given by

$$\overline{A}_{x:n}^{Acc} = \int_0^n {}_tp_x^{HH} \cdot \mu^{Acc}(x+t) \cdot e^{-\delta t} dt,$$

where $\mu^{Acc}(x+t)$ is the transition intensity for an issued aged $x+t$ transitioning from healthy (H) to death due transportation accident (D_2).

If a rider of the transportation accidental death is considered on the Full Accelerated policy described above assuming the benefit of the transportation accidental death being two times of the sum assured for death with other causes, then the NSP rate for this policy, denoted by P_x^{Acc}, can be express as

$$\begin{aligned} P_x^{Acc} =& B_C \times \overline{A}_{x:n}^{C} + B_S \times \overline{A}_{x:n}^{S} + B_A \times \overline{A}_{x:n}^{A} + B_D \times \overline{A}_{x:n}^{D} + B_D \times \overline{A}_{x:n}^{Acc} \\ &= \int_0^n {}_tp_x^{HH} \left[B_C \cdot \mu^{HC}(x+t) + B_S \cdot \mu^{HS}(x+t) + B_A \cdot \mu^{HA}(x+t) \right. \\ &\qquad \left. + B_D \cdot \mu^{HD}(x+t) + B_D \cdot \mu^{Acc}(x+t) \right] e^{-\delta t} dt. \end{aligned} \tag{2.25}$$

In Section 4.4 when comparing prices of several CI products under different approaches based on mortality models and Markovian multi-state models discussed and Canadian health data presented in earlier chapters, we set all the benefit amounts to be one dollar unit (the benefit for transportation accidental death is then two dollars) and the force of interest is assumed constant at 5% (i.e., $\delta = 0.05$).

Chapter 3

The Data

In this chapter, we present the Canadian health data used in the empirical application in detail. The prevalence rates, incidence rates of cancer, stroke and heart attack, and the mortality rates due to cancer, stroke, heart attack and other causes are discussed.

In the project, the incidence rate of cancer is used as the transition intensity from Healthy to Cancer and prevalence rates of stroke and heart attack are used to estimate the transition intensities from Healthy to Stroke and from Healthy to Heart Attack, respectively. (See Section 3.2 for details). The incidence rate of cancer can be found from Statistics Canada. According that cancer patients do not die due to cancer in a short time after diagnosis and considering there are thousands of new cases every year, it is appropriate to use the incidence rate of cancer as the transition intensity from Healthy to Cancer.

The data for stroke and heart attack from the Canadian Chronic Disease Surveillance System (CCDSS) is reported by hospitals from patients who have stayed in hospitals for a certain length of time. From their report of incidence rates, the number of new cases of stroke and heart attack in a certain year is small and may not be reliable. For example, the incidence rate of stroke for patients aged from 20 to 34 in 2015 is 15% more than the rate in 2014. The incidence rates vary from year to year since there are only single-digit or double-digit new cases reported by hospitals. By contrast, prevalence rates of stroke and heart attack are based on the cumulative cases that there are hundreds of cases reported every year. Hence, we use prevalence rates of stroke and heart attack to find the transition intensities from Healthy to Stroke and from Healthy to Heart Attack. The incidence rates of stroke and heart attack are used as the reference information when comparing mortality models in this project; we compare incidence rates with the estimated transition intensities under different mortality models.

However, there are some limitations. Firstly, the prevalence rates and incidence rates are not collected from the same source. Secondly, the Canadian data for stroke and heart attack excludes the province of Quebec, while the incidence rate of cancer includes all provinces and territories. Lastly, the age span of the prevalence rate is 15 years starting from 20 years old (a smaller age span is actually preferred). These facts may have some impacts

on the result we obtained. We believe our results may be improved if more accurate and appropriate data are available.

3.1 Prevalence Rate

From the Public Health Infobase provided by CCDSS[1], we can find the prevalence rates of Stroke and Heart Attack by genders and age groups of 15 years per age band. Table 3.1 shows prevalence rates for both males and females in 2015 in Canada (except Quebec); they are the latest available data and are used for this project.

Prevalence rates (percent)

Age Group	**Stroke**		**Heart Attack**	
(Years)	Males	Females	Males	Females
20-34	0.12	0.14	0.03	0.01
35-49	0.57	0.64	0.55	0.14
50-64	2.39	2.01	3.5	0.97
65-79	7.71	5.94	8.46	3.11
80+	18.38	16.36	13.41	7.96

Table 3.1: Prevalence rates of stroke and heart attack in 2015

From Table 3.1, we observe that the prevalence rates are increasing with age for each type of illness and each gender. The prevalence rates of stroke for females are higher than the corresponding rates for males from age 20 to 49, and then become lower than those for males after 50 years old. As for the prevalence rates of heart attack, female rates keep lower than male rates for all ages.

Prevalence rates of cancer is found from Canadian Cancer Registry (CCR)[2]. The prevalence proportions for all cancers combined in 2005 are the latest data we can find and are given in Table 3.2. From Table 3.2, it can be observed that female prevalence rates of cancer

Prevalence rates of cancer

Age Group	**Males**	**Females**
20-39	183.8	293.8
40-49	486.0	1,067.8
50-59	1,656.3	2,229.0
60-69	4,898.2	3,643.7
70-79	8,287.8	4,742.7
80+	8,945.3	4,934.9

Table 3.2: Prevalence rates of cancer (per 100,000)

[1]The data are found at https://health-infobase.canada.ca/ccdss/data-tool/

[2]Statistics Canada, 2009, "Cancer prevalence in the Canadian population", Health Reports, Vol. 20, No. 1, catalogue number 82-003-X.

are higher than those for males before age 60, and then become lower than male ones.

3.2 Incidence Rate

We can find the incidence rate of cancer directly from Statistics Canada[3], where the data are provided by CCR. Since it is hard for a patient to recover from cancer, which implies that all the new cases should be transitioned from healthy status, it is reasonable to use the rate of new cases of primary cancer as the incidence rate of cancer. The incidence rates of cancer in Canada 2015 for both males and females are shown in Table 3.3, which are collected in the same year as the prevalence rates of stroke and heart attack displayed in Table 3.1.

Observed from Table 3.3, the incidence rates of cancer are increasing from age 20 to 89 and then decreasing a little after age 90 for both males and females. Compared with female incidence rates of cancer, male incidence rates are slightly higher in the first 5 years starting from 20 years old and then become lower until age 59. After 60 years old, male incidence rates are much higher than female ones.

Rates of new cases of primary cancer

Age group	Males	Females
20-24	38.40	37.00
25-29	48.50	61.70
30-34	68.20	115.20
35-39	94.00	184.30
40-44	131.90	281.00
45-49	235.80	407.50
50-54	424.80	582.30
55-59	750.60	753.60
60-64	1,175.50	993.30
65-69	1,746.80	1,292.30
70-74	2,251.50	1,616.00
75-79	2,623.30	1,734.10
80-84	2,995.90	1,837.60
85-89	3,015.60	1,830.40
90+	2,690.00	1,533.80

Table 3.3: Rates of new cases of primary cancer in each age group (per 100,000)

Since the prevalence rates given by CCDSS are with a 15-year age span, incidence rates of cancer are then adjusted to age-standardized rates in the same age groups, following the formulae for age-standardized rates by Statistics Canada[4]. For example, the incidence rate

[3]The data are found from Statistics Canada, Table 13-10-0111-01, number and rates of new cases of primary cancer, by cancer type, age group and sex.

[4]https://www.statcan.gc.ca/eng/dai/btd/asr

in age group "20-34" should be

$$\sigma^C_{20-34} = \frac{\text{Population}_{20-24} \cdot \sigma^C_{20-24} + \text{Population}_{25-29} \cdot \sigma^C_{25-29} + \text{Population}_{30-34} \cdot \sigma^C_{30-34}}{\text{Population}_{20-24} + \text{Population}_{25-29} + \text{Population}_{30-34}},$$

where $\sigma^C_{x_1-x_2}$, the incidence rate of cancer in age group "$x_1 - x_2$", and $\text{Population}_{x_1-x_2}$, the population in age group "$x_1 - x_2$", are given in Table 3.4 by Statistics Canada[5]. These adjusted incidence rates of cancer are listed in Table 3.5.

Population		
Age Group	Males	Females
20-24	1,244,697	1,150,926
25-29	1,239,356	1,190,201
30-34	1,230,618	1,229,883
35-39	1,174,086	1,197,143
40-44	1,167,211	1,182,711
45-49	1,220,275	1,225,541
50-54	1,392,935	1,390,415
55-59	1,300,456	1,314,212
60-64	1,102,960	1,140,251
65-69	926,287	976,717
70-74	649,566	708,146
75-79	452,282	530,742
80-84	317,644	417,363
85-89	177,089	290,076
90+	77,062	191,352

Table 3.4: Population estimates in each age group

Incidence rates of cancer		
Age Group	Males	Females
20-34	0.000516421	0.000721651
35-49	0.001550046	0.002918912
50-64	0.007545062	0.007627395
65-79	0.021039070	0.015015921
80+	0.029666204	0.017820326

Table 3.5: Incidence rates of cancer in each age group

The incidence rates of stroke and heart attack for males and females obtained from CCDSS[1] are shown in Table 3.6, and can be used for comparing the mortality models in Chapter 4. From Table 3.6, we observe that male incidence rates of stroke are slightly lower than the ones for females in age group "20-34", and higher than female ones after 35 years

[5]The data are found from Statistics Canada, Table 17-10-0005-01, population estimates on July 1st, by age and sex.

old, while male incidence rates of heart attack are always higher than female ones at all ages.

Incidence rates				
Age	**Stroke**		**Heart Attack**	
Group	Males	Females	Males	Females
20-34	0.00022	0.00025	0.00006	0.00002
35-49	0.00081	0.00074	0.00088	0.00024
50-64	0.00295	0.00218	0.00341	0.00119
65-79	0.00855	0.00647	0.00691	0.00352
80+	0.02074	0.01913	0.01400	0.01019

Table 3.6: Incidence rates of stroke and heart attack in each age group

3.3 Mortality Rate

The mortality rates for 2015 are obtained from Statistics Canada[6]. The mortality rates by all causes and specified by causes for males and females are given in Table 3.7. From Table 3.7, it is clear that male mortality rates, the total number of death over $100,000$ population, are higher than female ones in all age groups. For mortality rates by causes, we observe that male mortality rates by cancer and by heart attack are higher than their corresponding female rates. As for the mortality rates by stroke, male rates are lower than female ones from age 20 to 29, and become larger than those for females from 30 years old to 89 years old, and then, after age 90, male rates are lower than female ones.

