Medical treatment decisions, such as glycemic control for type 2 diabetes patients, involve complex tradeoffs between the risks and benefits of treatment. The diversity of treatment options that patients can choose over time and uncertainty in future health outcomes results in a difficult sequential decision making problem. In this dissertation, we focus on developing quantitative models and methods for medical decision making, and we apply these models in the context of treatment decisions of glycemic control for patients with type 2 diabetes. We begin by reviewing the background of type 2 diabetes with the focus on treatment options for glycemic control. We also provide a review of the Markov decision process (MDP) and robust optimization literature which is relevant to the methodological aspect of this dissertation. We present a glycosylated hemoglobin (HbA1c) Markov model, and use this model to evaluate and compare effectiveness and cost of treatment regimens for new and old hyperglycemia lowering medications for individuals newly diagnosed with type 2 diabetes. Then, we present an MDP, based on the HbA1c Markov model, to optimize the sequence and time to initiate cost-effective medications with the objective of maximizing a patient’s quality-adjusted life-years prior to the first adverse event defined as a micro- and/or macro-vascular complication (e.g., heart attack, stroke, kidney failure) or death from any cause. The MDP optimally trades off the potential benefits from reducing HbA1c with the disutility associated with side effects of taking hyperglycemia lowering medications. We also analyze the impact of hypoglycemia on the MDP based optimal policy, and compare the MDP based optimal policy with published treatment guidelines. Finally, we extend the MDP to a robust MDP treatment model (RMDP-TM), which considers uncertainty in transition probability matrices. The RMDP-TM with a general uncertainty set is very difficult to solve. To deal with this, we relax the RMDP-TM and incorporate a new uncertainty set model for which we show that the relaxed RMDP-TM can be solved exactly. We also provide a
fast algorithm to solve the RMDP-TM approximately. We present theoretical analysis related to properties of the RMDP-TM which can be used to achieve computational efficiency of solving the RMDP-TM. We conclude by summarizing the most important conclusions that can be drawn from this dissertation.
Robust Optimal Control for Medical Treatment Decisions

by
Yuanhui Zhang

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Operations Research

Raleigh, North Carolina

2014

APPROVED BY:

__________________________  ____________________________
James R. Wilson            Yunan Liu

__________________________  ____________________________
Julie S. Ivy               Jennifer E. Mason

__________________________
Brian T. Denton
Chair of Advisory Committee
DEDICATION

This dissertation is dedicated to my family for their endless love and support:

Mingmin Lai, my dear Mom
Zhao Zhang, my dear Dad
Yu Du, my beloved husband
Derek Zhang Du, my lovely son
BIOGRAPHY

Yuanhui Zhang was born on January 5, 1983, in Hangzhou, Zhejiang, China. In 2005, Yuanhui received her Bachelor of Science in Mathematics and Applied Mathematics from Zhejiang University in Hangzhou, Zhejiang, China. She then received her Master of Operations Research from North Carolina State University in 2010.
ACKNOWLEDGEMENTS

First, I would like to thank my advisor Dr. Brian Denton for his guidance, support and encouragement over the last 4 years. I appreciated all he has done for me to help build my research skill and finish my dissertation. I would like to acknowledge the funding that supported this research from the National Science Foundation under Grant Number CMMI-0969885(Denton).

I would like to thank my committee members Dr. James Wilson, Dr. Julie Ivy, Dr. Yunan Liu, and Dr. Jennifer Mason for serving on my committee and providing me with useful suggestions and edits for this dissertation. In addition, a special thanks to Dr. Nilay Shah and Dr. Steven Smith for their help and guidance to help me understand medical problems over the past 4 years.

Finally, I would like to thank all of my friends for their support throughout graduate school. Thank you to Dr. Yu Du, my beloved husband, for his endless love, support, and encouragement. Thank you to my parents for loving me, supporting me, and believing in me.
# TABLE OF CONTENTS

**LIST OF TABLES** ................................................. vii

**LIST OF FIGURES** ................................................ ix

**Chapter 1 Introduction** ........................................ 1

**Chapter 2 Literature Review** .................................... 5
  2.1 Background on Glycemic Control for Type 2 diabetes .......... 5
  2.2 Diabetes Models .............................................. 9
  2.3 Markov Decision Process ...................................... 12
  2.4 MDPs Applied to Medical Decision Making ....................... 13
  2.5 Robust Optimization ......................................... 16
    2.5.1 General RO Formulation ................................. 17
    2.5.2 Uncertainty Set Models ................................. 18
    2.5.3 Robust MDP .............................................. 21
  2.6 Contributions of this Dissertation ............................ 24

**Chapter 3 Evaluation of Glycemic Control for Type 2 Diabetes** .... 26
  3.1 Introduction .................................................. 26
  3.2 Markov Chain .................................................. 28
    3.2.1 Treatment Regimens ...................................... 28
    3.2.2 Markov Model ............................................. 30
    3.2.3 Outcome Measures ........................................ 31
    3.2.4 Data Sources ............................................. 32
    3.2.5 Model Parameters for Base Case and Sensitivity Analysis .... 33
    3.2.6 Model Calibration and Validation ......................... 36
  3.3 Results ........................................................ 37
    3.3.1 Base Case Results ....................................... 37
    3.3.2 Sensitivity Analyses ..................................... 38
  3.4 Conclusions ................................................... 41

**Chapter 4 Optimization of Glycemic Control for Type 2 Diabetes** .... 46
  4.1 Introduction ................................................... 46
  4.2 Model Formulation ............................................ 47
  4.3 Model Properties and Insights ................................ 52
    4.3.1 Model Properties ........................................ 52
  4.4 Results ........................................................ 57
    4.4.1 Data Sources and Parameter Estimation ................... 57
    4.4.2 Base Case Analysis ....................................... 59
    4.4.3 Impact of Hypoglycemia on Optimal Policy ................ 60
    4.4.4 Comparison of the optimal policy and HbA1c threshold treatment policies ... 62
  4.5 Conclusions .................................................... 63
Chapter 5 Robust Optimization of Medical Treatment Decisions ............... 66
5.1 An RMDP Formulation for Medical Treatment Decisions ................. 67
5.2 A Polyhedral Uncertainty Set ............................................. 70
5.3 Bi-criteria RMDP-TM ......................................................... 74
5.4 Analysis of the RMDP-TM ..................................................... 76
5.5 Numerical Experiments ......................................................... 94
   5.5.1 Uncertainty Set Estimation ........................................... 95
   5.5.2 Base Case Analysis .................................................... 96
   5.5.3 Algorithm Performance Comparison ................................ 102
   5.5.4 RMDP-optimal Policy Performance Comparison ................. 104
   5.5.5 Sensitivity of Worst-Case Performance to Model Parameters ... 111
5.6 Conclusions ........................................................................... 114

Chapter 6 Conclusions ..................................................................... 116

References .................................................................................... 120

APPENDICES .................................................................................. 131
   A Risk Equations of Microvascular and Macrovascular complications for Patients with Type 2 Diabetes ......................................................... 132
   B CDC Mortality Table ............................................................ 133
   C Parameters for Chapter 5 ...................................................... 134
LIST OF TABLES

Table 2.1  Properties of frequently used hyperglycemia lowering medications [1]. . . . 8
Table 2.2  Summary of the HbA1c threshold of published treatment guidelines for treating type 2 diabetes. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
Table 2.3  Computational Complexity of the Robust Counterpart of Some Classical Optimization Problems . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

Table 3.1  Model parameters for the base case analysis and sensitivity analysis . . . 29
Table 3.2  Glycosylated hemoglobin (HbA1c) used in the Markov model for women. HbA1c range definition at diagnosis, the mean natural HbA1c values for each HbA1c state at diagnosis (prior to initiating medication), the initial HbA1c distributions at diagnosis, and 3-month HbA1c transition probability matrices for women. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
Table 3.3  Glycosylated hemoglobin (HbA1c) used in the Markov model for men. HbA1c range definition at diagnosis, the mean natural HbA1c values for each HbA1c state at diagnosis (prior to initiating medication), the initial HbA1c distributions at diagnosis, and 3-month HbA1c transition probability matrices for men. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32
Table 3.4  Base case comparison of four treatment regimens. Comparison of the total expected LYs, and total expected QALYs from birth to occurrence of first diabetes-related complication or death, expected medication cost per QALY and the ICER from diagnosis to occurrence of first diabetes-related complication or death, mean time from diagnosis to insulin initiation for men and women. Four treatment regimens are T1, metformin plus sulfonylurea plus insulin; T2, metformin plus DPP-4 inhibitor plus insulin; T3, metformin plus GLP-1 agonist plus insulin; and T4, metformin plus insulin. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38

Table 4.1  Summary of sources for model inputs. . . . . . . . . . . . . . . . . . . . . . 59
Table 4.2  Comparison of the expected total QALYs from birth to the first event where patients are assumed to have one QALY per year before diagnosis. Results were generated by evaluating the optimal policy, the observed average policy, and treatment guidelines with 6 different treatment goals under the base case model where hypoglycemia is considered as a transient symptom, and the alternative model where the hypoglycemia is modeled as an event. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63

Table 5.1  Comparison of the expected worst-case QALYs prior to first event for glycemic control case study under base case. Upper bounds are obtained by solving RMDP-TMs approximately using Algorithm 4, and lower bounds are obtained by solving the relaxed RMDP-TMs exactly using Algorithm 2. 97
Table 5.2  Comparison of the computation time (in seconds) of solving the NMDP and the relaxed RMDP-TM exactly, and solving the RMDP-TM approximately.103
Table 5.3  Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for men in base case when hypoglycemia is considered as a transient symptom. ................................................................. 104

Table 5.4  Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for women in base case when hypoglycemia is considered as a transient symptom. ................................................................. 106

Table 5.5  Comparison of the expected worst-case QALYs prior to first event for glycemic control case study when hypoglycemia is considered as an event. Upper bounds are obtained by solving RMDP-TMs approximately using Algorithm 4, and lower bounds are obtained by solving the relaxed RMDP-TMs exactly using Algorithm 2. ................................................................. 108

Table 5.6  Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for men when hypoglycemia is considered as an event. .......... 109

Table 5.7  Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for women when hypoglycemia is considered as an event. .......... 109

Table 1  Parameters and β coefficients for risk equations to estimate the probability of diabetes-related complications. ................................................................. 133

Table 2  The estimated 3-month all cause mortality table .................................................. 133

Table 3  Glycosylated hemoglobin (HbA1c) used in the RMDP-TM for women. HbA1c range definition at diagnosis, the mean natural HbA1c values for each HbA1c state at diagnosis (prior to initiating medication), the initial HbA1c distributions at diagnosis, and 3-month HbA1c transition probability matrices for women. ................................................................. 134

Table 4  Test TPM for IFR property of the MLE of the TPM for women. ......................... 134

Table 5  Left-hand-side maximum deviation of the TPM in RMDP-TM for women. ........ 135

Table 6  Right-hand-side maximum deviation of the TPM in RMDP-TM for women. ...... 135

Table 7  NIWC TPM for women. ......................................................................................... 136

Table 8  Test TPM for IFR property of the NIWC TPM for women. ............................... 136

Table 9  Glycosylated hemoglobin (HbA1c) used in the RMDP-TM for men. HbA1c range definition at diagnosis, the mean natural HbA1c values for each HbA1c state at diagnosis (prior to initiating medication), the initial HbA1c distributions at diagnosis, and 3-month HbA1c transition probability matrices for men. ................................................................. 137

Table 10 Test TPM for IFR property of the MLE of the TPM for men. ......................... 137

Table 11 Left-hand-side maximum deviation of the TPM in RMDP-TM for men. .......... 138

Table 12 Right-hand-side maximum deviation of the TPM in RMDP-TM for men. ...... 138

Table 13 NIWC TPM for men. ......................................................................................... 139

Table 14 Test TPM for IFR property of the NIWC TPM for men. ............................... 139
LIST OF FIGURES

Figure 2.1 A patient-centered approach: position statement of the ADA and the European association for the study of diabetes (EASD) from [2]. ......... 10

Figure 3.1 The HbA1c threshold treatment policy based on the ADA’s consensus algorithm for treating hyperglycemia for type 2 diabetes [3]. ............... 30