Using mortality rates by specified causes, if the extra mortality rates of cancer, stroke and heart attack (γ^s, γ^c, γ^a) are assumed to be zero, following equation (2.13), the intensity from healthy to death by causes other than these three illnesses for an insured aged x can be expressed as

$$\mu^{HD}(x) = m_x - m_x^C - m_x^S - m_x^A,$$

where m_x is the mortality rate by all causes at age x, and m_x^i is the mortality rate by the specific disease i, where $i = \mathrm{C}, \mathrm{S}, \mathrm{A}$.

The mortality rates of transportation accidents in 2015 by Statistics Canada[6] are also provided in Table 3.8 for males and females. It is clear that all the male rates are bigger than female rates. There are no increasing or decreasing trends for male and female mortality rates of transportation accidents, which indicates that it is not appropriate to fit the mortality models discussed in Section 2.3.2 with the mortality rates data due to transportation

[6]The data are found from Statistics Canada, Table 13-10-0392-01, deaths and age-specific mortality rates, by selected grouped causes.

Age-specific mortality rate per 100,000 population

Age at time	**Total (all causes of death)**		**Cancer**	
of death	Males	Females	Males	Females
20-24	71.2	31.3	4.8	2.6
25-29	85.7	36	4.7	6.2
30-34	90.8	46	10.1	11.8
35-39	103.5	57.9	14.7	19.3
40-44	147.9	93	24.8	35.8
45-49	223.8	137.3	50.7	62.9
50-54	352.3	241.7	106.1	120.6
55-59	574	377.4	209.8	209
60-64	907.8	593.3	379.4	317.5
65-69	1,371.6	889.6	579.4	446.6
70-74	2,181.9	1,452.5	902.8	664.7
75-79	3,571.1	2,377.9	1,310.1	878.2
80-84	6,308.3	4,390	1,884.7	1,198.6
85-89	11,109.3	7,891.9	2,509.4	1,429.7
90+	21,397.8	17,778.6	3,469.9	1,817.8
Age at time	**Heart Attack**		**Stroke**	
of death	Males	Females	Males	Females
20-24	0	0	0.1	0.4
25-29	0.2	0.1	0.4	0.6
30-34	0.6	0.1	1.1	0.7
35-39	2	0.7	1.5	1.2
40-44	4.6	1.2	3.6	2.7
45-49	12.1	2.6	4.8	4
50-54	22.1	4.7	8.5	7.7
55-59	41.5	10	14.6	12.1
60-64	64.3	22.9	22.6	16.9
65-69	95.5	34.6	41	33.6
70-74	142.2	62.7	86.3	64.2
75-79	199.5	115.8	164.1	135.3
80-84	376.1	218	358.6	310.7
85-89	670.8	417.1	665.8	618.3
90+	1,255.6	970.7	1,287.2	1,385.8

Table 3.7: Mortality rates of total death and by causes

accidents. Accordingly, piece-wise mortality rates shown in Figure 3.1 are considered as the transition intensities from Healthy to Transportation Accidental Death.

Transportation Accidental Death		
Age at time of death	**Males**	**Females**
20-24	14.3	4.8
25-29	12.4	2.9
30-34	9.6	2.6
35-39	6.5	2.3
40-44	9.2	2.6
45-49	9.2	3.6
50-54	10.8	2.5
55-59	13.4	2.8
60-64	10.7	3.1
65-69	8.6	4.7
70-74	11.5	5
75-79	15.3	6.4
80-84	20	10.3
85-89	23.3	9.1
90+	24	10

Table 3.8: Mortality rates of transportation accidents (per 100,000)

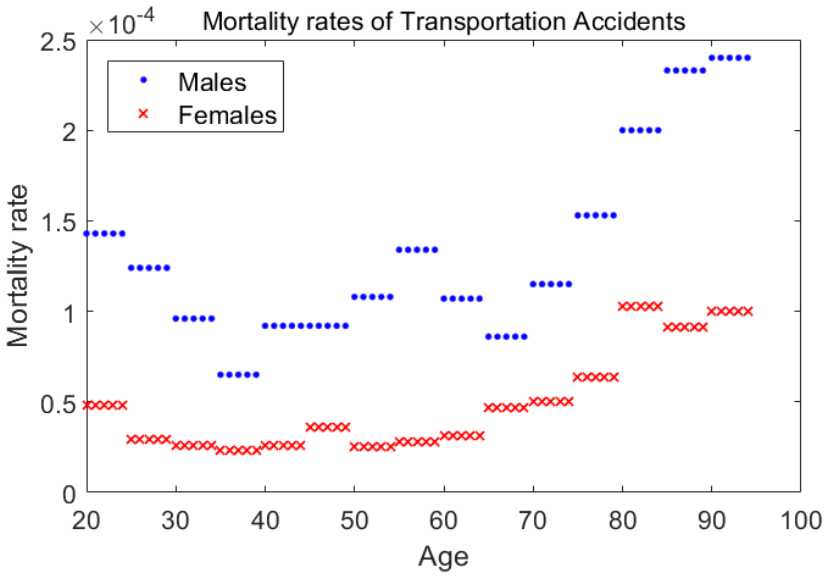

Figure 3.1: Mortality rates of transportation accidents

Chapter 4

Empirical Analysis

In this chapter, the mortality models and Markovian multi-state models discussed in Chapter 2 are applied to Canadian health data. Comparisons of transition intensities and premium rates under different approaches of graduations are presented in detail.

4.1 Mortality Models

Two-parameter Gompertz model and Weibull model introduced in Chapter 2 are given by equation (2.6) and (2.7), respectively. In this section, we fit the Gompertz model and Weibull model to mortality rates. The estimated parameters of the mortality models under the Nonlinear Least Squares method with 95% confidence interval (CI) and the goodness of fit measures are given in Tables 4.1-4.4. The goodness of fit is provided by Matlab where smaller SSE (Sum of Squares due to Error) and RMSE (Root Mean Squared Error) and larger R^2 or adjusted R^2 are preferred. The adjusted R^2 is a modified version of R^2; it takes the number of predictors in the model into account. From Tables 4.1-4.4, the Gompertz model always provides the lower SSE and RMSE, and the larger R^2 and adjusted R^2 than the Weibull model, which implies that the Gompertz model is better fitted with the mortality rates data. All Gompertz models are fitted well with R^2 and Adjusted R^2 values around 0.9, except for the male mortality model for heart attack, where both values are around 0.82.

The fitted mortality rates under the Gompertz and Weibull models for males and females are displayed in Figure 4.1. It shows that for mortality models of healthy, both the Gompertz model and the Weibull model give similar curves for males. Under female mortality models of healthy, the Gompertz model presents a similar result to that for males, while female rates are lower than those for males under the Weibull model.

In addition, for the mortality models due to Stroke and Heart Attack, the fitted models are not very close to the observed mortality rates in the first few years. From equation (2.9), the mortality rates from stroke to death due to stroke, μ^{SDs}, are given by the mortality rate of stroke over the prevalence rate of stroke. Since the mortality rate by a specified

Model	**Gompertz** $f(x) = exp(\beta_1 + \beta_2 x)$	**Weibull** $f(x) = \beta_1 x^{\beta_2}$
	Males	
Coefficients (95% CI)	$\hat{\beta}_1$ =-12.99 (-13.27, -12.71) $\hat{\beta}_2$ =0.119 (0.1159, 0.1221)	$\hat{\beta}_1$ = 7.657e-22 (-3.495e-22, 1.881e-21) $\hat{\beta}_2$ = 10.31 (9.989, 10.64)
SSE	0.0003951	0.000582
R^2	0.9967	0.9952
Adjusted R^2	0.9967	0.9951
RMSE	0.002326	0.002823
	Females	
Coefficients (95% CI)	$\hat{\beta}_1$ =-14.88 (-15.18, -14.58) $\hat{\beta}_2$ =0.139 (0.1357, 0.1423)	$\hat{\beta}_1$ =7.29e-16 (-1.67e-15, 3.128e-15) $\hat{\beta}_2$ =7.086 (6.352, 7.821)
SSE	0.0002861	0.00141
R^2	0.9969	0.9846
Adjusted R^2	0.9968	0.9844
RMSE	0.00198	0.004395

Table 4.1: Estimated parameters of mortality models (from Healthy to Death)

cause has an age span of 5 years while the age span of the prevalence rate is 15 years, there exists an increasing trend within each 15-year age group, which results in the increasing and decreasing trend of 'Mortality Rate' (dots) with the increase of age as shown in the lower four subgraphs of Figure 4.1. This is clearly not the case in real situations and could be improved if prevalence rates of a 5-year age span are available. From Figure 4.1, we observe that mortality rates in younger ages are underestimated under both models, since we try to find a mortality model for 74 years (from age 20 to 94). This may cause the underestimation of transition intensities and may be improved in further studies.

For both male and female mortality models of stroke and female mortality models of heart attack, mortality rates are underestimated using Weibull models at younger ages and overestimated at older ages. On the opposite, Gompertz models show a better fit in around first fifty years from age 20 and slightly underestimate the mortality rates in the last twenty years, which leads to a better goodness of fit in general. For male mortality models of heart attack, there is a huge fluctuation of the mortality rate in the beginning years, which may cause the underestimation in both Gompertz and Weibull models in the first few years. Likewise, for male mortality models of cancer, mortality rates are underestimated using both models for the beginning ages and work well after around age 70. The Gompertz models and Weibull models provide similar curves for female transition intensities of cancer and are close to the 'Mortality Rate' (dots).

Model	Gompertz $f(x) = exp(\beta_1 + \beta_2 x)$	Weibull $f(x) = \beta_1 x^{\beta_2}$
	Males	
Coefficients (95% CI)	$\hat{\beta}_1$ =-8.01 (-8.542, -7.477) $\hat{\beta}_2$ =0.05229 (0.04607, 0.05852)	$\hat{\beta}_1$ = 1.741e-15 (-6.011e-15, 9.492e-15) $\hat{\beta}_2$ = 6.894 (5.901, 7.888)
SSE	0.001312	0.002773
R^2	0.942	0.8773
Adjusted R^2	0.9412	0.8756
RMSE	0.00424	0.006164
	Females	
Coefficients (95% CI)	$\hat{\beta}_1$ =-8.531 (-9.049, -8.012) $\hat{\beta}_2$ =0.05901 (0.05301, 0.065)	$\hat{\beta}_1$ = 14.43e-18 (-1.578e-17, 2.465e-17) $\hat{\beta}_2$ = 8.25 (7.235, 9.265)
SSE	0.001076	0.002578
R^2	0.9666	0.9211
Adjusted R^2	0.9666	0.92
RMSE	0.00384	0.005943

Table 4.2: Estimated parameters of mortality models (from Stroke to Death due to Stroke)

4.2 Transition Intensities

In this section, transition intensities under different models are calculated and compared. Under the Markovian multi-state models, transition intensities can be obtained based on various continuous mortality models. Estimated beta parameters of Gompertz model and Weibull model are given in Tables 4.1-4.4; transition probabilities given by the equations (2.15)-(2.20) are used to estimate the transition under both Gompertz and Weibull models (see Section 2.3.2 for details). Extra mortalities are assumed to be zero (i.e., $\gamma^s = \gamma^a = \gamma^c = 0$), since the transition probability from a disease to death due to other causes should be very close to that of death from healthy. Using the iterative approach described in Section 2.3.2, the transition intensities are calculated and given in Tables 4.5 and 4.6. In addition, the incidence rates in Table 3.6 are also provided and compared with the calculated transition intensities.

In Tables 4.5 and 4.6, the transition intensities obtained using Gompertz model and Weibull model are compared with the incidence rates of stroke and heart attack given in Table 3.6. The comparisons are also presented in Figure 4.2. From the results, transition intensities of stroke for males are similar under both models in the first four age groups; after 80 years old, the transition intensity under Weibull model is larger than that under Gompertz model. Transition intensities of stroke for females show the same trend with the

Model	Gompertz $f(x) = exp(\beta_1 + \beta_2 x)$	Weibull $f(x) = \beta_1 x^{\beta_2}$
	Males	
Coefficients (95% CI)	$\hat{\beta}_1$ =-9.048 (-10.01, -8.085) $\hat{\beta}_2$ =0.07067 (0.0597, 0.08165)	$\hat{\beta}_1$ = 1.897e-14 (-7.865e-14, 1.166e-13) $\hat{\beta}_2$ = 6.434 (5.285, 7.584)
SSE	0.006826	0.008687
R^2	0.8227	0.7744
Adjusted R^2	0.8203	0.7713
RMSE	0.00967	0.01091
	Females	
Coefficients (95% CI)	$\hat{\beta}_1$ =-6.605 (-7.095, -6.114) $\hat{\beta}_2$ =0.04035 (0.03444, 0.04625)	$\hat{\beta}_1$ = 11.22e-15 (-5.62e-15, 8.053e-15) $\hat{\beta}_2$ = 7.09 (5.836, 8.343)
SSE	0.00417	0.01181
R^2	0.9354	0.8171
Adjusted R^2	0.9346	0.8146
RMSE	0.007558	0.01272

Table 4.3: Estimated parameters of mortality models (from Heart Attack to Death due to Heart Attack)

male ones given in the upper two subgraphs of Figure 4.2. Observed from the lower two subgraphs of Figure 4.2, male transition intensities of heart attack are similar under both models, while female intensities are similar under both models before age 80; after 80 years old, female intensities under Weibull model are much larger than those under Gompertz model. In Figure 4.2, it is obvious that male intensities under Weibull models are larger than the results under Gompertz in the late years, which is due to the overestimation of Weibull models in old ages. Shown in Tables 4.5 and 4.6, all transition intensities are less than the incidence rates collected by CCDSS. As discussed in Section 4.1, the Gompertz mortality model shows a better fit compared to the Weibull model. However, the transition intensities under Weibull models give a closer result to the incidence rates of stroke and heart attack.

Comparing the male and female transition intensities, male intensities of heart attack are always larger than female intensities under both models, which is consistent with the mortality rates data given in Table 3.7. As for the intensities of stroke, male intensities of stroke are lower than female intensities in the starting and late ages, and larger then female ones in the middle ages, which shows the same trend as mortality rates data of stroke as well.

Model	Gompertz $f(x) = exp(\beta_1 + \beta_2 x)$	Weibull $f(x) = \beta_1 x^{\beta_2}$
	Males	
Coefficients (95% CI)	$\hat{\beta}_1$ =-10.09 (-11.08, -9.093) $\hat{\beta}_2$ =0.08673 (0.07555, 0.09791)	$\hat{\beta}_1$ = 3.323e-17 (-1.379e-16, 2.044e-16) $\hat{\beta}_2$ = 7.936 (6.788, 9.083)
SSE	0.009983	0.01179
R^2	0.8918	0.8722
Adjusted R^2	0.8903	0.8705
RMSE	0.01169	0.01271
	Females	
Coefficients (95% CI)	$\hat{\beta}_1$ =-13 (-13.3, -12.69) $\hat{\beta}_2$ =0.1222 (0.1189, 0.1256)	$\hat{\beta}_1$ = 2.19e-23 (-8.656e-23, 1.304e-22) $\hat{\beta}_2$ = 11.16 (10.06, 12.25)
SSE	0.0007861	0.009486
R^2	0.9959	0.95071
Adjusted R^2	0.9959	0.95
RMSE	0.003282	0.0114

Table 4.4: Estimated parameters of mortality models (from Cancer to Death due to Cancer)

Gompertz model				
Age Group	Stroke (per 100,000)	Heart Attack (per 100,000)	Incidence rate (per 100,000) Stroke	Heart Attack
20-34	8.13	2.02	22	6
35-49	31.97	36.41	81	88
50-64	146.83	236.92	295	341
65-79	572.27	605.08	855	691
80+	1,315.72	1,223.68	2074	1400
Weibull model				
Age Group	Stroke (per 100,000)	Heart Attack (per 100,000)	Incidence rate (per 100,000) Stroke	Heart Attack
20-34	8.04	2.01	22	6
35-49	30.93	35.83	81	88
50-64	141.53	233.38	295	341
65-79	578.95	615.45	855	691
80+	1,677.56	1,269.30	2074	1400

Table 4.5: Transition intensities for males under Gompertz and Weibull models

4.3 Graduation of Intensities

In Section 4.2, transition intensities for each age group are obtained. Graduation functions can be used to help obtain the transition intensities for each age, which would make the

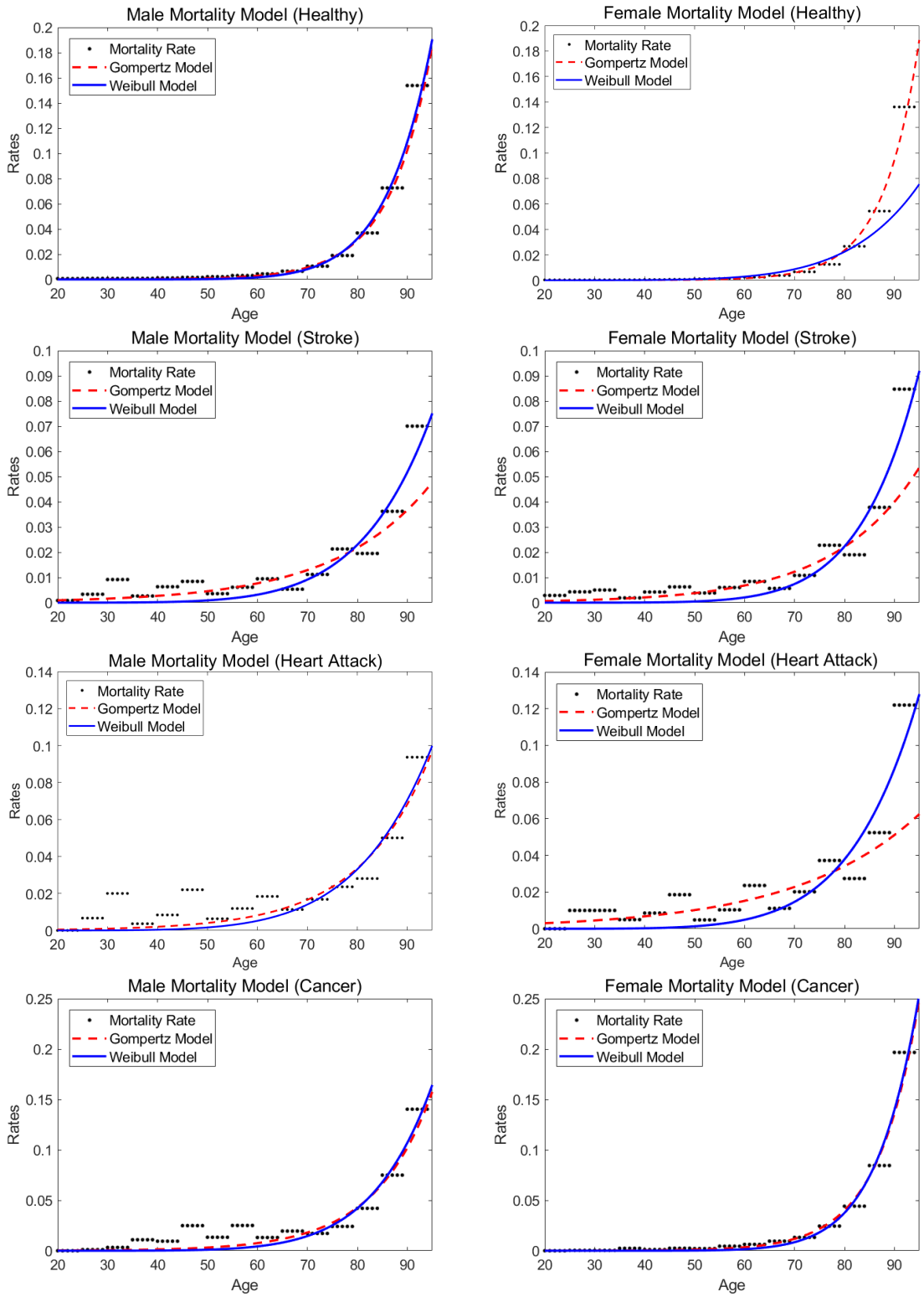

Figure 4.1: Estimated mortality models (Gompertz v.s. Weibull)