Figure 3.2 QALYs versus cost incurred by the four different treatment regimens as a function of glycemic control goal. Comparison of the total expected QALYs versus the expected medication cost per QALY incurred from diagnosis to first event (diabetes-related complication or death) for men (5.10a) and women (5.10b). Each of the four treatments is compared as the glycemic control goal is varied from 6.5% (48 mmol/mol) to 8% (64 mmol/mol). Results are presented using HbA1c of 6.5% (48 mmol/mol) (circle), 7% (53 mmol/mol) (triangle), and 8% (64 mmol/mol) (square) as the glycemic control goal. ..................... 39

Figure 3.3 Sensitivity analysis on the medication cost. The x-axis represents the difference in the expected medication cost per QALY from the base-case cost: metformin costs 81.75 USD per month, sulfonylurea costs 54.85 USD per month, DPP-4 inhibitor costs 232.84 USD per month, GLP-1 agonist costs 325.97 USD per month, and insulin therapy costs 245.70 USD per month. The y-axis represents the treatment regimen. The solid bar represents men, and the hatched bar represents women. met, metformin; sulf, sulfonylurea. ..................... 40

Figure 4.1 A flow diagram illustrating the entire treatment decision process. ....... 47

Figure 4.2 State transition diagram illustrating transitions among the HbA1c states, the absorbing states, and the medication states. The dashed box represents the set of absorbing states. The solid lines represent the HbA1c state transitions. ..................... 50

Figure 4.3 The first 6 years optimal policy from diagnosis for patients who are not on insulin, including patients not on any medications, patients on metformin only, patient’s on metformin and sulfonylurea, when hypoglycemia is modeled as a transient symptom (base case). The following medication states are never optimal in the figure: no medication, metformin only, and sulfonylurea only. ..................... 60

Figure 4.4 The first 6 years optimal policy from diagnosis for patients who are not on insulin, including patients not on any medications, patients on metformin only, patient’s on metformin and sulfonylurea, when hypoglycemia is modeled as an event. ..................... 61
Figure 5.1 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently not on any medications when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin. 

Figure 5.2 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently on metformin only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only. 

Figure 5.3 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently on sulfonylurea only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin. 

Figure 5.4 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently on metformin and sulfonylurea together when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only. 

Figure 5.5 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently not on any medications. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin. 

Figure 5.6 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently on metformin only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only. 

Figure 5.7 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently on sulfonylurea only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin. 

Figure 5.8 A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently on metformin and sulfonylurea together when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only. 

Figure 5.9 The NMDP-optimal and RMDP-optimal policies performance comparisons based on the sampling approach and the worst-case approach when hypoglycemia is modeled as a transient symptom.
Figure 5.10 The NMDP-optimal and RMDP-optimal policies performance comparisons based on the sampling approach and the worst-case approach when hypoglycemia is modeled as an event. 110

Figure 5.11 The difference in the expected worst-case QALYs prior to the first event obtained from the NMDP-optimal policy and the RMDP-optimal policy without the uncertainty budget constraint. The adjustment factor is a parameter ranging from 0 to 1. When the adjustment factor equals zero, the model is equivalent to the base case where hypoglycemia is modeled as a transient symptom, and when the adjustment factor equals 1, the model is equivalent to the alternative case when hypoglycemia is modeled as an event. 113
Chapter 1

Introduction

Medical treatment decisions involve complex tradeoffs between the risks and benefits of treatment. These decisions are often made by physicians based on results from randomized control trials (RCTs) and/or observational studies. However, decision approaches are rarely codified or quantified in a way that makes clear the short term costs of medication and harms from side effects versus the long term benefits of avoiding disease complications. For complex diseases, like diabetes, there are many treatment options that can potentially be selected over the course of a patient’s lifetime. This results in a difficult sequential decision making process which seek to optimally trade off short term harms with uncertain long term benefits.

Glycemic control for patients with type 2 diabetes involves the regulation of blood glucose levels over time. It is a good example of a complex medical decision making problem involving multiple medications that may be used over a patient’s lifetime. Glycemic control aims to avoid acute daily symptoms of hyperglycemia, to avoid instability in blood glucose over time, and to prevent or delay the development of diabetes complications associated with the high blood glucose level. Glycosylated hemoglobin (HbA1c), as a percentage of total haemoglobin, is commonly used as a measure of average blood glucose concentration over time. A high HbA1c indicates poor glycemic control. Therefore, glycemic control is also referred to as HbA1 control. HbA1c control is difficult for several reasons. First, the medication selection is difficult since
there are many hyperglycemia lowering medications with varying effects on HbA1c, and with various side effects. Second, there are uncertainties in future HbA1c progression, effects of medications, and risks of diabetes-related complications. Third, current treatment guidelines are not consistent on the HbA1c threshold for treating type 2 diabetes patients [4, 5, 6]. In light of recently discovered uncertainties regarding the benefits of intensive glycemic control that were found in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) Trial [7], the best HbA1c threshold for treating diabetes patients is not clear.

In this dissertation we formulate and analyze mathematical models, first to evaluate and compare different treatment guidelines for HbA1c control, and next to optimize treatment decisions with regard to the sequence and time to initiate medications. In addition to developing new models for treatment decisions, we also analyze the structure of these models and develop computationally efficient methods for solving these models.

This dissertation is structured as follows. First, in Chapter 2 we provide a literature review of glycemic control for type 2 diabetes, and relevant methods for optimal control including Markov decision processes (MDPs) and robust MDPs (RMDPs). For the MDP and the RMDP, we discuss some important theoretical properties and solution algorithms. Furthermore, we summarize their applications in the medical decision making context.

In Chapter 3, we present a new population-based HbA1c control model based on a Markov chain. The Markov state of the model is a patient’s HbA1c level. Treatment decisions are revisited every three months. The model is calibrated with an retrospective administrative claim data set. The eligible population for our study consisted of 37,501 patients who were diagnosed with type 2 diabetes from 1995 to 2010, and were 40 years and older with at least 5 years of continuous enrollment with at least two HbA1c records, and had complete pharmacy claim data. This model aims to answer the following question: “After initiating metformin, which is generally accepted as the first-line medication in treating type 2 diabetes, which medication should be the best second-line medication?” The recent development of several new and competing medications makes this an important research question. We compare four hyperglycemia lowering medications:
sulfonylurea, dipeptidyl peptidase-4 inhibitor, glucagon-like peptide-1 receptor agonist, and insulin. We consider both the real-world effectiveness of medications from the administrative claim data set, and RCT results from literature. Outcome measures include life-years (LYs), and quality-adjusted life-years (QALYs) to the first micro- or macro-vascular complication (ischemic heart disease, myocardial infarction, congestive heart failure, stroke, blindness, renal failure, amputation) or death from any causes, mean time to insulin dependence, expected medication cost per QALY, and incremental cost-effectiveness ratio (ICER) from diagnosis to the first micro- or macro-vascular complication. We find that use of sulfonylurea as second-line medication results in similar LYs and QALYs as compared to those associated with other agents but at lower cost.

In Chapter 4, we present an MDP model, based on the Markov chain developed in Chapter 3, to determine the optimal policy (and thus the optimal sequence and time to initiate the hyperglycemia lowering medications) for HbA1c control. The goal of this MDP is to maximize a patient’s expected QALYs prior to the first adverse event defined as a micro- and/or macro-vascular complication or death from any cause. This model aims to answer the following research questions: 1) when and in which order should hyperglycemia lowering medications be prescribed? 2) and what are the benefits of using the optimal policy vs. a recommended treatment guideline based on a specified HbA1c treatment threshold? We evaluate the effect of the risk of hypoglycemia on the optimal sequence to initiate medications. A series of experiments are performed to compare the expected outcomes for the optimal policy to the outcomes for the current American Diabetes Association (ADA) treatment goal. We also perform experiments to investigate the influence of hypoglycemia risk on the optimal sequence and time to initiate medications.

In Chapter 5, we extend the MDP presented in Chapter 4, which we refer to in this chapter as the nominal MDP (NMDP), to a robust MDP treatment model (RMDP-TM), which considers uncertainty in transition probability matrices (TPMs). The term nominal refers to the fact that the model parameters are not subject to uncertainty. The RMDP-TM, on the other hand,
assumes TPM may vary within a defined uncertainty set. The purpose of RMDP-TM is to protect the performance of the optimal solution against the parameter ambiguity. It is particularly valuable for our HbA1c control problem since our model parameters are estimated from “noisy” observational data. As a result, estimates are subject to statistical variation which might degrade the performance of the optimal policy.

The RMDP-TM with a general uncertainty set is difficult to solve. Therefore, we relax the RMDP-TM and incorporate a new uncertainty set model named the interval model with uncertainty budget (IMUB) for which we show that the relaxed RMDP-TM can be solved exactly using the robust dynamic programming algorithm. We provide a fast algorithm to solve the RMDP-TM approximately. We conduct theoretical analysis to derive structural properties of the relaxed RMDP-TM, and prove sufficient conditions for the optimality gap between the RMDP-TM and its relaxation to be zero. Numerical experiments show that for the glycemic control problem, the difference between the NMDP-optimal policy and RMDP-optimal policies are significantly affected by the methods of modeling the risk of hypoglycemia. The magnitude of applying the RMDP-optimal policy over the NMDP-optimal policy in the worst case is comparable to some well-know prevention programs.

Finally, in Chapter 6, we summarize the most significant findings from the work presented in this thesis, and discuss opportunities for future research.
Chapter 2

Literature Review

In this chapter, we review the literature that is most related to our study. In Section 2.1 we provide some background on type 2 diabetes. In Section 2.2 we present several mathematical models that are related to glycemic control for type 2 diabetes patients. In Section 2.3 we give a brief introduction to the MDP models, and solution techniques. In Section 2.4 we summarize the previous studies that applied MDP models to medical decision making problems. In Section 2.5 we provide a brief review on robust optimization and their applications. Finally, in Section 2.6, we highlight the main contributions of this proposal.

2.1 Background on Glycemic Control for Type 2 diabetes

Diabetes is a chronic disease in which a person has high blood sugar, either because the body does not produce enough insulin, or because cells do not respond to the insulin that is produced. Type 2 diabetes is characterized by insulin resistance, a condition in which the natural hormone insulin becomes less effective at lowering blood sugars. Insulin resistance may also occur in combining with reduced insulin secretion.

There are seven classes of hyperglycemia lowering medications that are commonly used in the United States: biguanides, thiazolidinediones (TZDs), sulfonylureas, meglitinides, dipep-
tidyl peptidase-4 (DPP-4) inhibitors, glucagon-like peptide-1 (GLP-1) receptor agonists, and insulin. The first six noninsulin medications can be categorized into 3 types: *insulin sensitizers*, *secretagogues* and *peptide analogs*.

Insulin sensitizers include biguanides, such as the commonly prescribed drug metformin, and TZDs. Biguanides are the most widely used oral medications. These medications lower blood glucose level by reducing the amount of glucose the liver produces and by making muscle tissue more sensitive to insulin, which increases the absorption of glucose by cells. These medications usually do not cause hypoglycemia when used alone and may help with weight loss. Because of their benefits and typical absence of significant side effects, these medications are commonly adopted as the first-line medication for patients newly diagnosed with type 2 diabetes. TZDs are also effective at reducing HbA1c by improving the function of insulin in muscle and fat, and decreasing the production of glucose in the liver. However, on May 21, 2007, the Food and Drug Administration (FDA) issued a safety alert regarding the risk of cardiovascular and other health complications when taking rosiglitazone, one of the agents in this drug class [8].

Sulfonylureas and meglitinides are both secretagogues. The medications stimulate the beta cells, a type of cell located in the pancreas, to release more insulin. Because of this mechanism, these two drug classes may cause episodes of low blood glucose levels, known as hypoglycemia. Sulfonylureas have been used for glycemic control since the 1950s while meglitinides have been approved by the FDA for treating type 2 diabetes since 1997. Meglitinides have a rapid onset and short duration of action as compared to sulfonylureas, so they are particular effective in increasing insulin levels immediately after a meal.

DPP-4 inhibitors and GLP-1 receptor agonists are two classes of hyperglycemia lowering medications that enhance insulin production by increasing the half-life of the active form of GLP-1. GLP-1 is a incretin hormone that increases insulin production when blood sugar is high. However, GLP-1 is active for a very short time (approximately 2 min) in the blood since it is broken down quickly by an enzyme called DPP-4. DPP-4 inhibitors are a new class of oral medication that prolong the half-life of GLP-1 by inhibiting the action of DPP-4. GLP-1
receptor agonists are a new class of injectable drugs which acts as a GLP-1 but are not broken as quickly as GLP-1 [9].

Insulin is the oldest and most commonly used injectable drug for treating diabetes. There are more than 20 types of insulin sold in the United States. These insulins differ in how they are made, how they work in the body, and how much they cost. There are human and analogue insulins, and in general human insulins are cheaper than analogs. The regimen of taking insulin is complex. Most patients need at least 2 insulin shots a day while some need 3 or 4 shots. Each type of insulin works at a different speed and lasts for a different length of time. In general, rapid-acting insulin analogs begin to work within 5 to 15 minutes, and last for 3 to 4 hours. Short-acting regular insulin starts working within 30 minutes, and lasts about 5 to 8 hours. Intermediate-acting insulin starts working in 1 to 3 hours, and lasts 12 to 16 hours. Long-acting insulin is delivered at a steady level and lasts 24 to 28 hours.

Medications vary in their ability to reduce HbA1c and their side effects. Table 2.1 summarizes the properties of the frequently used hyperglycemia lowering medications including the mechanism of action, advantages, side effects and the medication effect in terms of absolute reduction in HbA1c. Notice that the unit of HbA1c level is percentage, thus the medication effect shown in Table 2.1 is actually the absolute change in HbA1c level.

Some medications may be used together, such as metformin and sulfonylurea combination therapy, TZD and metformin combination therapy, and others. Some medications are not commonly used together either because they use the same mechanism of action, such as sulfonylurea and insulin, or because they are not approved to be used together by the FDA, such as GLP-1 receptor agonist and insulin.

Hypoglycemia is the most common adverse effect of diabetes treatment. It is an important limiting factor in the glycemic management for patients with diabetes, and it is a significant barrier to adherence to medications. Metformin alone is rarely associated with hypoglycemia [10]. The reported severity of sulfonylurea-induced and insulin-induced hypoglycemia range widely across studies [1]. Over 6 years of follow-up of patients with type 2 diabetes in the United
Table 2.1: Properties of frequently used hyperglycemia lowering medications [1].

<table>
<thead>
<tr>
<th>Medication</th>
<th>Mechanism of Action</th>
<th>Effect</th>
<th>Advantages</th>
<th>Side Effect</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metformin</td>
<td>Inhibits glucose production by the liver</td>
<td>1.0%–2%</td>
<td>Weight neutral, Rarely cause hypoglycemia</td>
<td>Gastrointestinal side effect</td>
<td>Low</td>
</tr>
<tr>
<td>TZD</td>
<td>Increase glucose uptake by skeletal muscle</td>
<td>1.0%–1.5%</td>
<td>Rarely cause hypoglycemia</td>
<td>Weight gain, Increase the risk of heart failure</td>
<td>High</td>
</tr>
<tr>
<td>Sulfonulurea</td>
<td>Increase insulin secretion</td>
<td>1.0%–2.0%</td>
<td>Rapidly effective</td>
<td>Weight gain, Hypoglycemia</td>
<td>Low</td>
</tr>
<tr>
<td>DPP-4 inhibitor</td>
<td>Increase GLP-1 activity</td>
<td>0.5%–0.8%</td>
<td>Weight neutral, Well tolerated</td>
<td>Urinary tract infection, Angioedema</td>
<td>High</td>
</tr>
<tr>
<td>GLP-1 receptor agonist</td>
<td>Increase insulin secretion, Decrease glucagon secretion</td>
<td>0.5%–1.0%</td>
<td>Weight loss, Rarely cause hypoglycemia</td>
<td>Gastrointestinal side effect, Inject medication</td>
<td>High</td>
</tr>
<tr>
<td>Insulin</td>
<td>Increase glucose disposal, Decrease hepatic glucose production</td>
<td>Theoretically unlimited efficacy</td>
<td>Universally effective</td>
<td>Weight gain, Inject medication, Hypoglycemia</td>
<td>Depends</td>
</tr>
</tbody>
</table>

Kingdom Prospective Diabetes Study (UKPDS), 0.3% (95% CI: 0.1% to 1.1%) of those using metformin, 1.2% (95% CI: 0.4% to 3.3%) of those using a sulfonylurea, and 3.8% (95% CI: 1.2% to 11.1%) of those using insulin reported hypoglycemia [11].

According to the clinical recommendations of the ADA, a reasonable HbA1c goal for many non-pregnant adults is HbA1c < 7%. A more stringent HbA1c goal (6.0–6.5%) is suggested if an individual can achieve it without significant hypoglycemia events or other adverse effects of treatment. Conversely, a less stringent HbA1c goal (7.5–8.0%) is recommended if the patient has a history of severe hypoglycemia events, limited life expectancy, advanced microvascular or macrovascular complications, extensive comorbid conditions, or longstanding diabetes. A summary of HbA1c goals according to different treatment guidelines is presented in Table 2.2.

The ADA recommends that HbA1c tests should be performed at least twice a year in patients that meet the treatment goal, and have stable glycemic control. For patients whose therapy
Table 2.2: Summary of the HbA1c threshold of published treatment guidelines for treating type 2 diabetes.

<table>
<thead>
<tr>
<th>Treatment Guideline</th>
<th>HbA1c Threshold</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>International Diabetes Federation 2009</td>
<td>$\leq 6.5%$</td>
<td>[12]</td>
</tr>
<tr>
<td>American Association of Clinical Endocrinologists 2013</td>
<td>$\leq 6.5%$</td>
<td>[4]</td>
</tr>
<tr>
<td>American Diabetes Association 2013</td>
<td>$&lt; 7%$</td>
<td>[5]</td>
</tr>
<tr>
<td>Institute for Clinical Systems Improvement 2013</td>
<td>$\leq 8%$</td>
<td>[6]</td>
</tr>
<tr>
<td>Veterans Affairs Clinical Practice Guideline 2010</td>
<td>$&lt; 9%$</td>
<td>[13]</td>
</tr>
</tbody>
</table>

has changed or who are not meeting glycemic goals [5], the HbA1c test should be performed quarterly. Figure 2.1 shows the ADA and the European association for the study of diabetes (EASD) suggested implementation strategies for glycemic control [2].

## 2.2 Diabetes Models

Several mathematical models have been developed to help validate, estimate the benefits, and improve medical decision making for type 2 diabetes management. The types of models include systems of differential equations, simulation models, and Markov chains.

Schlessinger et al. [14] report on the “Archimedes” model which uses a system of differential equations to represent normal physiology and diseases and disorders related to diabetes and its complications. They simulated 74 clinical trials and found that their model was able to replicate the results of 71 clinical trials with no statistical difference from the actual results. They concluded that it was possible to build a mathematical model that could replicate the pathophysiology of diabetes at a level of biological and clinical detail and that could be used to simulate clinical trials.

Simulation models have been widely used to estimate the benefits of glycemic control in preventing microvascular complications, and to evaluate the cost-effectiveness of different treatment strategies for patients with type 2 diabetes. Bagust et al. [15] developed a computer simulation model which generated dynamic risk factor trajectories under prespecified treatment
strategies: lifestyle modification, sulfonylurea, metformin, and insulin. The authors used their model as an input of existing economic model to evaluate the cost-effectiveness of treatment strategies. They found that metformin outperformed other therapies when treatment continues for 10 or more years from diagnosis.

Chen et al. [16] also used a simulation model to investigate the impact of alternative HbA1c thresholds, ranging from 6.5% to 9.0%, for treatment intensification on long-term health outcomes. In their model, HbA1c changed after initiating a new intervention with the followed three stages:
a drop in HbA1c during the first 6-month cycle, a stable HbA1c period, and an increase in HbA1c at a fixed rate thereafter. The magnitude of the HbA1c drop was randomly drawn from a normal distribution, with parameters based on data from randomized controlled trials. The rate of deterioration in HbA1c was also randomly determined from a normal distribution and applied in subsequent cycles while the patient remained on the regimen. The authors concluded that strategies that intensify treatment with a HbA1c threshold of 7.0% or less were associated with enhanced projected long-term health outcomes.

Eastman et al. [17, 18] used a Markov model with Monte Carlo sampling to predict vascular complications occurring in type 2 diabetes patients. Their model used 14 health states to model the natural history of vascular complications including retinopathy, nephropathy, neuropathy and cardiovascular disease. They found the predictions of complications and mortality are consistent with the known epidemiology of the disease, and they concluded that such a model is appropriate for evaluating the effect of preventative interventions for type 2 diabetes. Furthermore, they used their model to analyze the health benefits and economics of treating type 2 diabetes with the goal of HbA1c < 7%. The incremental effectiveness of treating type 2 diabetes with the goal of HbA1c < 7% was estimated to be $16,000/QALY gained.

Herman et al. [19, 20] used five independent Markov chains to model the progression of five complications associated with type 2 diabetes: nephropathy, neuropathy, retinopathy, coronary heart disease, and stroke. The Markov chains were used to evaluate the lifetime cost–utility (the ratio between the cost of a health–related intervention and the benefit it produces in terms of the number of years lived in full health by the beneficiaries) of the Diabetes Prevention Program (DPP) interventions including metformin and intensive lifestyle intervention. DPP was a multicenter clinical research study aimed at discovering whether modest weight loss through dietary changes and exercise, or treatment with metformin could prevent or delay the onset of type 2 diabetes. They estimated the DPP lifestyle intervention to have a cost of $8,800/QALY and metformin a cost of $29,900/QALY.
2.3 Markov Decision Process

Sequential decision making under uncertainty involves decisions in a system that evolves according to a stochastic process. The decision maker take actions to intervene the evolution of the system at a series of time points (decision epochs) in order to achieve a pre-specified goal. The key ingredients of a sequential decision model includes the decision epochs, states (the status of the system), actions (decisions), rewards (the outcome associated with the state and action), and the transition probabilities between states. An MDP is a particular type of sequential decision model in which the transition probability and reward functions depend only on the current state of the system and the action selected by the decision maker in that state. In the remainder of this section, we provide an short description of finite–horizon, discrete–time, discrete–state MDPs. The reader is referred to books by Howard [21], Puterman [22], and Bertsekas [23] for more details on different types of MDP formulations, solution techniques, as well as the methodological results.

A finite-horizon, discrete-time, discrete-state MDP can be mathematically represented by a collection of objects,

\[(T, S, A_s, P_t^a, r_t(s, a))\] (2.1)

where \(T = \{0, 1, \ldots, T - 1\}, T < \infty\) denotes the set of decision epochs that decisions are made; \(S = \{1, 2, \ldots, S\}\) is the state space which includes all possible states that the system can occupy at each decision epoch; for each state \(s \in S\), \(A_s\) is the action space which includes all possible actions that the decision maker can take at state \(s\); \(P_t^a\) is the transition probability matrix which determines the state distribution for the system in the next decision epoch; and \(r_t(s, a)\) is the reward function which is the result of taking action at each state. A decision rule, denoted by \(d_t : S \rightarrow A_s\), prescribes a procedure for action selection in each state at a specified decision epoch. A policy \(\pi = \{d_0, d_1, \ldots, d_{T-1}\}\) specifies the decision rule to be used at all decision epochs. It provides the decision maker with a prescription for action selection under any possible system state.
The objective of solving an MDP is to find the policy that maximizes the expected reward for a given optimality criterion. The most commonly used criterion in health applications is the expected total discounted reward criterion which is also used in our model. In decision epoch $t \in T$ for state $s \in S$, the optimal value function $v_t(s)$ is defined as follows:

$$v_t(s) = \max_{\pi \in \Pi} E_{\pi}^s \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, d_k(s_k)) + \lambda^{T-t} r_T(s_T) \right\}$$  \hspace{1cm} (2.2)

where $\lambda \in (0, 1]$ is the discount factor, $\Pi$ is the policy space which includes all possible policies, and $r_T(s_T)$ denotes the terminal reward that occurs at the end of the process when the system is in state $s_T$. The optimal value function (2.2) can be computed by iteratively solving the following recursive equations which are also called optimality equations or Bellman equations:

$$v_t(s) = \begin{cases} 
\max_{a \in A} \left\{ r_t(s, a) + \lambda \sum_{s_{t+1} \in S} p_t(s_{t+1} \mid s, a) v_{t+1}(s_{t+1}) \right\}, & \forall s \in S, t \neq T, \\
r_T(s), & \forall s \in S, t = T.
\end{cases}$$  \hspace{1cm} (2.3)

The optimal action, $a^*_t(s)$, at decision epoch $t$ for state $s$ is chosen based on the optimality equations:

$$a^*_t(s) = \arg\max_{a \in A} \left\{ r_t(s, a) + \lambda \sum_{s_{t+1} \in S} p_t(s_{t+1} \mid s, a) v_{t+1}(s_{t+1}) \right\}, \hspace{0.5cm} t \neq N$$  \hspace{1cm} (2.4)

The above equations can be solved to obtain the optimal value function and optimal policies using the backwark induction algorithm [22].