Gompertz model

Age Group	Stroke (per 100,000)	Heart Attack (per 100,000)	Incidence rate (per 100,000) Stroke	Heart Attack
20-34	9.47	0.69	25	2
35-49	35.70	9.66	74	24
50-64	113.17	72.59	218	119
65-79	417.96	260.06	647	352
80+	1,217.16	673.08	1913	1019

Weibull model

Age Group	Stroke (per 100,000)	Heart Attack (per 100,000)	Incidence rate (per 100,000) Stroke	Heart Attack
20-34	9.39	0.67	25	2
35-49	34.68	9.05	74	24
50-64	106.80	66.15	218	119
65-79	405.93	252.09	647	352
80+	1,458.80	974.64	1913	1019

Table 4.6: Transition intensities for females under Gompertz and Weibull models

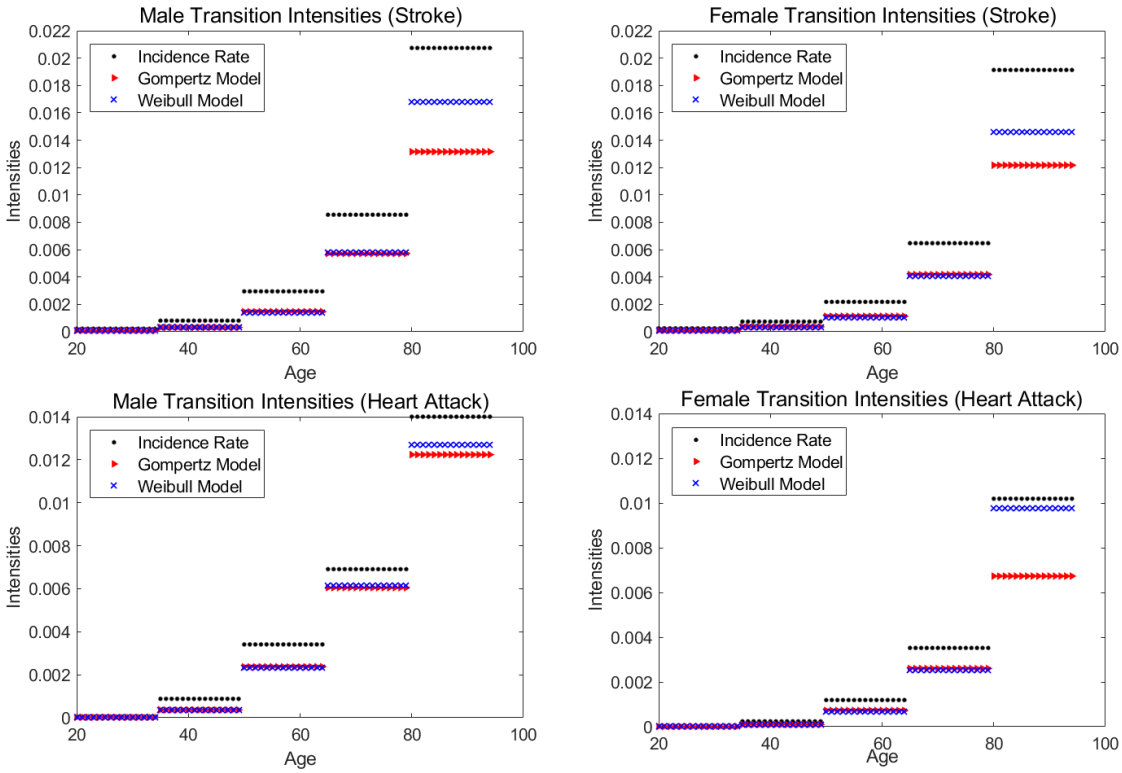

Figure 4.2: Comparison of transition intensities under Gompertz and Weibull models

pricing of CI policies more smooth. Following the idea of Baione and Levantesi (2014), in this section we obtain the transition intensities using graduated prevalence rates of each age and for comparison we also obtain transition intensities for each age using the graduation

method directly. Since the Gompertz model fits the mortality rates better according to our previous study, we use the Gompertz model to calculate the intensities and premium rates in this section and the section it follows. Three approaches are used to obtain the prevalence rates and the transition intensities of stroke and heart attack. We first describe these approaches. The results for the prevalence rates and the transition intensities by different approaches are graphed and explained later.

- Approach 1: "Raw" prevalence rates and "Raw" transition intensities

  The "Raw" prevalence rates of stroke and heart attack for males and females are shown in Table 3.1. Note that these rates are constant within each age span. They are compared with corresponding graduated prevalence rates based on Approach 2 where different graduation functions are used to perform the graduation on "Raw" ones.

  The incidence rates of cancer, stroke and heart attack for males and females shown in Tables 3.5-3.6 are considered as "Raw" transition intensities. They are also constant within each age span. They are compared with graduated transition intensities by age using shape-preserving interpolant function calculated by Approach 3; note that the base transition intensities in each age span are calculated in Section 4.2 based on the Gompertz mortality model assumption.

- Approach 2: Transition intensities by age from graduated prevalence rates

  According to Haberman and Pitacco (1998), graduation is a method that transition probabilities are adjusted for further inferences and practical calculations. It ensures the resulting multi-state model shows some degree of smoothness so that the transition intensities and premium rates calculated from graduated models also have the smoothness.

  In this project, we use power, linear, cubic spline and shape-preserving interpolant as graduation functions to perform graduations. They are built-in graduation functions in Matlab. Similar to the cubic spline function, shape-preserving interpolant function is actually a piece-wise cubic Hermite interpolating polynomial. For example, the shape-preserving interpolant use the cubic function on interval (x_1, x_2) in the following form

  $$f(x) = a(x - x_1)^3 + b(x - x_1)^2 + c(x - x_1) + d, \quad x_1 \leqslant x \leqslant x_2.$$

  The main difference between the cubic functions used for cubic spline and shape-preserving interpolant is that the cubic spline function gives a smoother curve as it ensures a continuous second derivative at sample points, while the shape-preserving function guarantees the resulting curve not being oscillated between sample points[1].

[1] Details are provided on MathWorks Documentation: https://www.mathworks.com/help/matlab/ref/pchip.html

Applying the four graduation functions mentioned above, we perform graduations in "Raw" prevalence rates of stroke and heart attack for both males and females, respectively, to obtain corresponding graduated prevalence rates by age. These graduated prevalence rates by age are compared with "Raw" ones in Figure 4.3.

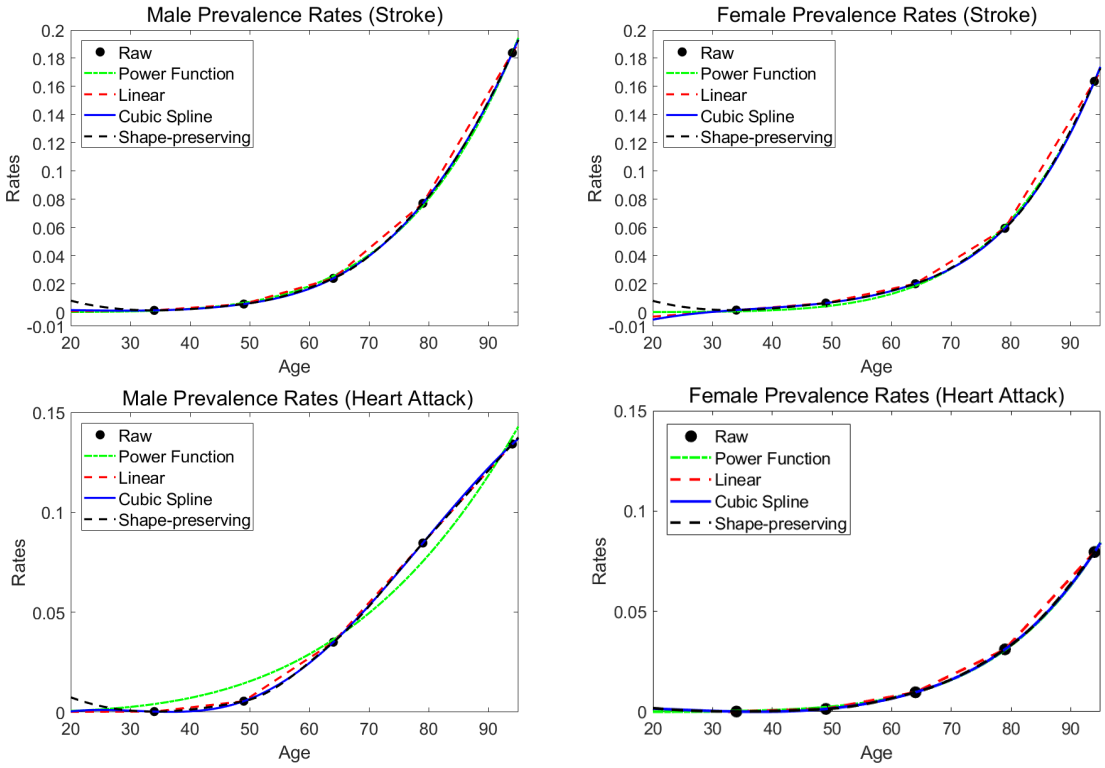

Figure 4.3: Graduated prevalence rates of stroke and heart attack

We then use the suitable graduated prevalence rates by age for stroke and heart attack of males and females to calculate the transition intensities. Note that the mortality rates calculated in Section 4.1 and the incidence rates of cancer calculated under the same graduation function are used when computing the transition intensities.

- Approach 3: Graduation of transition intensities under the Gompertz model

  The transition intensities of stroke and heart attack for each age span of both males and females under Gompertz mortality model assumption are obtained, respectively, in Section 4.2. We apply the graduation method to these rates using both the cubic spline and shape-preserving interpolant functions to get the corresponding intensities by age.

  These graduated transition intensities are compared with the "Raw" ones (described in Approach 1) in Figure 4.4. Comparisons of transition intensities by age under three approaches are also done and shown in Figure 4.5 for both stroke and heart attack,

males and females, respectively. In Figure 4.5, "Raw" transition intensities by age are described by Approach 1, transition intensities calculated by age by Approach 2 are based on the graduated prevalence rates using the shape-preserving interpolant function, and graduated transition intensities by age by Approach 3 are graduated based on the ones by age spans under Gompertz model using the shape-preserving interpolant function.