### 2.4 MDPs Applied to Medical Decision Making

Since the 1960s there has been a significant amount of research on MDPs with applications to production control [24, 25], finance [26], machine maintenance [27, 28, 29], and inventory control [30, 31]. Recently, there has been considerable research into the application of discrete-time,
discrete-state MDPs to medical decision making. They can be categorized into infinite-horizon stationary MDPs [32, 33, 34, 35] and finite-horizon non-stationary MDPs [36, 37, 38, 39, 40]. In this section, we provide a brief review of the most relevant literature. For a broader survey of applications the reader is referred to [41, 42].

Alagoz et al. [32] presented an infinite-horizon stationary MDP model for deciding the optimal time to perform a living-donor liver transplantation. The goal of the model was to determine when to perform the surgery in order to maximize the patient’s expected life years. The markov state in the model is patient’s MELD score, which represents the patient’s health status, and is calculated based on patient’s total bilirubin, creatinine, and prothrombin time. The model considered the daily decision of whether or not to transplant. Alagoz et al. used the policy iteration algorithm to solve the MDP and generated an optimal stationary policy to transplant, or wait at least another day, as a function of the liver quality and the patient health at the start of the day. Further, Alagoz et al. [33] extended the model to considerate both living-donor and cadaveric organs transplantation and derived structural properties of the MDP model such as the optimal control-limit policy.

Shechter et al. [34, 35] presented an infinite-horizon stationary MDP model to study the optimal timing to initiate the HIV therapy. The goal of the model was to maximize the expected life years of the patient. The markov states in the model are patient’s CD4 counts which is the number of CD4 cells (a type of white blood cell that fights infection) per mm$^3$ of blood. The model considers a monthly decision of whether or not to initiate treatment for patients. The authors also prove that if it is optimal to initiate treatment at a given CD4 count, it is also optimal to initiate treatment for patients with higher CD4 counts, i.e., that the optimal policy is of control-limit type.

Denton et al. [36] presented a finite-horizon non-stationary MDP model to study the optimal timing of statin initiation for cholesterol management for patients with type 2 diabetes. The states in the model are the patient’s total cholesterol (TC) and high-density lipoprotein (HDL) levels, and the decision was made annually. The authors compared the optimal start times of
statin treatment by using different cardiovascular risk models and found that the choice of cardiovascular risk model influenced the optimal time for starting statins. Further Mason et al. [43] extended the model to incorporate a Markov model linking adherence to treatment effectiveness. They concluded that patients who improved their adherence could increase their expected QALYs by up to 1.5 years without significantly increasing the expected cost.

Kurt et al. [37] also considered the optimal timing of statin initiation for patients with type 2 diabetes. They formulated an infinite-horizon non-stationary MDP to maximize the patient’s QALYs prior to the first CHD or stroke event. The authors derived structural properties of the MDP model including sufficient conditions for the optimality of control-limit policies with respect to the patient’s lipid ratio levels and age. They also found that the optimal policy was very sensitive to the disutility of statins.

Mason et al. [38] presented a finite-horizon non-stationary MDP model to determine the optimal coordinated treatment decisions for blood pressure and cholesterol control in patients with type 2 diabetes. The states in the model were the patient’s health status including systolic blood pressure (SBP), TC, HDL, current medications, number of stroke and CHD. The model considered annually decision of whether to initiate various blood pressure and cholesterol lowering medications. The goal of the model was to find the optimal sequence and timing of medications to simultaneously maximize quality of life and minimize cost of treatment. By comparing with the U.S. guidelines which have separate treatment recommendations for blood pressure control and cholesterol control, the author concluded that the model-based optimal coordinated treatment policy can reduce the costs of while improving the quality of life.

Chhatwal et al. [39] formulated a finite-horizon MDP to address the optimal breast biopsy decision for an individual patient. The states in the model were the patient’s risk score (the current probability of cancer). The decision in the model, which is made annually, is whether to biopsy or wait to biopsy. The goal of the model was to find the optimal breast biopsy decision which maximizes the patient expected QALYs. The authors derived structural properties of the MDP model including sufficient conditions for the optimality of control-limit policies with
respect to the patient’s risk score, and sufficient conditions to ensure the optimal control limits
does not decrease with time. They found that the optimal policy is age-dependent with older
women having a higher biopsy threshold than younger women.

Hsih et al. [40] presented a finite-horizon MDP to optimize a patient’s health condition by
achieving blood sugar levels within a pre-specified range. They focused on patients who were
either unable to take metformin (due to a contraindication) or were unable to control their blood
sugar by only using metformin. The state in the model was defined by fasting plasma glucose,
HbA1c, body mass index (BMI) and an indicator of whether current side effects were intolerable.
The actions included no treatment, initiate TZD, initiate Sulfonylurea, initiate α-glucosidase
inhibitor, and initiate GLP-1 agonist. The reward function was defined to be a combination
of piecewise linear functions each representing the health impact of one of the components in
the state. They used backward induction to solve the problem for 25 periods (6 years and 3
months). Based on their results they found that sulfonylurea should be used first.

2.5 Robust Optimization

Robust optimization (RO) is a modeling methodology for optimizing problems with uncertain
parameters. It seeks to protect the decision maker against parameter ambiguity and stochastic
uncertainty. The main paradigm in RO is the worst-case analysis which was originated from
decision theory in the 1950s. The development of RO in the field of operations research goes
back to the pioneering work on robust linear programming done by Soyster in 1973 [44]. In
recent years, there has been significant research on the theoretical development of RO, as well as
real-world applications including logistics [45, 46, 47], scheduling [48, 49], inventory management
[50, 51, 52], portfolio optimization [53, 54, 55], revenue management [56], energy systems [57],
queueing networks [58, 59], and machine learning [60]. Interested readers are referred to the
survey papers by Bertsimas in 2011 and Gabel in 2012 [61, 62], and the recent book of Ben-Tal
and El Ghaoui [63] for more comprehensive review of RO.

The remainder of the section is structured as follows. In Section 2.5.1, we give a brief
introduction to the general RO formulation for a given optimization problem. In Section 2.5.2, we review different uncertainty sets that have been used in literature. In Section 2.5.3, we focus on introducing the theoretic concepts, algorithms, and applications of robust MDP.

### 2.5.1 General RO Formulation

For any given nominal optimization problem (2.5)

\[
\begin{align*}
\max & \quad f_0(x, u_0) \\
\text{s.t.} & \quad f_i(x, u_i) \geq 0, \ i = 1, \ldots, m
\end{align*}
\]  

(2.5)

where \( u_i = (u_{i,1}, u_{i,2}, \ldots, u_{i,k_i}) \) is a vector of model parameters for constraint \( i \) that is subject to uncertainty, \( u_0 \) is a vector of parameters in the objective function that is subject to uncertainty, \( x \in \mathbb{R}^n \) is a vector of decision variables, \( f_0, f_i : \mathbb{R}^n \to \mathbb{R} \) are functions, and \( k_i \) is the number of uncertain parameters in the \( i^{th} \) constraint. The uncertain parameters \( u_i, i = 1, 2, \ldots, m \) can lead to the uncertainty in the feasibility and/or the optimality of the solution. The RO seeks to obtain a solution that performs well and is also feasible for any realization of the uncertain parameter taken from its uncertainty set [62]. Given uncertainty sets \( \mathcal{U}_i \) for the parameter \( u_i, i = 0, 1, \ldots, m \), the worst-case optimization formulation can be written as follows [63]:

\[
\begin{align*}
\max & \quad \inf_{u_0 \in \mathcal{U}_0} f_0(x, u_0) \\
\text{s.t.} & \quad f_i(x, u_i) \geq 0, u_i \in \mathcal{U}_i, i = 1, \ldots, m
\end{align*}
\]  

(2.6)

The optimization problem (2.6) can be reformulated as follows:

\[
\begin{align*}
\max & \quad z \\
\text{s.t.} & \quad f_i(x, u_i) \geq 0, u_i \in \mathcal{U}_i, i = 1, \ldots, m \quad (2.6')
\end{align*}
\]

\[
\begin{align*}
f_0(x, u_0) & \geq z, \ u_0 \in \mathcal{U}_0
\end{align*}
\]

(2.6')

The formulation (2.6') is called the robust counterpart of the nominal problem (2.5).
The two central issues in RO include the *conservativeness* of the RO formulation and the *tractability* of RO models which we summarize as follows:

- **Conservativeness.** An often cited shortcoming of RO models is over-conservatism. The uncertain set selection and construction should achieve a tradeoff between the performance and protection against uncertainty. Bertstimas and Brown [61] showed that there are several convenient, efficient, and well-motivated parameterizations of different classes of uncertainty sets that provide a notion of a *budget of uncertainty* which allow the designer a level of flexibility in choosing the tradeoff between robustness and performance. We provide a summary of those uncertainty sets in Section 2.5.2.

- **Tractability.** In general, the robust counterpart of a tractable optimization problem may not be tractable where tractable means that the problem can be reformulated into an equivalent problem that is solvable by a polynomial-time algorithm [63]. The tractability depends on the structure of the nominal problem as well as the class of the uncertainty set. In general, the robust counterpart of a tractable optimization problem is intractable. However, there are some well-known classes of optimization problems which have tractable robust counterpart or the computational complexity of the robust counterpart has been proved to be NP-hard when coupled with a specific uncertainty set. We summarize those RO problems in Table 2.3.

### 2.5.2 Uncertainty Set Models

The uncertainty set in an RO model defines the region within which model parameters may vary. As can be seen from Table 2.3, the choice of uncertainty set can significantly affect the complexity of an RO model. In this section, we describe some commonly used uncertainty set models, most of which are presented in the survey paper Bertsimas 2011 [61].

1. **Interval Model**

   The interval model, proposed by Soyster [44], is the earliest and the most widely used
Table 2.3: Computational Complexity of the Robust Counterpart of Some Classical Optimization Problems

<table>
<thead>
<tr>
<th>Nominal problem</th>
<th>Uncertainty set</th>
<th>Robust counterpart</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Programming (LP)</td>
<td>Ellipsoidal model</td>
<td>SOCP</td>
<td>[64]</td>
</tr>
<tr>
<td></td>
<td>Polyhedral model</td>
<td>LP</td>
<td>[64]</td>
</tr>
<tr>
<td></td>
<td>Cardinality constrained model</td>
<td>LP</td>
<td>[65]</td>
</tr>
<tr>
<td></td>
<td>$l_1$ Norm model</td>
<td>LP</td>
<td>[66]</td>
</tr>
<tr>
<td></td>
<td>$l_\infty$ Norm model</td>
<td>LP</td>
<td>[66]</td>
</tr>
<tr>
<td></td>
<td>$l_2$ Norm model</td>
<td>SOCP</td>
<td>[66]</td>
</tr>
<tr>
<td>Quadratic Optimization</td>
<td>Single ellipsoidal model</td>
<td>SDP</td>
<td>[67]</td>
</tr>
<tr>
<td></td>
<td>Polyhedral model</td>
<td>NP-hard</td>
<td>[67]</td>
</tr>
<tr>
<td>Semidefinite Optimization (SDP)</td>
<td>Ellipsoidal model</td>
<td>NP-hard</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td>Polyhedral model</td>
<td>NP-hard</td>
<td>[69]</td>
</tr>
</tbody>
</table>

SOCP: Second-Order Cone Programming

uncertainty model [70, 71, 72]. It is motivated by the concept of the statistical estimates of intervals of confidence. This model describes the uncertainty on parameter $u_i$ for constraint $i$ with the following form:

$$\mathcal{U}_i = \{ u_i : \underline{u}_i \leq u_i \leq \bar{u}_i \}, \forall i$$

(2.7)

where $\underline{u}_i$ and $\bar{u}_i$ are the given componentwise lower and upper bounds of the vector $u_i$.

2. Ellipsoidal Model

The ellipsoidal model was proposed by Ben-Tal and Nemirovski [73] and El-Ghaoui et al. [74]. It is very convenient from the mathematical point of view because it has a simple parametric representation and can be solved easily. The ellipsoidal model describes the uncertainty on parameter $u_i$ for constraint $i$ with the following form:

$$\mathcal{U}_i = \{ u_i : (u_i - E[u_i])^T \Sigma_i^{-1} (u_i - E[u_i]) \leq \rho_i \}, \forall i$$

(2.8)

where $E[u_i]$ is the expectation of parameter $u_i$, $\Sigma_i$ is the covariance matrix of $u_i$, and parameter $\rho_i$ has the interpretation of an uncertainty budget which allows the decision
maker to control the size of the ellipsoidal set in order to easily tradeoff robustness and performance. In practice, the maximal likelihood estimate $\hat{u}_i$ is used to replace $E[u_i]$.

3. Polyhedral Model

The polyhedral model can be viewed as a special case of the ellipsoidal model where the uncertainty set $\mathcal{U}$ has the following polyhedral representation:

$$\mathcal{U} = \{(u_0, u_1, \ldots, u_m) : a_i u_i \leq b_i, \forall i\} \quad (2.9)$$

Notice that the interval model is a special case of the polyhedral model.

4. Cardinality Constrained Model

Bertsimas and Sim [65] proposed the cardinality constrained model in 2004. It is defined as a family of polyhedral models with a budget of uncertainty in terms of cardinality constraints on the number of parameters of the problem that are allowed to vary from their mean values. The authors define parameter $\Gamma_i$ to denote the maximum number of $u_{i,j}$’s that can deviate from its mean value, and interpret it as the budget of uncertainty for the $i^{th}$ row that the decision–maker can control. The cardinality constrained model can be mathematically formulated as follows:

$$\mathcal{U} = \{(u_0, u_1, \ldots, u_m) : a_i u_i \leq b_i, |S_i| = \Gamma_i, \forall i\} \quad (2.10)$$

where $S_i$ denotes index set for the $u_{i,j}$’s that deviate from the mean value.

5. D-Norm model

The D-norm model is an instance of cardinality constrained model with the following form:

$$\mathcal{U}_i = \{u_i : \overline{u}_i + (\overline{u}_i - u_i)z_i, z_i \leq 1, \sum_{i=1}^{k_i} z_i \leq \Gamma_i\}, \forall i \quad (2.11)$$

It has been used portfolio selection in financial applications [61].
2.5.3 Robust MDP

As motivated in Section 2.3, MDPs provide an appealing mathematical modeling framework for optimizing sequential decision making problems. An optimal policy for an MDP can be computed efficiently by dynamic programming techniques. For large-scale problems, approximate solution methods are available [75, 76]. However, the application of MDPs is limited by the sensitivity of the optimal policy to the uncertainty in the model parameters. An optimal policy for a given set of model parameters might perform poorly on a system with slightly different parameter values, as illustrated recently in [45, 77].

Silver [78] initiated the pioneering study of MDPs with uncertain transition probabilities or rewards in 1963, and the work was expanded by Martin [79] in 1967. The framework is that some prior distribution is assumed at the start of the decision process, and Bayes formula is used to update information about unknown transition probabilities. The decision maker’s problem is to maximize the expected reward while updating relevant information. Later, Satia et al. [80], White et al. [81], and Bagnell et al. [82] studied MDPs with uncertain transition probabilities in the max-min framework. Due to the similarity with RO paradigm, the worst-case optimization of MDPs with uncertain transition probabilities is thereafter referred as robust MDP or RMDP.

RMDP Formulation

Similar to an MDP, a RMDP can be categorized into finite-horizon RMDPs and infinite-horizon RMDPs. We focus on model formulations and solution techniques for finite-horizon RMDPs in the reminder of this section to build the necessary foundation for the RMDP presented in Chapter 5. We refer interested readers to papers [45, 80, 83, 84] for more details on infinite-horizon RMDPs.

A general finite-horizon, discrete-time, discrete-state RMDP can be mathematically represented by a collection of objects,

\[(T, S, A_s, U_t^a, r_t(s, a))\]  \hspace{1cm} (2.12)
where $\mathcal{T} = \{1, 2, \ldots, T\}, T < \infty$ denotes the decision epochs, $\mathcal{S} = \{1, 2, \ldots, S\}$ denotes the state space, $A_s$ denotes the action space that the decision maker can take at state $s$, $r_t(s,a)$ denotes the reward function. Compared with the nominal MDP defined in Equation (2.1), the transition probability matrix $P_t^a$ is replaced by a collection of (possibly time dependent) transition matrices denoted by $\mathcal{U}_t^a$ which is referred to the uncertainty set of transition probability matrix $P_t^a$. In RMDP, a *policy of controller* is defined as $\pi = \{d_1, d_2, \ldots, d_{T-1}\} \in \Pi$, and for a given controller’s policy $\pi$, a *policy of nature* is defined as a specific collection of time-dependent transition matrices $\theta$ defined as follows:

$$\theta = \left( P_1^{d_1(s_1)}, P_2^{d_2(s_2)}, \ldots, P_{T-1}^{d_{T-1}(s_{T-1})} \right) \in \Theta(\pi) \triangleq \mathcal{U}_1^{d_1(s_1)} \times \mathcal{U}_2^{d_2(s_2)} \times \cdots \times \mathcal{U}_{T-1}^{d_{T-1}(s_{T-1})}. \quad (2.13)$$

For decision epoch $t \in \mathcal{T}$ and state $s \in \mathcal{S}$, the optimal robust value function $v_t(s)$ is defined as follows:

$$v_t(s) = \max_{\pi \in \Pi} \min_{\theta \in \Theta(\pi)} E_s^\theta \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, d_k(s_k)) + \lambda^{T-t} r_T(s_T) \right\} \quad (2.14)$$

where $\lambda \in (0, 1]$ is the discount factor, and $r_T(s_T)$ denotes the terminal reward that occurs at the end of the process when the system is in state $s_T$.

For a general uncertainty set $\mathcal{U}_t^a$, the optimal policy for a RMDP may not be Markovian, and the computational complexity is PSPACE (polynomial space)-hard [77]. However, two concurrent papers by Nilim and El Gaoui [45] and Iyengar [83] show that when the uncertainty set $\mathcal{U}_t^a$ satisfies the *rectangular uncertainty property*, namely,

$$\mathcal{U}_t^a = \mathcal{U}_{1,t}^a \times \mathcal{U}_{2,t}^a \times \cdots \times \mathcal{U}_{S,t}^a, \quad (2.15)$$

where $\mathcal{U}_{i,t}^a$ denote the uncertainty set of row $i$ of the transition probability matrix $P_t^a$, then the problem (2.14) reduces to a zero-sum sequential Markov game, and there exist an deterministic Markovian policy which is optimal. The robust optimal value function (2.14) can be written
Recursively as follows:

\[
v_t(s) = \max_{a \in A} \left\{ r_t(s, a) + \min_{p \in U_{t+1}} \sum_{s_{t+1} \in S} p(s_{t+1}) v_{t+1}(s_{t+1}) \right\}, \quad \forall s \in S, \forall t \neq T, \\
v_T(s) = r_T(s), \quad \forall s \in S, t = T.
\] (2.16)

The minimization problem in Equation (2.16), referred to as the inner problem.

Similar to the backward induction algorithm for finite-horizon MDP, Nilim and El Ghaoui [45] and Iyengar [83] both proposed the Robust finite-horizon dynamic programming algorithm (Algorithm 1) to solve the finite-horizon RMDP.

**Algorithm 1** Robust Finite-Horizon Dynamic Programming Algorithm

1. Step 1. Initialize the value function to its terminal value
2. for \( s_T \in S \) do
3. \( v_T(s_T) \leftarrow r_T(s_T) \)
4. end for
5. Step 2. Recursion
6. for \( t = T - 1 \rightarrow 0 \) do
7. for \( s_t \in S \) do
8. for \( a \in A_{s_t} \) do
9. Solve the inner problem, and to get the optimal value \( \sigma^*_t(s_t, a) \)
10. end for
11. Update Value Function: \( v_t(s_t) \leftarrow \max_{a \in A_{s_t}} \left\{ r_t(s_t, a) + \lambda \sigma^*_t(s_t, a) \right\} \)
12. Update Optimal Action Set: \( A^*_t(s_t) \leftarrow \arg\max_{a \in A_{s_t}} \left\{ r_t(s_t, a) + \lambda \sigma^*_t(s_t, a) \right\} \)
13. end for
14. end for

The choice of a specific uncertainty model determines the tractability of the RMDP, and the conservativeness of the robust optimal solution. Much research effort has been devoted to develop an understanding of the special structure of the uncertainty model to guarantee a computationally tractable RMDP. For instance, Mannor et al. [85] present a uncertainty model, termed “Lightning Does not Strike Twice” (LDST). The LDST model requires total number of uncertain parameters among different states to be less than a budget parameter. They provide
conditions under which the RMDP is tractable.

**RMDP Applications**

Theoretical advances have opened the door for the practical applications of RMDPs. However, RMDP applications are not nearly as broad as MDP applications. To the best of our knowledge, RMDPs have not yet been applied to health applications, such as medical decision making problems. Following are some recent examples of RMDP applications.

Nilim et al. [45] formulate a finite-horizon RMDP coupled with likelihood uncertainty model to optimize the aircraft routing problem. The authors find that the robust solution outperforms the nominal solution even with a small error in the estimation of transition probabilities, and the robust solution performs better than the nominal solution as the estimation error increases.

In order to compare the performance of a new robust modified policy iteration (RMPI) algorithm proposed in [86] with the robust value iteration proposed in [83], Kaufman et al. [86] formulated an infinite-horizon RMDP coupled with a relative entropy uncertainty model to optimize an inventory control problem. The numerical experiments showed that RMPI can reduce total computation time significantly compared to robust value iteration, and the degree to which computational effort is reduced can depend heavily on the optimal policy.

2.6 Contributions of this Dissertation

This dissertation contributes to the existing literature in the following ways. First, to our knowledge, we develop the first stochastic HbA1c model for glycemic control. Our model is more realistic than previously published deterministic HbA1c models because we consider both the expected rise in HbA1c over time caused by age and anticipated deterioration of glycemic control, and variations in the HbA1c evolution among patients. Moreover, we present a framework for calibrating and validating the HbA1c model using large claims data set. This framework can be used for researchers to study other populations of interest or using other data sets. We also present a method to compare the benefits and costs of new and old medications, and it can also
be used to evaluate other coming medications for glycemic control.

Second, we present two new optimization models: the MDP and the RMDP-TM for optimizing treatment decisions, including the sequence and starting times to initiate medications, for glycemic control. Both models allow for more patient specific decisions which depend on a patient’s gender, age, and HbA1c level as compared to the current treatment guideline which only depends on HbA1c level.

Third, to our knowledge, we present the first robust optimization model for optimizing medical treatment decisions when there are uncertainties in TPMs. We propose a new uncertainty set model for the transition probability which we show leads to a computationally tractable RMDP. We provide a fast algorithm to solve the RMDP approximately, and prove sufficient conditions under which the fast algorithm generates the exact robust optimal solution. We analyze the structure of the optimal value function and the resulting optimal policy for the RMDP and its relationship to its nominal MDP counterpart. Finally, we present a new framework for comparing robust optimal policies based on the method of sampling TPMs.
Evaluation of Glycemic Control for Type 2 Diabetes

3.1 Introduction

Diabetes is one of the most prevalent and costly chronic medical conditions worldwide, causing significant burdens on individuals, society, and the health care system. It has been estimated that 25.8 million Americans, or 8.3% of the population, have diabetes [87]. Glucose-lowering therapies are the cornerstone of diabetes management, with multiple epidemiological studies linking glycemic control to a lower risk of diabetes-related complications and mortality. Large randomized controlled trials have demonstrated a reduction in microvascular complications with intensive glycemic control, e.g., lowering HbA1c to, 6.5–8.0% (48–64 mmol/ mol), depending on the study [88, 89, 90, 91, 92, 93, 94, 95]. Evidence linking glycemic control to lower macrovascular disease risk and mortality has been less conclusive; lowering HbA1c among younger patients with newly diagnosed diabetes did reduce cardiovascular event rates and mortality in the UK Prospective Diabetes Study (UKPDS) [91, 92], but further reductions among people with long-standing diabetes in the Action to Control Cardiovascular Risk in Diabetes (ACCORD) and Action in Diabetes and Vascular Disease: Preterax and Diamicron MR Controlled Evaluation
(ADVANCE) studies and Veterans Affairs Diabetes Trial (VADT) did not yield similar results [93, 94, 95]. The exact glycemic target in the treatment of diabetes therefore remains controversial, with professional groups and regulatory organizations currently recommending lowering HbA1c to, 6.5% (48 mmol/mol) [4], 7.0% (53 mmol/mol) [5], or 8.0% (64 mmol/mol) [6], except in patients at high risk for hypoglycemia or those with limited life expectancy or multiple comorbid conditions that preclude safe intensive control.