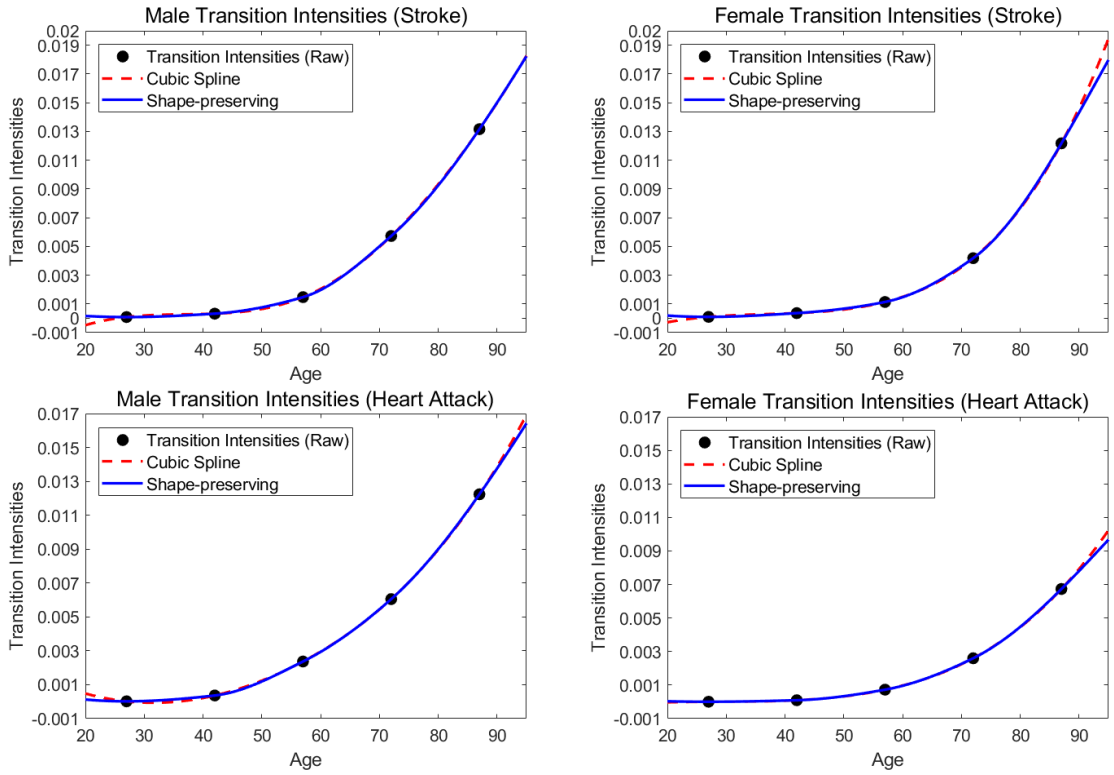

Figure 4.4: Graduated transition intensities of stroke and heart attack

Figure 4.3 displays comparisons of the "Raw" prevalence rates of stroke and heart attack by age, described in Approach 1, and the graduated prevalence rates based on "Raw" ones using four different graduation functions (Approach 2). Note these prevalence rates are respectively for males and females of each age from age 20 to age 95. It can be observed that the impact of using different graduation functions on the resulting graduated prevalence rates is minimal as the obtained prevalence rates curves are very similar in all the cases except for the curves of males' prevalence rates of heart attack under linear interpolant and power function. From Figure 4.3, the shape-preserving interpolant shows decreasing trend before around 30 years old for all four subgraphs, and there are some negative values of females' prevalence rates of stroke under linear and cubic spline interpolants. Note that the unrealistic shape of these curves at younger ages is due to the fact that there is a lack of data before age 20. As seen in Figure 4.3, all the prevalence rates curves show an increasing

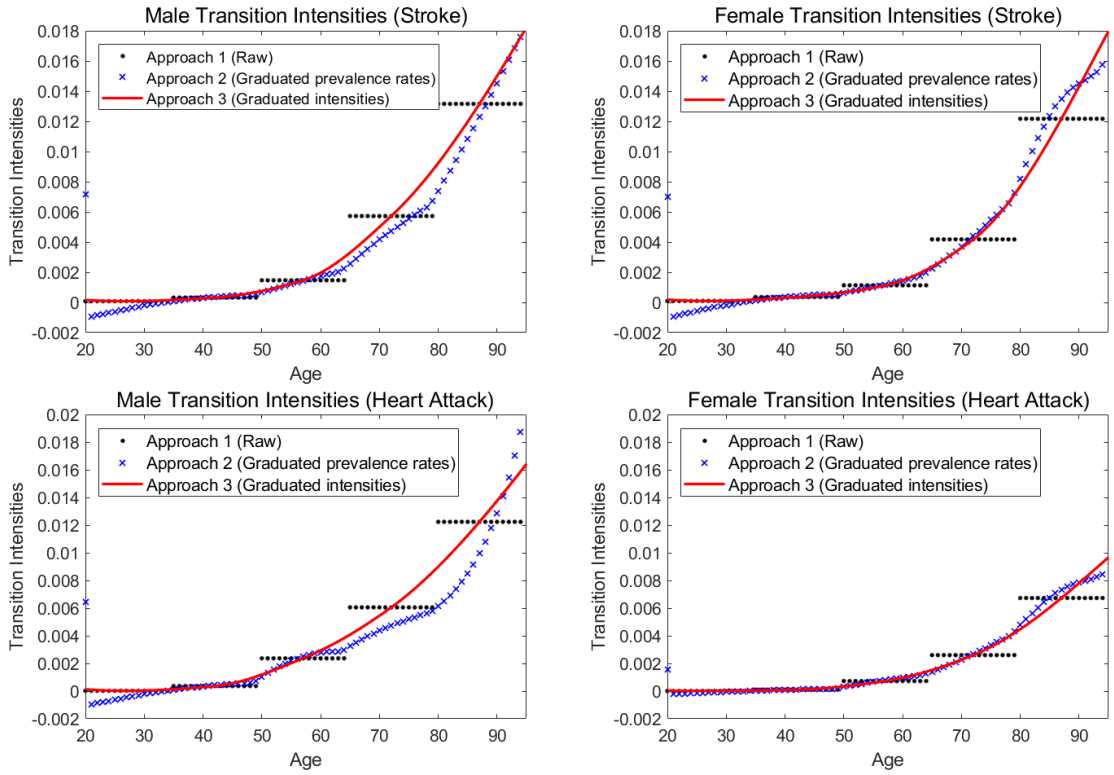

Figure 4.5: Transition intensities of stroke and heart attack under different approaches

trend after around 32 years old. We can also observe that the graduated prevalence rates of stroke for males and females are similar, while the graduated prevalence rates of heart attack for females are always less than those for males.

Figure 4.4 shows the graduated transition intensities of stroke and heart attack by age under both the cubic spline and shape-preserving interpolant functions (Approach 3) with the corresponding "Raw" ones (described in Approach 1). These transition intensities are respectively for males and females of each age from age 20 to age 95. It can be seen from the figure that the curves of graduated transition intensities of stroke for both genders are similar with both cubic spline and shape-preserving interpolant graduation functions, except that the cubic spline method produces larger intensity values than the shape-preserving interpolant does in extreme elderly age group and some negative values before age 30. We also observe that graduated transition intensities of heart attack for males decrease in early ages, and then increase in age under the cubic spline method, while the curve for females under the shape-preserving method is relatively stable in early ages and increasing all the way until age 95.

Considering our results shown in Figures 4.3 and 4.4 with Approaches 2 and 3 together, we see that the shape-preserving method gives transition intensities, in all cases (both stroke and heart attack, and both genders), closer to their corresponding "Raw" transition intensities than the other methods of graduation. Thus, we apply the shape-preserving inter-

polant function to graduation of prevalence rates, incidence rates and transition intensities along with the Gompertz mortality models to find transition intensities by age, that is, the piece-wise constant transition intensities of cancer, stroke and heart attack, respectively, for each age from age 20 to age 95. The numerical results are presented in Tables A.1-A.3 in Appendix A.

Figure 4.5 shows transition intensities of stroke and heart attack under three approaches. "Raw" transition intensities described in Approach 1 are compared with transition intensities by age obtained from graduated prevalence rates (Approach 2) and with the graduated transition intensities by age (Approach 3), where the shape-preserving interpolant function is used for both graduations. These transition intensities are respectively for males and females of each age from age 20 to age 95. From Figure 4.5, we observe some negative values of transition intensities by Approach 2 at early ages. As we have explained earlier, they are caused by the lack of data before the smallest observed age 20. Starting from age 34, all the transition intensities are positive, and moreover graduated transition intensity values are close to transition intensities obtained from graduated prevalence rates, except for the male transition intensities of heart attack in the late ages. For graduated transition intensities of heart attack, males' ones are always larger than those for females after age 30; this is also the case for prevalence rates, incidence rates and mortality rates of heart attack when comparing male rates with female ones. For graduated transition intensities of stroke, the intensities are about the same for males and females at first few ages and then, in middle ages, males' intensities are larger than those for females and in late ages, both males' intensities and female ones are similar. This shows a similar movement with mortality rates of stroke where only in middle ages, mortality rates for males are larger than female ones. However, the prevalence rates and incidence rates of stroke show different trends, where both male prevalence and incidence rates are larger than the corresponding female rates in late ages.

4.4 Comparison of Premium Rates

In this section, premium rates are calculated using transition intensities by age under three approaches. A general pricing model was introduced in Section 2.4 with continuous-time multi-state models. Here, we use the sum of sub-integrals in age to find premium rates, since our transition intensities are obtained for each integer age.

Under the policies in Chapter 2, the basic CI insurance plan is a 25-year term Stand Alone (SA) insurance ($n = 25$) with a payment of \$1 payable immediately for Cancer, Stroke and Heart Attack. The issuing age x is between ages 35 and 55. A Full Accelerated (FA) policy is also considered where an additional \$1 benefit is paid if death occurs due to other causes. A 'with Rider' option is included in the comparison, which is the FA policy with a rider increasing the payment to \$2 in case of Transportation Accidental death.