There are currently 11 classes of approved glucose-lowering medications, and the usage of these medications has varied from 1994 to 2007 [96]. The 2011 Centers for Disease Control and Prevention diabetes fact sheet reported that 58% of adults with diabetes are being treated with oral agent(s), 12% with insulin, and 14% with both insulin and oral agent(s) [87]. Diabetes medications alone accounted for 11.8% of all prescriptions issued in the U.S. in 2012 at a cost of more than 18.3 billion USD [97]. Metformin has a long-standing evidence base for efficacy and safety, is inexpensive, and is regarded by most as the primary first-line agent in the treatment of type 2 diabetes [4, 5, 98]. When metformin fails to achieve or maintain glycemic goals, another agent should be added; however, there is no consensus or sufficient empirical evidence supporting the use of one second-line agent over another [99]. Over the past decade, the mix of secondary agents used in the treatment of diabetes has changed significantly, with increasing use of newer glucose-lowering agents such as dipeptidyl peptidase-4 (DPP-4) inhibitors and glucagon-like peptide-1 (GLP-1) receptor agonists in place of older and less expensive drugs such as sulfonylureas. This has resulted in a dramatic rise in the cost of diabetes medications and management; yet, the long-term clinical benefit of this shift is uncertain [96].

In the absence of clinical trials directly comparing alternative treatment regimens and considering the high cost and challenges of running any such trials, we developed and validated a new population-based glycemic control model based on a Markov chain to compare the real-world effectiveness and cost of different treatment regimens for individuals newly diagnosed with type 2 diabetes. We focus on pharmacologic approaches for glycemic control because lifestyle interventions are usually not enough for achieving glycemic control goals. Research has
shown that 91% of diabetic patients take prescribed medications [100]. We used this model to quantify differences among the regimens in terms of life-years (LYs), quality-adjusted life-years (QALYs), and medication cost per QALY necessary to achieve and maintain glycemic control from the time of diagnosis to the development of first major diabetes-related complication, specifically, ischemic heart disease, stroke, blindness, renal failure, amputation, or death from other cause. We specifically chose these micro- and macrovascular complications of diabetes, as they have been used in most large observational and interventional studies of diabetes therapies [91, 93, 94, 95]. Each regimen was tested using the range of currently recommended glycemic control goals between HbA1c 6.5% (48 mmol/mol) and 8% (64 mmol/mol) both to confirm model generalizability and to identify the potential impact of different glycemic control goals on patient health, quality of life, and expenditure. The results presented in this chapter were published in [101]

3.2 Markov Chain

In this section, we present the model construction, validation, and data sources that we used to calibrate the model.

3.2.1 Treatment Regimens

We considered an HbA1c threshold treatment policy, as shown in Figure 3.1. This policy is based on the ADA’s consensus algorithm for treating hyperglycemia for type 2 diabetes [3]. We considered four different treatment-intensification regimens: metformin, sulfonylurea, and insulin (T1); metformin, DPP-4 inhibitor, and insulin (T2); metformin, GLP-1 agonist, and insulin (T3); and metformin and insulin (T4). In each regimen, patients started metformin monotherapy when HbA1c reached the prespecified glycemic control goal. In T1–T3, treatment was sequentially intensified by addition of a second-line agent other than insulin, and if or when HbA1c again exceeded the glycemic control goal, insulin was initiated (in place of the second-line agent) as the third-line agent in combination with metformin. In T4, treatment was intensified by directly
Table 3.1: Model parameters for the base case analysis and sensitivity analysis

<table>
<thead>
<tr>
<th>Parameter (reference no.)</th>
<th>Base case value (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient’s characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Diagnose age [102]</td>
<td>Women: 55.2; Men 53.6</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>None Afro-Caribbean</td>
</tr>
<tr>
<td>Body mass index (kg/m²) [103]</td>
<td>32.6</td>
</tr>
<tr>
<td>Smoking status</td>
<td>Non-smoker</td>
</tr>
<tr>
<td>Concurrent comorbidity at diagnosis *</td>
<td>No</td>
</tr>
<tr>
<td>Blood pressure (mmHg) [5]</td>
<td>140 mmHg</td>
</tr>
<tr>
<td>Total cholesterol (mg/dL) [104]</td>
<td>200 mg/dL</td>
</tr>
<tr>
<td>HDL (mg/dL) [104]</td>
<td>40 mg/dL</td>
</tr>
<tr>
<td><strong>Glycemic control goals, % [4, 5, 6]</strong></td>
<td>7% (53 mmol/mol), 6.5% (48 mmol/mol), 8% (64 mmol/mol)</td>
</tr>
<tr>
<td><strong>Disutility of hypoglycemia [105]</strong></td>
<td></td>
</tr>
<tr>
<td>Metformin</td>
<td>-0.0002</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>-0.0064</td>
</tr>
<tr>
<td>DPP-IV inhibitor</td>
<td>-0.0002</td>
</tr>
<tr>
<td>GLP-1 agonist ‡</td>
<td>-0.0005</td>
</tr>
<tr>
<td>Insulin ‡</td>
<td>-0.0143</td>
</tr>
<tr>
<td><strong>Disutility of weight gain [105]</strong></td>
<td></td>
</tr>
<tr>
<td>Metformin</td>
<td>0</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>-0.0031</td>
</tr>
<tr>
<td>DPP-IV inhibitor</td>
<td>0</td>
</tr>
<tr>
<td>GLP-1 agonist §</td>
<td>0.0013</td>
</tr>
<tr>
<td>Insulin §</td>
<td>-0.0031</td>
</tr>
<tr>
<td><strong>Disutility of injectable medication [105]</strong></td>
<td></td>
</tr>
<tr>
<td>Metformin</td>
<td>0</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>0</td>
</tr>
<tr>
<td>DPP-IV inhibitor</td>
<td>0</td>
</tr>
<tr>
<td>GLP-1 agonist ‡</td>
<td>-0.0032</td>
</tr>
<tr>
<td>Insulin</td>
<td>-0.0032</td>
</tr>
<tr>
<td><strong>Month medication cost (USD) [99, 106]</strong></td>
<td></td>
</tr>
<tr>
<td>Metformin</td>
<td>81.75 (25.87–181.09)</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>54.85 (9.31–165.57)</td>
</tr>
<tr>
<td>DPP-IV inhibitor</td>
<td>232.84 (227.66–238.01)</td>
</tr>
<tr>
<td>GLP-1 agonist</td>
<td>325.97 (165.57–486.37)</td>
</tr>
<tr>
<td>Insulin</td>
<td>245.70 (189.39–327.54)</td>
</tr>
<tr>
<td><strong>Base case medication effect [10]</strong></td>
<td></td>
</tr>
<tr>
<td>Metformin</td>
<td>0.0661 (0.0620–0.0703)</td>
</tr>
<tr>
<td>Sulfonylurea</td>
<td>0.0937 (0.0852–0.1022)</td>
</tr>
<tr>
<td>DPP-IV inhibitor</td>
<td>0.0520 (0.0378–0.0662)</td>
</tr>
<tr>
<td>GLP-1 agonist</td>
<td>0.0558 (0.0472–0.0644)</td>
</tr>
<tr>
<td>Insulin</td>
<td>Maintain HbA1c at 7% (53 mmol/mol)</td>
</tr>
<tr>
<td><strong>Randomized control trial medication effect</strong></td>
<td></td>
</tr>
<tr>
<td>Sulfonylurea [107, 108]</td>
<td>(0.1282–0.2090)</td>
</tr>
<tr>
<td>DPP-IV inhibitor [107]</td>
<td>(0.0588–0.1149)</td>
</tr>
<tr>
<td>GLP-1 agonist [109, 110]</td>
<td>(0.0886–0.1744)</td>
</tr>
</tbody>
</table>

* Concurrent comorbidities include peripheral vascular disease, atrial fibrillation, ischemic heart disease, congestive heart failure, and blindness.
† Patients blood pressure, total cholesterol, and HDL were assumed to be well controlled by anti-hypertension and anti-hyperlipidemia medications.
‡ The disutility of hypoglycemia associated with insulin is set to be 2.24 times the disutility of hypoglycemia associated with sulfonylurea. This choice is motivated by the incidence rate of severe hypoglycemia among patients using each medication provided in ref. [111].
§ Weight loss is reflected in terms of gains in quality of life; therefore, it is associated with positive number.
¶ Values in the range represent the 95% CI of the estimated relative effect in reducing HbA1c. Sample sizes for estimating clinical effect were 2,118 for metformin, 765 for sulfonylurea, 204 for DPP-4 inhibitor, and 477 for GLP-1 agonist.
adding insulin once HbA1c exceeded the glycemic control goal. For all regimens, there were no further treatment changes once insulin was initiated, as it was assumed to maintain glycemic control.

Figure 3.1: The HbA1c threshold treatment policy based on the ADA’s consensus algorithm for treating hyperglycemia for type 2 diabetes [3].

3.2.2 Markov Model

The Markov chain has based on the 10 discrete HbA1c states presented in Table 3.2 and 3.3. Each state is defined by the conditional mean HbA1c in a given interval for a patient newly diagnosed with type 2 diabetes. The mean HbA1c value for each state increases linearly with respect to age according to a linear trend factor. This common assumption, based on other published glycemic control models [16, 112], reflects the expected rise in HbA1c with age and anticipated deterioration of glycemic control. At the beginning of each 3-month period, treatment is initiated/intensified if HbA1c exceeds the glycemic control goal. Treatment results in
a proportional decrease in HbA1c according to a medication effect estimated from observational
data (Table 3.1). If no diabetes complications or death occurs, patients undergo continued HbA1c
state transition based on the 3-month transition probability matrices provided in Table 3.2 and
3.3. Each treatment regimen was evaluated using the Markov model by backward induction [22].
All analyses were conducted using MATLAB R2012b (MathWorks, Inc., Natick, MA).

Table 3.2: Glycosylated hemoglobin (HbA1c) used in the Markov model for women. HbA1c
range definition at diagnosis, the mean natural HbA1c values for each HbA1c state at diagnosis
(prior to initiating medication), the initial HbA1c distributions at diagnosis, and 3-month HbA1c
transition probability matrices for women.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c Range</td>
<td>&lt;6</td>
<td>[6.6.5)</td>
<td>[6.5.7)</td>
<td>[7.7.5)</td>
<td>[7.5.8)</td>
<td>[8.8.5)</td>
<td>[8.5.9)</td>
<td>[9.9.5)</td>
<td>[9.5.10)</td>
<td>≥ 10</td>
</tr>
<tr>
<td>Mean HbA1c value (%)</td>
<td>5.7</td>
<td>6.25</td>
<td>6.74</td>
<td>7.24</td>
<td>7.73</td>
<td>8.23</td>
<td>8.73</td>
<td>9.22</td>
<td>9.72</td>
<td>11.73</td>
</tr>
<tr>
<td>Initial HbA1c Distribution</td>
<td>0.0771</td>
<td>0.1543</td>
<td>0.2125</td>
<td>0.18</td>
<td>0.1105</td>
<td>0.0848</td>
<td>0.0502</td>
<td>0.035</td>
<td>0.0273</td>
<td>0.0683</td>
</tr>
<tr>
<td>TPM HbA1c state 1</td>
<td>0.6379</td>
<td>0.3042</td>
<td>0.0481</td>
<td>0.0088</td>
<td>0.0010</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>0.1717</td>
<td>0.5085</td>
<td>0.2692</td>
<td>0.0412</td>
<td>0.0064</td>
<td>0.0020</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0010</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>0.0299</td>
<td>0.1731</td>
<td>0.5213</td>
<td>0.2258</td>
<td>0.0374</td>
<td>0.0085</td>
<td>0.0018</td>
<td>0.0004</td>
<td>0.0011</td>
<td>0.0007</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>0.0114</td>
<td>0.0538</td>
<td>0.2830</td>
<td>0.1467</td>
<td>0.1716</td>
<td>0.0446</td>
<td>0.0114</td>
<td>0.0029</td>
<td>0.0021</td>
<td>0.0025</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>0.0048</td>
<td>0.0240</td>
<td>0.1055</td>
<td>0.0650</td>
<td>0.3329</td>
<td>0.1678</td>
<td>0.0568</td>
<td>0.0199</td>
<td>0.0055</td>
<td>0.0089</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>0.0045</td>
<td>0.0116</td>
<td>0.0191</td>
<td>0.1438</td>
<td>0.2482</td>
<td>0.2768</td>
<td>0.1598</td>
<td>0.0661</td>
<td>0.0268</td>
<td>0.0134</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>0.0015</td>
<td>0.0120</td>
<td>0.0316</td>
<td>0.0648</td>
<td>0.1687</td>
<td>0.2364</td>
<td>0.2184</td>
<td>0.1370</td>
<td>0.0768</td>
<td>0.0527</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>0.0043</td>
<td>0.0065</td>
<td>0.0281</td>
<td>0.0562</td>
<td>0.0864</td>
<td>0.1533</td>
<td>0.1879</td>
<td>0.1965</td>
<td>0.1555</td>
<td>0.1253</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>0</td>
<td>0.0166</td>
<td>0.0194</td>
<td>0.0332</td>
<td>0.0831</td>
<td>0.1357</td>
<td>0.1662</td>
<td>0.1717</td>
<td>0.1828</td>
<td>0.1911</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>0.0078</td>
<td>0.0111</td>
<td>0.0277</td>
<td>0.0532</td>
<td>0.0831</td>
<td>0.0920</td>
<td>0.0854</td>
<td>0.0976</td>
<td>0.1042</td>
<td>0.4379</td>
</tr>
</tbody>
</table>

3.2.3 Outcome Measures

We considered five outcome measures related to primary prevention: total expected LYs, and
total expected QALYs from birth to occurrence of first diabetes-related complication or death,
mean time to insulin dependence, expected medication cost per QALY for maintaining glycemic
control, and incremental cost-effectiveness ratio (ICER) relative to no treatment from diagnosis
to occurrence of first diabetes-related complication or death. The expected total LYs/QALYs
from birth to first event are the sum of LYs to diagnosis and the LYs/QALYs from diagnosis to
Table 3.3: Glycosylated hemoglobin (HbA1c) used in the Markov model for men. HbA1c range definition at diagnosis, the mean natural HbA1c values for each HbA1c state at diagnosis (prior to initiating medication), the initial HbA1c distributions at diagnosis, and 3-month HbA1c transition probability matrices for men.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c Range</td>
<td>&lt;6</td>
<td>[6,6.5)</td>
<td>[6.5,7)</td>
<td>[7,7.5)</td>
<td>[7.5,8)</td>
<td>[8,8.5)</td>
<td>[8.5,9)</td>
<td>[9,9.5)</td>
<td>[9.5,10)</td>
<td>≥ 10</td>
</tr>
<tr>
<td>Mean HbA1c value (%)</td>
<td>5.69</td>
<td>6.25</td>
<td>6.73</td>
<td>7.24</td>
<td>7.74</td>
<td>8.24</td>
<td>8.74</td>
<td>9.21</td>
<td>9.73</td>
<td>11.59</td>
</tr>
<tr>
<td>Initial HbA1c Distribution</td>
<td>0.0694</td>
<td>0.1388</td>
<td>0.1968</td>
<td>0.1626</td>
<td>0.1138</td>
<td>0.0919</td>
<td>0.0619</td>
<td>0.0424</td>
<td>0.0328</td>
<td>0.0896</td>
</tr>
</tbody>
</table>

**TPM**

| HbA1c state 1 | 0.6245 | 0.2885 | 0.0685 | 0.0093 | 0.0034 | 0.0025 | 0.0008 | 0.0008 | 0 | 0.0017 |
| HbA1c state 2 | 0.1574 | 0.4949 | 0.2953 | 0.0402 | 0.0072 | 0.0038 | 0.0004 | 0 | 0.0004 | 0.0004 |
| HbA1c state 3 | 0.0349 | 0.2061 | 0.4715 | 0.2279 | 0.0441 | 0.0078 | 0.0024 | 0.0004 | 0.0024 | 0.0018 |
| HbA1c state 4 | 0.0130 | 0.0592 | 0.2462 | 0.4014 | 0.1971 | 0.0549 | 0.0166 | 0.0043 | 0.0029 | 0.0043 |
| HbA1c state 5 | 0.0098 | 0.0237 | 0.1058 | 0.2606 | 0.3029 | 0.1852 | 0.0686 | 0.0243 | 0.0083 | 0.0108 |
| HbA1c state 6 | 0.0058 | 0.0134 | 0.0645 | 0.1335 | 0.2313 | 0.2888 | 0.1514 | 0.0550 | 0.0294 | 0.0268 |
| HbA1c state 7 | 0.0104 | 0.0142 | 0.0455 | 0.0796 | 0.1308 | 0.2284 | 0.2351 | 0.1422 | 0.0645 | 0.0493 |
| HbA1c state 8 | 0.0111 | 0.0249 | 0.0456 | 0.0526 | 0.0982 | 0.1674 | 0.1840 | 0.1646 | 0.1328 | 0.1189 |
| HbA1c state 9 | 0.0125 | 0.0333 | 0.0412 | 0.0376 | 0.0789 | 0.1057 | 0.1595 | 0.1792 | 0.1344 | 0.2276 |
| HbA1c state 10 | 0.0098 | 0.0249 | 0.0537 | 0.0688 | 0.0629 | 0.0799 | 0.0911 | 0.0096 | 0.1134 | 0.3958 |

The ICER was calculated as follows:

\[ \text{ICER} = \frac{\text{costs of the treatment}}{\text{Expected QALYs gained from the treatment}} \]

### 3.2.4 Data Sources

A retrospective administrative claims data set that included medical claims, pharmacy claims, laboratory data, and eligibility information from a large, national U.S. health plan was used to estimate 3-month HbA1c state transition probabilities (Table 3.2 and 3.3), to estimate the medication effect on reducing HbA1c, and to calibrate and validate our model. The individuals covered by this health plan are geographically diverse across the U.S. with greatest representation in the south and midwest U.S. census regions. The plan provides fully insured coverage for...
professional (e.g., physician), facility (e.g., hospital), and outpatient prescription medication services. Medical (professional, facility) claims include ICD-9, Clinical Modification (ICD-9-CM) diagnosis codes, ICD-9 procedure codes, Current Procedural Terminology, version 4 procedure codes, Healthcare Common Procedure Coding System procedure codes, site of service codes, provider specialty codes, and health plan and patient costs. Outpatient pharmacy claims provide National Drug Codes for dispensed medications, quantity dispensed, drug strength, days supply, provider specialty code, and health plan and patient costs. Laboratory results linked to the administrative claims data are available for a subset of these patients. All study data were accessed using techniques that are in compliance with the Health Insurance Portability and Accountability Act of 1996, and no identifiable protected health information was extracted during the course of the study. Because this study involved analysis of preexisting, de-identified data, it was exempt from institutional review board approval.

The population meeting criteria for our study (37,501 individuals) were age 40 years or older, diagnosis with type 2 diabetes between 1995 and 2010, prescription for their first non-insulin glucose-lowering medication at least 6 months after enrollment, and at least 5 years of continuous enrollment with at least two HbA1c records and complete pharmacy claim data. Type 2 diabetes was defined using the Healthcare Effectiveness Data and Information Set criteria [114]. Healthcare Effectiveness Data and Information Set requirements for pharmacy data include at least one anti-hyperglycemia medication prescription and, for claim encounter data, the presence of at least one diabetes-specific ICD-9 diagnosis codes 250.XX (exclude 250.X1 and 250.X3), 357.2X, 362.0X, or 366.41 with two annual face-to-face outpatient encounters with different dates of service or one face-to-face in an acute inpatient or emergency department encounter.

3.2.5 Model Parameters for Base Case and Sensitivity Analysis

Model parameters, including base case values and ranges for sensitivity analysis, are shown in Table 3.1. We assumed a diagnosis age of 55.2 years for women and 53.6 years for men based on the median age at time of diagnosis of diabetes in the U.S. as of 2011 [102]. The initial HbA1c
state distributions for men and women are shown in Table 3.2 and 3.3. Treatment regimens were assumed to be fixed for patients living beyond 100 years, and future life expectancy at age 100 years was assumed to be 2.24 years for women and 2.05 years for men based on a 2008 U.S. life table [115].

To estimate the 3-month HbA1c transition probabilities, we selected all pairs of HbA1c records from the 37,501 eligible patients such that the period between tests was between 2.5 and 3.5 months and the patient was not on insulin during that time period. This resulted in 30,249 pairs (multiple pairs permitted per patient). Using the observed HbA1c value, $h_{i,t}$ of patient $i$ at time epoch $t$, the corresponding natural HbA1c value (without medication), $h_{i,t}^n$, was estimated as:

$$h_{i,t}^n = h_{i,t} \frac{1}{1 - \omega(m_{i,t})}, \forall i, t \quad (3.1)$$

Where $m_{i,t}$ denotes patient $i$’s current treatment regimen and $\omega(m_{i,t})$ is the estimated relative reduction in HbA1c shown in Table 3.1 when patient $i$ is using treatment regimen $m_{i,t}$ at time period $t$. We discretized all natural HbA1c values into 10 HbA1c states as defined in Tables 3.2 and 3.3. For any two HbA1c states, $a$ and $b$, we denoted the total number of transitions from state $a$ to state $b$ as $n_{ab}$. The maximum likelihood estimate of the transition probability from state $a$ to state $b$ was estimated as:

$$q_a(b) = \frac{n_{ab}}{\sum_{b \in L} n_{ab}}, \forall a \in L \quad (3.2)$$

where $L$ is the set of HbA1c states.

The probabilities of diabetes complications were determined by a patient’s age, sex, ethnicity (Afro-Caribbean or not), smoking status, BMI, HbA1c, systolic blood pressure, total cholesterol, and HDL cholesterol; history of peripheral vascular disease, atrial fibrillation, ischemic heart disease, and congestive heart failure; and blindness at diagnosis using the UKPDS outcomes model [116] (see Appendix A for a summary of this model) to calculate the probability of having macro- and micro-vascular complications. Probability of death from other cause was estimated...
based on the Centers for Disease Control and Prevention 2007 mortality tables [117].

The cost of medications other than insulin was based on the Federal median price for generic agents and the average wholesale price for brand name agents provided by the Agency for Healthcare Research and Quality Evidence Practice Centers [99]. The cost of insulin therapy, including the cost related to self-monitoring of blood glucose, insulin, and insulin-related supplies, was taken from Yeaw et al. [106]. All costs were inflation adjusted to 2013 dollars using the consumer price index method [118]. For medications other than insulin, the base case cost was the mean price of all brand name and generic (if available) medicines, and the cost in the range represents the least and the most expensive medicines. The base case cost for insulin was the mean cost of all insulin regimens including basal insulin regimens, premixed insulin regimens, and basal-bolus insulin regimens. The cost in the range represents the average cost for basal insulin therapy (the least expensive insulin therapy) and the average cost for basal-bolus insulin therapy (the most expensive insulin therapy), respectively.