The force of interest is assumed constant at 5% so the discount factor can be expressed as

$$v_t = e^{-\delta t} = e^{-0.05t}, \quad t \geqslant 0.$$

With the transition intensities from healthy to death given in Section 3.3, and, those from healthy to cancer, stroke or heart attack under graduation approaches given in Section 4.3, the net single premium (NSP) for insured at different ages, when the CI policy is issued, can be calculated. Based on the pricing formulas given in equations (2.23)-(2.25) for SA, FA and 'with Rider' options, respectively, we can find the formulae for NSP under the Gompertz models. Here, we only consider the Gompertz models, which are the same models used for graduated transition intensities. For an insured aged x, the NSPs for a 25-year term CI SA, CI FA and CI FA 'with Rider' policies can be calculated, respectively, by

$$\begin{aligned}
P_x^{SA} &= \int_0^{25} {}_tp_x^{HH} \left[\mu^{HC}(x+t) + \mu^{HS}(x+t) + \mu^{HA}(x+t)\right] e^{-\delta t} dt \\
&= \sum_{h=0}^{24} \int_h^{h+1} {}_{t-h}p_{x+h}^{HH} \left[\mu^{HC}(x+h) + \mu^{HS}(x+h) + \mu^{HA}(x+h)\right] e^{-\delta t} dt, \\
P_x^{FA} &= \int_0^{25} {}_tp_x^{HH} \left[\mu^{HC}(x+t) + \mu^{HS}(x+t) + \mu^{HA}(x+t) + \mu^{HD}(x+t)\right] e^{-\delta t} dt \\
&= \sum_{h=0}^{24} \int_h^{h+1} {}_{t-h}p_{x+h}^{HH} \left[\mu^{HC}(x+h) + \mu^{HS}(x+h) + \mu^{HA}(x+h) + e^{\beta_1^d + \beta_2^d(x+t)}\right] e^{-\delta t} dt, \\
P_x^{Acc} &= \int_0^{25} {}_tp_x^{HH} \left[\mu^{HC}(x+t) + \mu^{HS}(x+t) + \mu^{HA}(x+t) + \mu^{HD}(x+t) + \mu^{Acc}(x+t)\right] \\
&\quad \times e^{-\delta t} dt \\
&= \sum_{h=0}^{24} \int_h^{h+1} {}_{t-h}p_{x+h}^{HH} \left[\mu^{HC}(x+h) + \mu^{HS}(x+h) + \mu^{HA}(x+h) + e^{\beta_1^d + \beta_2^d(x+t)}\right. \\
&\quad \left. + \mu^{Acc}(x+h)\right] e^{-\delta t} dt,
\end{aligned}$$

where transition probability ${}_{t-h}p_{x+h}^{HH}$ under the Gompertz model is given by

$${}_{t-h}p_{x+h}^{HH} = e^{-\left\{ \frac{e^{\beta_1^d}}{\beta_2^d} \left[e^{\beta_2^d(x+t)} - e^{\beta_2^d(x+h)}\right] + \left[\mu^{HC}(x+h) + \mu^{HS}(x+h) + \mu^{HA}(x+h)\right](t-h) \right\}}, \quad h \leqslant t \leqslant h+1. \tag{4.1}$$

Transition intensities, $\mu^{Hi}(x)$, for all ages (i.e., $x = 35, 36, \ldots, 79$) under Approaches 2 and 3 can be found in Tables A.2-A.3, for $i = \text{S, A}$. The graduated transition intensities of cancer, $\mu^{HC}(x)$, for all ages under both approaches can be found in Table A.1. Mortality rates, $\mu^{Acc}(x)$, for all ages are given in Table 3.8. Using the pricing formulae above and the transition probability expression in equation (4.1), the NSP rates can be calculated.

Net Single Premium Rates per $1000

Approach 1 (Raw Transition Intensities)

	Males			Females		
Age	SA	FA	with Rider	SA	FA	with Rider
35	56.66	65.09	66.28	63.94	67.54	67.90
40	76.83	91.69	93.01	77.68	84.77	85.16
45	115.07	138.17	139.46	105.44	118.18	118.59
50	170.50	207.84	209.14	144.61	168.35	168.73
55	197.02	258.67	259.94	163.72	208.54	208.98

Approach 2 (Graduated Prevalence Rates)

	Males			Females		
Age	SA	FA	with Rider	SA	FA	with Rider
35	51.02	59.54	60.74	58.74	62.36	62.72
40	77.60	92.01	93.33	80.22	87.12	87.50
45	112.36	136.14	137.45	105.93	118.86	119.26
50	154.96	193.56	194.88	136.08	159.96	160.35
55	197.56	259.62	260.90	167.95	211.46	211.89

Approach 3 (Graduated Transition Intensities)

	Males			Females		
Age	SA	FA	with Rider	SA	FA	with Rider
35	51.37	59.89	61.08	58.59	62.21	62.57
40	78.51	92.82	94.13	80.08	86.97	87.35
45	114.89	138.33	139.63	106.23	119.13	119.54
50	158.02	195.78	197.08	136.14	160.04	160.43
55	203.28	263.19	264.46	167.98	211.58	212.02

Table 4.7: NSP rates under three graduation approaches

The NSP rates (per $1000) are displayed in Table 4.7 for males and females insured issuing at 35, 40, 45, 50 and 55 with SA, FA policies and with a rider of Transportation Accidental death. A comparison of these premium rates is shown in Figure 4.6. From Table 4.7, the premium rates for FA policies are about 20% more expensive than those for SA policies under all three approaches. Premium rates for FA policies increase with the issuing age, which is due to the increase of mortality rates in age. FA premium rates are about the same as the corresponding premium rates with the rider. Comparing premium rates of the insured at different ages, NSP rates under Approaches 2 and 3 are close to each other and both are less than the premium rates under Approach 1 for insureds' ages at 35 and 50. Ages 35 and 50 under Approach 1 are at the beginning of age groups 35-49 and 50-64, which would cause the transition intensities to be larger than those under Approaches 2 and 3 (as shown in Figure 4.5), and lead to higher premium rates. For the insureds at issuing ages 40 and 55 under Approach 1, the NSP rates are lower than those under Approaches 2 and 3. Premium rates of Approach 3 for male insured aged 55 are around $5 more expensive than the rates under Approach 1 and Approach 2, which may due to the overestimated transition

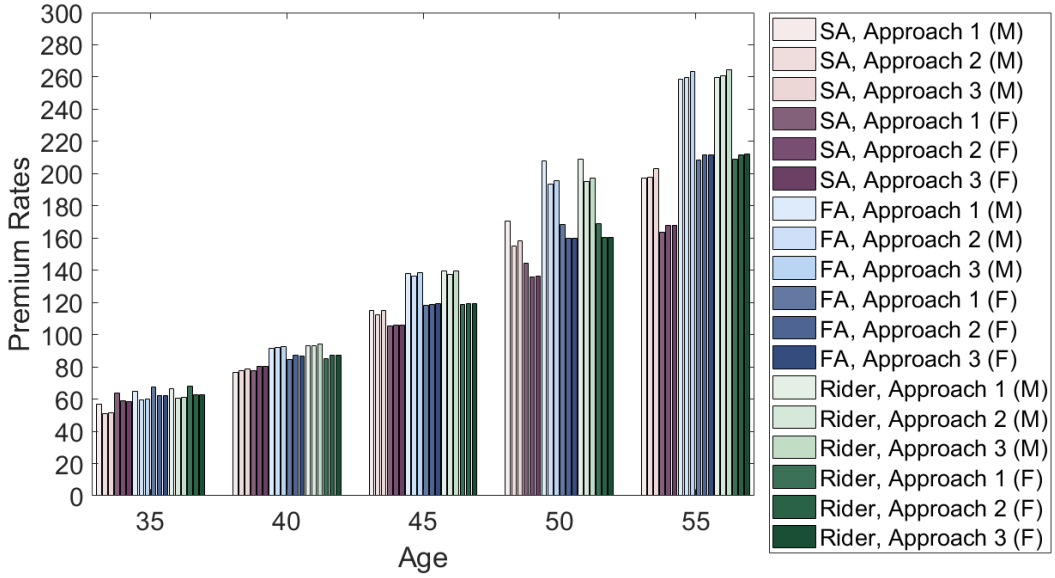

Figure 4.6: NSP rates for males and females under three approaches

intensities for ages around 70 to 80 under Approach 3. Premium rates for females in elder ages, for all policies, are always lower than the corresponding rates for males.

4.5 Impact of the Force of Interest

After the analysis of premium rates under different graduation approaches, we now study the impacts of the force of interest on the premium rates of CI policies in this section. We set the premium rates under Approach 3 given in Table 4.7 as the base scenario.

Under the base assumption, a 5% constant force of interest is used for pricing. We test scenarios in which the constant force of interest ($\delta = 5\%$) is increased and decreased, respectively, by 10% and 20% in the pricing formulas with graduated transition intensities under Approach 3. The resulting of premium rates under these scenarios are displayed in Table 4.8 for males and Table 4.10 for females. Tables 4.9 and 4.11 show the percent change of premium rates for males and females, respectively.

From the results shown in Tables 4.9 and 4.11, we find that with the interest rate being decreased by 20%, the premium rates increase by about 14% in general for all policies, both males and females, and issuing ages from 35 to 55; while premium rates increase by about 6.5% with the interest rate being decreased by 10%. Compared with the decreased interest rates, if the interest rate goes up by 10%, the premium rates decrease by around 6%, which is a little lower than the increment of the "-10%" scenario. Likewise, the premium rates decrease by about 12% with the interest rate being increased by 20%. Moreover, it can be observed that male premium rates have changed more than those for females in younger ages for all types of policies. Comparing the premium rates for SA, FA and 'with Rider'

Male premium rates (per $1000)

Scenarios	**-20%**	**-10%**	**Base**	**+10%**	**+20%**
Age			SA		
35	59.60	55.30	51.37	47.77	44.48
40	90.47	84.23	78.51	73.25	68.43
45	131.08	122.64	114.89	107.74	101.16
50	178.26	167.74	158.02	149.04	140.73
55	226.84	214.62	203.28	192.75	182.97
Age			FA		
35	69.54	64.50	59.89	55.67	51.81
40	107.12	99.66	92.82	86.55	80.80
45	158.21	147.85	138.33	129.57	121.51
50	221.73	208.23	195.78	184.28	173.67
55	295.43	278.69	263.19	248.84	235.54
Age			with Rider		
35	70.88	65.76	61.08	56.80	52.89
40	108.57	101.04	94.13	87.80	81.99
45	159.63	149.21	139.63	130.81	122.70
50	223.14	209.59	197.08	185.54	174.88
55	296.80	280.00	264.46	250.06	236.72

Table 4.8: Male premium rates under different scenarios of the force of interest

Percent change of male premium rates

Scenarios	**-20%**	**-10%**	**+10%**	**+20%**
Age			SA	
35	16.02%	7.65%	-7.00%	-13.41%
40	15.25%	7.29%	-6.69%	-12.83%
45	14.09%	6.75%	-6.22%	-11.95%
50	12.81%	6.15%	-5.69%	-10.95%
55	11.59%	5.58%	-5.18%	-9.99%
Age			FA	
35	16.12%	7.70%	-7.04%	-13.48%
40	15.41%	7.37%	-6.75%	-12.95%
45	14.37%	6.88%	-6.33%	-12.16%
50	13.25%	6.36%	-5.87%	-11.29%
55	12.25%	5.89%	-5.45%	-10.51%
Age			with Rider	
35	16.04%	7.66%	-7.00%	-13.41%
40	15.34%	7.33%	-6.73%	-12.90%
45	14.33%	6.86%	-6.31%	-12.12%
50	13.23%	6.35%	-5.86%	-11.27%
55	12.23%	5.88%	-5.44%	-10.49%

Table 4.9: Percent change of male premium rates from "Base"

Female premium rates (per $1000)					
Scenarios	**-20%**	**-10%**	**Base**	**+10%**	**+20%**
Age			SA		
35	66.93	62.58	58.59	54.92	51.54
40	91.08	85.35	80.08	75.23	70.76
45	120.26	112.95	106.23	100.02	94.30
50	153.24	144.35	136.14	128.55	121.53
55	187.85	177.53	167.98	159.13	150.92
Age			FA		
35	71.20	66.51	62.21	58.26	54.62
40	99.19	92.82	86.97	81.59	76.64
45	135.41	126.93	119.13	111.96	105.34
50	181.22	170.20	160.04	150.68	142.04
55	238.70	224.59	211.58	199.57	188.47
Age			with Rider		
35	71.60	66.89	62.57	58.60	54.95
40	99.61	93.22	87.35	81.96	76.99
45	135.85	127.36	119.54	112.34	105.71
50	181.65	170.61	160.43	151.05	142.39
55	239.18	225.05	212.02	199.99	188.87

Table 4.10: Female premium rates under different scenarios of the force of interest

Percent change of female premium rates				
Scenarios	**-20%**	**-10%**	**+10%**	**+20%**
Age		SA		
35	14.24%	6.82%	-6.27%	-12.03%
40	13.73%	6.58%	-6.06%	-11.64%
45	13.21%	6.33%	-5.84%	-11.23%
50	12.57%	6.03%	-5.58%	-10.73%
55	11.82%	5.69%	-5.27%	-10.16%
Age		FA		
35	14.46%	6.92%	-6.35%	-12.19%
40	14.05%	6.73%	-6.18%	-11.88%
45	13.66%	6.54%	-6.03%	-11.58%
50	13.23%	6.34%	-5.85%	-11.25%
55	12.81%	6.15%	-5.68%	-10.92%
Age		with Rider		
35	14.43%	6.91%	-6.35%	-12.18%
40	14.03%	6.72%	-6.18%	-11.86%
45	13.65%	6.54%	-6.02%	-11.57%
50	13.23%	6.34%	-5.85%	-11.24%
55	12.81%	6.15%	-5.68%	-10.92%

Table 4.11: Percent change of female premium rates from “Base”

policies, we find that the premium rates for FA policies are most sensitive to a change in interest rate for each age and gender than the other types of policies. Premium rates for SA policies change the least among all policies.

Chapter 5

Conclusion

In this project, critical illness insurance plans are priced using Markovian multiple state models. First, we describe critical illness insurance products available in the Canadian market and compare them with CI products offered in some other countries. We present those CI products that are commonly provided by Canadian, U.S. and Asian insurers. Previous studies in Markovian multiple state models with applications and methods for estimating transition intensities from prevalence rates are reviewed. Parametric and non-parametric mortality models and methods of graduation are also discussed.

We then design several typical CI insurance plans to be studied in this project. The basic coverage of cancer, stroke and heart attack are considered in the Stand Alone policy. We also propose Full Accelerated policy that includes a death benefit and the optional Transportation Accidental death rider. A Markovian multi-state model is introduced to describe the health status under our design and the required transition probabilities and intensities are calculated for the pricing of CI plans.

Inspired by Baione and Levantesi (2014), a pricing model is studied in this project with a Markovian multi-state model that considers three critical illnesses (Cancer, Stroke and Heart Attack). It extends the model in Baione and Levantesi (2014) in the sense that they consider all the illnesses covered as one state, while in our model we set separate states for each illness covered so that the benefit payments can be different and more data can be used for pricing accurately. In their paper, they estimate transition intensities of multi-state models based only on the prevalence rate. In this project, we estimate the transition intensities using prevalence rates and incidence rates, which provide flexibility in using the existing data and may give more reliable pricing results. Besides, since Canadian health data and statistics are used in this project, the pricing model we introduce and illustrate may be suitable for the Canadian insurance market.

To obtain the transition intensities using prevalence rates, mortality rates by specified causes and continuous mortality models are required. Following the idea presented in Baione and Levantesi (2018), we obtain transition intensities from prevalence rates with different mortality models. Two mortality models are considered; they are Gompertz model and

Weibull model and both are two-parameter models. We fit Canadian mortality rates of healthy, stroke and heart attack to these models; parameters are estimated and goodness of fit tests are performed. We find that Gompertz models provide a better goodness of fit in general. We then obtain and compare transition intensities based on these mortality models with their estimated parameters. We conclude that these two popular continuous-time mortality models can be easily used in pricing the CI products under our model due to their simple analytical form and the estimation and inference can be efficiently done using existing software packages.

We compare transition intensities under fitted Gompertz and Weibull mortality models with Canadian health data. We find that both models produce similar transition intensities from stroke to death due to stroke and from heart attack to death due to heart attack in most years and both intensities are smaller than corresponding incidence rates of stroke and heart attack from data. Since Canadian health data has an age span of 15 years, which implies that an assumption of constant transition intensities in each age group is unrealistic; methods of graduation are then used and investigated for more smooth pricing. Similar to Baione and Levantesi (2014), we obtain transition intensities by each age from graduated prevalence rates (Approach 2) and also through graduated transition intensities (Approach 3). Different methods of graduation are considered, and the obtained transition intensities for each age are compared with the 'Raw' intensities. In our empirical study, among the graduation functions considered, we find that the shape-preserving interpolant method performs the best. In addition, we find that graduation can be easily implemented using existing software packages.

Finally, with estimated transition intensities under different approaches and calculated transition probabilities under our Markovian multiple state model, premium rates are calculated and compared within and across three approaches. The net single premium rates are calculated for male and female insureds with ages at issue ranging from age 35 to age 55. A comparison of premium rates indicates that with relatively more smooth transition intensities by each age, premium rates might be lower for younger insureds and be higher for older insureds. In addition, graduated transition intensities under Approach 3 always provide similar premium rates with transition intensities obtained from graduated prevalence rates under Approach 2.

Moreover, we examine the impact of forces of interest on NSP rates under Approach 3. It shows that with the base force of interest being decreased by 20%, premium rates are around 13% higher for male insureds and 14% higher for female insureds. Premium rates of Full Accelerated policy have the most significant changes with the change of the interest rate among all policies. We conclude that the pricing based on our model assumptions and all the estimation methods is quite efficient and reliable, and we believe it can be easily applied in Canadian insurance market.

However, there are some limitations due to the lack of data. Firstly, as mentioned in Chapter 3, the health data of prevalence, incidence rates and the mortality rates are not provided from the same sources. Secondly, some critical illness insurance products (e.g., Aviva's critical illness product mentioned in Section 2.1) consider payments at different stages of diagnosis of cancer, which is not addressed in our current model due to the difficulties in obtaining statistics about the number of patients under each stage of a specified cancer. These two limitations can be relaxed if appropriate and reliable data are available. The estimation of transition intensities may also be improved using Gompertz mortality models with more parameters. In addition, we only consider the constant force of interest in this project. In future study, we could consider stochastic models for the force of interest and other mortality models to price CI products that would be closer to the real market conditions.

Bibliography

F. Baione and S. Levantesi. A health insurance pricing model based on prevalence rates: Application to critical illness insurance. *Insurance: Mathematics and Economics*, 58: 174–184, 2014.

F. Baione and S. Levantesi. Pricing critical illness insurance from prevalence rates: Gompertz versus Weibull. *North American Actuarial Journal*, 22(2):270–288, 2018.

H. Booth and L. Tickle. Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, 3(1-2):3–43, 2008.

M. C. Christiansen. Multistate models in health insurance. *AStA Advances in Statistical Analysis*, 96(2):155–186, 2012.

A. C. Dash and D. L. Grimshaw. Dread disease cover- an actuarial perspective. *Journal of the Staple Inn Actuarial Society*, 33(1):149–193, 1993.

D. C. Dickson, M. R. Hardy, and H. R. Waters. *Actuarial mathematics for life contingent risks*. Cambridge University Press, 2013.

S. Haberman. Mortality studies: measurement of experience and comparison with a standard. *Journal of the Institute of Actuaries*, 109(2):203–223, 1982.

S. Haberman. Scientific underwriting: 'Stroke' as a model. *Journal of the Institute of Actuaries*, 110(3):445–455, 1983.

S. Haberman. Decrement tables and the measurement of morbidity: II. *Journal of the Institute of Actuaries*, 111(1):73–86, 1984.

S. Haberman and E. Pitacco. *Actuarial Models for Disability Insurance*. CRC Press, 1998.

B. L. Jones. Actuarial calculations using a Markov model. *Transactions of the Society of Actuaries*, 46:227–250, 1994.

D. A. Juckett and B. Rosenberg. Comparison of the Gompertz and Weibull functions as descriptors for human mortality distributions and their intersections. *Mechanisms of Ageing and Development*, 69(1-2):1–31, 1993.

S. A. Klugman, G. E. Willmot, and H. H. Panjer. *Loss models: From data to decisions*. John Wiley & Sons, 2008.

R. D. Lee and L. R. Carter. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419):659–671, 1992.

L. Meira-Machado, J. de Uña-Álvarez, C. Cadarso-Suarez, and P. K. Andersen. Multi-state models for the analysis of time-to-event data. *Statistical Methods in Medical Research*, 18(2):195–222, 2009.

E. Pitacco. Actuarial models for pricing disability benefits: Towards a unifying approach. *Insurance: Mathematics and Economics*, 16(1):39–62, 1995.

D. L. Wilson. The analysis of survival (mortality) data: fitting Gompertz, Weibull, and logistic functions. *Mechanisms of Ageing and Development*, 74(1-2):15–33, 1994.

Appendix A

Graduated incidence rates and transition intensities

Graduated Incidence Rates of Cancer (per 100,000)

Age	Males	Females	Age	Males	Females	Age	Males	Females
20	83.86	70.15	45	213.36	360.81	70	1,945.72	1,433.83
21	74.99	64.82	46	242.16	387.46	71	2,029.03	1,470.95
22	67.63	61.49	47	275.06	415.75	72	2,103.91	1,501.59
23	61.73	60.10	48	311.72	445.56	73	2,173.45	1,528.56
24	57.24	60.55	49	351.81	476.78	74	2,241.69	1,555.15
25	54.09	62.76	50	394.97	509.28	75	2,308.53	1,581.16
26	52.25	66.66	51	440.86	542.94	76	2,373.87	1,606.42
27	51.64	72.17	52	489.15	577.65	77	2,437.62	1,630.74
28	52.23	79.19	53	539.49	613.29	78	2,499.69	1,653.93
29	53.95	87.66	54	591.54	649.74	79	2,559.97	1,675.83
30	56.75	97.48	55	644.95	686.87	80	2,618.37	1,696.23
31	60.58	108.59	56	699.39	724.58	81	2,674.80	1,714.97
32	65.38	120.89	57	754.51	762.74	82	2,729.16	1,731.85
33	71.10	134.30	58	815.54	803.87	83	2,781.36	1,746.70
34	77.68	148.76	59	887.00	849.98	84	2,831.30	1,759.32
35	85.08	164.16	60	967.43	900.19	85	2,878.89	1,769.54
36	93.23	180.44	61	1,055.37	953.62	86	2,924.03	1,777.17
37	102.08	197.51	62	1,149.38	1,009.39	87	2,966.62	1,782.03
38	111.59	215.29	63	1,248.00	1,066.61	88	3,006.58	1,783.94
39	121.69	233.70	64	1,349.79	1,124.40	89	3,043.80	1,782.71
40	132.32	252.66	65	1,453.29	1,181.89	90	3,078.19	1,778.16
41	143.45	272.08	66	1,557.05	1,238.19	91	3,109.66	1,770.11
42	155.00	291.89	67	1,659.63	1,292.41	92	3,138.11	1,758.37
43	169.44	312.90	68	1,759.57	1,343.68	93	3,163.44	1,742.76
44	189.00	335.92	69	1,855.41	1,391.12	94	3,185.56	1,723.10

Table A.1: Graduated incidence rates of cancer used in Approach 2 and Approach 3

Transition Intensities under Approach 2 (per 100,000)

	Males		Females			Males		Females	
Age	**Stroke**	**HA**	**Stroke**	**HA**	**Age**	**Stroke**	**HA**	**Stroke**	**HA**
20	713.61	643.87	697.60	157.98	58	155.49	261.75	125.98	85.45
21	-95.26	-97.78	-95.01	-22.99	59	165.50	271.88	135.22	91.52
22	-82.72	-85.69	-84.93	-20.79	60	175.22	279.67	144.88	97.40
23	-77.69	-80.12	-75.24	-18.66	61	184.52	284.92	155.01	103.10
24	-68.47	-70.86	-65.91	-16.60	62	193.36	287.51	165.64	108.60
25	-59.70	-62.00	-56.97	-14.59	63	201.70	287.35	176.80	113.91
26	-51.31	-53.49	-48.39	-12.65	64	223.12	299.97	195.76	124.17
27	-43.31	-45.34	-40.17	-10.77	65	257.21	325.92	222.71	139.85
28	-35.68	-37.52	-32.32	-8.95	66	290.58	350.60	250.56	156.12
29	-28.41	-30.02	-24.82	-7.19	67	323.38	374.25	279.38	173.01
30	-21.49	-22.83	-17.68	-5.48	68	355.26	396.39	309.22	190.56
31	-14.91	-15.95	-10.89	-3.84	69	386.25	417.15	340.13	208.79
32	-8.65	-9.35	-4.44	-2.24	70	416.38	436.60	372.13	227.72
33	-2.72	-3.04	1.66	-0.70	71	445.69	454.89	405.20	247.33
34	2.90	2.99	7.42	0.79	72	474.24	472.17	439.28	267.61
35	8.21	8.75	12.85	2.23	73	502.11	488.64	474.31	288.49
36	13.22	14.24	17.95	3.62	74	529.42	504.53	510.14	309.91
37	17.92	19.46	22.72	4.96	75	556.05	519.77	546.63	331.76
38	22.32	24.43	27.16	6.26	76	581.73	534.18	583.53	353.90
39	26.42	29.14	31.28	7.51	77	606.37	547.79	620.55	376.16
40	30.23	33.59	35.07	8.72	78	629.88	560.66	657.32	398.30
41	33.74	37.79	38.54	9.88	79	673.83	581.26	724.42	431.80
42	36.95	41.74	41.68	11.01	80	739.18	612.40	820.77	476.48
43	39.88	45.43	44.51	12.09	81	805.41	648.20	914.22	520.37
44	42.50	48.88	47.01	13.13	82	873.00	689.51	1,003.73	562.93
45	44.83	52.08	49.19	14.14	83	941.56	736.41	1,088.10	603.54
46	46.84	55.02	51.05	15.11	84	1,011.24	789.57	1,166.06	641.52
47	48.53	57.68	52.59	16.04	85	1,082.08	849.70	1,236.33	676.20
48	49.89	60.08	53.80	16.94	86	1,154.14	917.60	1,297.83	706.97
49	56.43	76.29	57.53	21.48	87	1,227.41	994.19	1,349.82	733.37
50	68.18	105.06	64.11	29.50	88	1,301.87	1,080.47	1,392.20	755.22
51	79.75	131.79	70.90	37.28	89	1,377.43	1,177.55	1,425.72	772.73
52	91.19	156.55	77.92	44.81	90	1,453.89	1,286.70	1,452.30	786.66
53	102.41	179.25	85.19	52.11	91	1,531.01	1,409.30	1,475.17	798.35
54	113.43	199.92	92.72	59.20	92	1,608.40	1,546.90	1,498.94	809.83
55	124.25	218.52	100.54	66.07	93	1,685.57	1,701.21	1,529.48	823.67
56	134.87	235.05	108.68	72.73	94	1,761.86	1,874.16	1,573.51	842.86
57	145.28	249.47	117.15	79.19					

Table A.2: Transition intensities from graduated prevalence rates under shape-preserving interpolant (HA for Heart Attack)

Transition Intensities under Approach 3 (per 100,000)

	Males		Females			Males		Females	
Age	**Stroke**	**HA**	**Stroke**	**HA**	**Age**	**Stroke**	**HA**	**Stroke**	**HA**
20	15.94	12.73	18.81	3.36	58	160.51	254.78	122.40	79.38
21	13.78	9.78	16.16	2.63	59	177.28	273.70	133.55	87.16
22	12.00	7.34	13.99	2.02	60	196.91	293.67	146.56	95.89
23	10.57	5.38	12.27	1.53	61	219.18	314.66	161.34	105.53
24	9.48	3.88	10.98	1.16	62	243.84	336.64	177.80	116.05
25	8.72	2.84	10.10	0.90	63	270.68	359.59	195.88	127.39
26	8.28	2.22	9.61	0.74	64	299.47	383.47	215.49	139.52
27	8.13	2.02	9.47	0.69	65	329.97	408.27	236.55	152.39
28	8.27	2.22	9.68	0.74	66	361.95	433.95	258.99	165.98
29	8.69	2.79	10.21	0.89	67	395.20	460.48	282.71	180.23
30	9.35	3.72	11.03	1.12	68	429.47	487.85	307.65	195.10
31	10.26	4.99	12.12	1.45	69	464.54	516.02	333.73	210.57
32	11.39	6.59	13.46	1.86	70	500.19	544.97	360.86	226.57
33	12.74	8.49	15.03	2.34	71	536.17	574.66	388.96	243.08
34	14.28	10.68	16.81	2.91	72	572.27	605.08	417.96	260.06
35	16.00	13.14	18.77	3.54	73	609.43	636.69	449.42	278.12
36	17.90	15.85	20.89	4.24	74	648.69	669.91	484.86	297.85
37	19.94	18.80	23.14	5.01	75	689.98	704.70	524.07	319.20
38	22.13	21.96	25.52	5.83	76	733.23	741.00	566.86	342.07
39	24.44	25.32	27.98	6.72	77	778.35	778.75	613.02	366.40
40	26.86	28.86	30.51	7.65	78	825.26	817.90	662.36	392.10
41	29.38	32.56	33.10	8.63	79	873.90	858.38	714.69	419.11
42	31.97	36.41	35.70	9.66	80	924.18	900.15	769.81	447.35
43	34.98	41.30	38.45	10.98	81	976.03	943.14	827.52	476.74
44	38.72	48.03	41.49	12.83	82	1,029.37	987.30	887.62	507.21
45	43.18	56.47	44.82	15.19	83	1,084.12	1,032.58	949.92	538.68
46	48.35	66.46	48.47	18.03	84	1,140.21	1,078.91	1,014.22	571.08
47	54.20	77.86	52.44	21.32	85	1,197.55	1,126.24	1,080.33	604.33
48	60.73	90.52	56.76	25.02	86	1,256.09	1,174.52	1,148.04	638.35
49	67.90	104.29	61.42	29.12	87	1,315.72	1,223.68	1,217.16	673.08
50	75.71	119.03	66.46	33.58	88	1,376.39	1,273.68	1,287.50	708.44
51	84.15	134.59	71.88	38.37	89	1,438.00	1,324.44	1,358.85	744.34
52	93.18	150.82	77.69	43.47	90	1,500.50	1,375.93	1,431.02	780.72
53	102.81	167.58	83.91	48.85	91	1,563.79	1,428.07	1,503.81	817.50
54	113.00	184.72	90.56	54.48	92	1,627.80	1,480.83	1,577.03	854.60
55	123.75	202.08	97.64	60.33	93	1,692.45	1,534.12	1,650.48	891.96
56	135.03	219.53	105.17	66.38	94	1,757.68	1,587.91	1,723.96	929.49
57	146.83	236.92	113.17	72.59					

Table A.3: Graduated transition intensities under shape-preserving interpolant (HA for Heart Attack)