Medication effect (other than for insulin) was estimated based on HbA1c changes seen 3-month before and after use of these agents from patients included in the data set. We also assumed that medications other than insulin had additive effect in reducing HbA1c [99]; therefore, each medication effect was estimated independently. For each medication other than insulin, we selected patients who had at least one HbA1c record within 3 months before and after its initiation, and who were treated with this medication for at least 3 consecutive months. For each selected patient, we calculated the pre-treatment HbA1c and the post-treatment HbA1c by taking the mean of his/her HbA1c records during the 3-month intervals before and after the date of initiation, respectively. The medication effect shown in Table 3.1 was then calculated as the overall mean relative change between the pre-treatment HbA1c and the post-treatment HbA1c of all the selected patients.
3.2.6 Model Calibration and Validation

To calibrate and validate the model we used all HbA1c pairs of the eligible 37,501 patients such that the period between HbA1c tests was greater than or equal to 3.5 months (in order to have at least one 3-month transition) and the patient was not on insulin during that time period. This resulted in 97,667 pairs. For each value of the linear trend factor, $\beta$, between 0 and 0.25, and for each initial test result in each pair, we simulated the second test result in the pair one hundred times using the 3-month HbA1c transition probability matrix (Tables 3.2 and 3.3) and the number of transitions, $t_k$, determined by the time interval between the two HbA1c tests of each pair $k$. Using the model-generated natural HbA1c state, $\ell^k(t_k)$, for each pair $k$, we calculated the model-generated HbA1c value, $h_{k,t_k}(\ell^k(t_k))$ with medications initiated during the time interval as follows:

$$h_{k,t_k}(\ell^k(t_k)) = h^n(\ell^k(t_k)) + \beta \times t_k - \omega(m_{k,t_k}) \times h^n(\ell^k(t_k))$$ (3.3)

where $h^n(\ell^k(t_k))$ is the mean natural HbA1c value of being in the HbA1c state $\ell^k(t_k) \in \mathcal{L}$ at diagnosis as shown in Tables 3.2 and 3.3, and $\omega(m_{k,t_k})$ is the medication effect of using medications $m_{k,t_k}$. Finally, we determined the model-generated HbA1c state for that pair based on the model-generated HbA1c value.

Given the one hundred model-generated 97,667 HbA1c pairs, we calculated the mean of the sum of the squared errors (SSE) between the model-generated HbA1c state distribution and the observed HbA1c state distribution as:

$$\bar{F}(\beta) = \sum_{i=1}^{n} (\pi - p^i(\beta))^T(\pi - p^i(\beta)) \overline{\pi}$$ (3.4)

where $\pi = (\pi(\ell_1), \pi(\ell_2), \cdots, \pi(\ell_{10}))$ represents the observed HbA1c state probability distribution (based on the second HbA1c values in all pairs) and the vector $p^i(\beta)$ represents the model-generated HbA1c state probability distribution for the $i^{th}$ simulation with a fixed linear trend
value $\beta$. The best linear trend was selected as the one that minimizes the mean of the SSE.

We found that the optimal trend factor was 0.1075 for men (mean SSE of 0.0022) and 0.105 for women (mean SSE of 0.0015) with the median difference between the observed HbA1c distribution and the simulated HbA1c distribution of 0.0096 (minimum: 0.0027, maximum 0.0271) for men and 0.0055 (minimum: 0.0000, maximum: 0.0197) for women.

### 3.3 Results

In this section, we present numerical results for base case analysis, and sensitivity analysis.

#### 3.3.1 Base Case Results

The Markov model-based results showed that the total expected LYs and QALYs from birth to first event produced by the four treatment regimens were similar (Table 3.4). The maximum difference among regimens in the total expected LYs to first event, specifically, the difference between T4 and T1, was 0.03 years (12.73 days) for women and 0.03 years (11.06 days) for men. Similarly, the maximum difference among regimens in the total expected QALYs to the first event, specifically, the difference between T4 and T1, was 0.04 QALYs (16.12 quality-adjusted days) for women and 0.04 QALYs (14.20 quality-adjusted days) for men. The observed differences in total expected LYs and QALYs among regimens were primarily the result of different expected durations of sustained glycemic control with the three second-line agents (in combination with metformin). The mean time elapsed between failure of metformin monotherapy and the need for insulin initiation was 1.05 years (381.99 days) for women and 1.0 year (364.65 days) for men using T1, 0.62 years (224.50 days) for women and 0.53 years (194.84 days) for men using T2, and 0.68 years (247.96 days) for women and 0.62 years (225.46 days) for men using T3.

Significant differences were observed in the expected medication cost per QALY incurred by the four treatment regimens. Compared with using sulfonylurea as a second-line agent, which was the least expensive treatment regimen, use of DPP-4 inhibitor (T2) was associated with a mean per-person additional medication cost of 141 USD per QALY for women and 160 USD per
Table 3.4: Base case comparison of four treatment regimens. Comparison of the total expected LYS, and total expected QALYs from birth to occurrence of first diabetes-related complication or death, expected medication cost per QALY and the ICER from diagnosis to occurrence of first diabetes-related complication or death, mean time from diagnosis to insulin initiation for men and women. Four treatment regimens are T1, metformin plus sulfonylurea plus insulin; T2, metformin plus DPP-4 inhibitor plus insulin; T3, metformin plus GLP-1 agonist plus insulin; and T4, metformin plus insulin.

<table>
<thead>
<tr>
<th>Treatment regimen</th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T1</td>
<td>T2</td>
</tr>
<tr>
<td>Total expected LYS</td>
<td>68.66</td>
<td>68.63</td>
</tr>
<tr>
<td>Total expected QALYs</td>
<td>68.41</td>
<td>68.39</td>
</tr>
<tr>
<td>Total expected medication cost (USD) per QALY</td>
<td>2600</td>
<td>2741</td>
</tr>
<tr>
<td>Mean time to use insulin (years)</td>
<td>2.76</td>
<td>2.33</td>
</tr>
<tr>
<td>ICER (cost/QALY)</td>
<td>9310</td>
<td>9865</td>
</tr>
</tbody>
</table>

QALY for men. Use of GLP-1 agonist (T3) incurred a mean additional medication cost of 191 USD per QALY for women and 216 USD per QALY for men compared with T1, and use of insulin as a second-line agent (T4) incurred a mean additional medication cost of 150 USD per QALY for women and 170 USD per QALY for men compared with T1.

3.3.2 Sensitivity Analyses

For any fixed glycemic control goal ranging between 6.5% (48 mmol/mol) and 8.0% (64 mmol/mol), use of sulfonyl as the second-line agent incurred the lowest expected medication cost per QALY, and GLP-1 agonist use incurred the highest expected medical cost per QALY, among both men and women (Fig. 3.2). Targeting a treatment goal of 6.5% (48 mmol/mol) vs. 7% (53 mmol/mol) incurred significantly higher expected medication cost per QALY and a small reduction in the total expected QALYs for all treatment regimens (Fig. 3.2). All treatment regimens resulted in increased total expected QALYs and increased medication cost per QALY when targeting a treatment goal of 7% (53 mmol/mol) compared with 8% (64 mmol/mol) (Fig. 3.2).

The expected medication cost per QALY of each of the four treatment regimens varied significantly (Fig. 3.3) as a result of differential costs incurred by generic (metformin, sulfonylurea)
Figure 3.2: QALYs versus cost incurred by the four different treatment regimens as a function of glycemic control goal. Comparison of the total expected QALYs versus the expected medication cost per QALY incurred from diagnosis to first event (diabetes-related complication or death) for men (5.10a) and women (5.10b). Each of the four treatments is compared as the glycemic control goal is varied from 6.5% (48 mmol/mol) to 8% (64 mmol/mol). Results are presented using HbA1c of 6.5% (48 mmol/mol) (circle), 7% (53 mmol/mol) (triangle), and 8% (64 mmol/mol) (square) as the glycemic control goal.
Figure 3.3: Sensitivity analysis on the medication cost. The x-axis represents the difference in the expected medication cost per QALY from the base-case cost: metformin costs 81.75 USD per month, sulfonylurea costs 54.85 USD per month, DPP-4 inhibitor costs 232.84 USD per month, GLP-1 agonist costs 325.97 USD per month, and insulin therapy costs 245.70 USD per month. The y-axis represents the treatment regimen. The solid bar represents men, and the hatched bar represents women. met, metformin; sulf, sulfonylurea.
compared with brand name (DPP-4, GLP-1) medications and basal insulin compared with basal plus bolus insulin regimens. T3 exhibited the largest variation in the expected medication cost per QALY (503 USD per QALY difference for women and 453 USD per QALY difference for men), while T2 was associated with the smallest variation in the expected medication cost per QALY (291 USD for women and 261 USD for men).

When the effects of sulfonylurea, DPP-4 inhibitor, and GLP-1 agonist on HbA1c were simultaneously set to be the lower bound or upper bound of the randomized control trial (RCT) results on the efficacy of medications (Table 3.1), the four treatment regimens still resulted in similar total expected LYs and QALYs from birth to first event. The treatment regimen with sulfonylurea as the second-line agent resulted in the lowest cost per QALY (2,537 USD per QALY for women and 2,612 USD per QALY for men at lower bound and 2,388 USD per QALY for women and 2,454 USD per QALY for men at upper bound), while the treatment regimen with GLP-1 agonist as the second-line agent still produced the highest cost per QALY (2,809 USD per QALY for women and 2,911 USD per QALY for men at lower bound and 2,867 USD per QALY for women and 2,971 USD per QALY for men at upper bound).

3.4 Conclusions

Direct comparison of four different diabetes treatment regimens by the Markov model developed and validated in this study demonstrated that all four treatment regimens resulted in similar expected benefits in LYs and QALYs irrespective of glycemic control goal. However, for all glycemic control goals ranging between the currently recommended targets of HbA1c 6.5% (48 mmol/mol) and 8% (64 mmol/mol), the use of sulfonylurea as the second-line agent incurred the lowest expected medication cost per QALY. These findings hold for both observed effects of medications from real-world data and randomized control trial results. The differences in cost per patient among the four treatment regimens were substantial and thus of potential importance to patients as well as health care providers and payers. In addition, the treatment regimen with a sulfonylurea as the second-line agent resulted in the longest time of insulin independence.
compared with all other regimens—an important factor to be considered by patients who wish to delay insulin initiation as long as possible. Conversely, the more expensive treatment options that use a DPP-4 inhibitor or a GLP-1 agonist as the second-line agent were associated with slightly less expected benefit in terms of both LYs and QALYs, and a shorter time of insulin independence, compared with the use of sulfonylurea. Use of insulin as the second-line agent resulted in the shortest time to insulin dependence, and was also significantly more expensive than using sulfonylurea with no added benefit in terms of LYs or QALYs.

To date, there has been no comprehensive side-by-side evaluation of the clinical benefits, effects on quality of life, and costs incurred by different diabetes treatment regimens for glycemic control. Our model fills this gap by integrating real-world knowledge of treatment costs, benefits, and harm, thereby allowing clinicians, payers, and patients to directly compare treatment regimens to select the one that is best suited for each individual patient given his/her specific glycemic control goal, cost sensitivity, and preference. Given that >25 million patients have been diagnosed with type 2 diabetes in the U.S., the potential policy implications of these differences uncovered by our model are also significant.

The Glycemia Reduction Approaches in Diabetes: A Comparative Effectiveness Study (GRADE), which is in the recruitment phase now, seeks to compare the same treatment regimens using a prospective clinical trial design; however, our model is significantly different from that of GRADE in that our results compare QALYs and costs for newly diagnosed patients and because our treatment efficacy is based on data that captures long-term adherence effects that are typically much smaller in clinical trials.

Several models have been developed to predict the natural history of diabetes-related complications progression and to gauge their sequelae on patient quality of life [15, 16, 17, 18, 112, 119]; however, none of these models were based on real-world data describing the rate of and variations in HbA1c progression caused by both biological changes and patient behavior with and without different treatment modalities. Moreover, none of the previous published models explicitly compared and contrasted different treatment regimens with regard to their practical
efficiency, cost, and clinical benefit based on real-world inputs rather than clinical trial data or select observational study population groups. To our knowledge, this study is the first to develop and validate a glycemic control model that takes into consideration the known adverse effects of treatment, such as hypoglycemia, current medication cost, and various suggested glycemic control goals.

Our model can serve as an adjunctive decision aid to facilitate treatment selection for people newly diagnosed with type 2 diabetes in a way that trades off health and economic implications for patients. It can also be used by health care providers and payers to determine whether a particular treatment option is consistent with the goal of high-value care, e.g., providing a clinically justified benefit given the incurred cost. While no clinical study has yet definitively established the clinical benefit of using incretins in place of sulfonylureas as second-line agents and there is increasing concern regarding sulfonylurea use owing to its association with severe hypoglycemia [4], our model, which considers the side effect of severe hypoglycemia, suggests that for a glycemic control goal of 6.5% (48 mmol/mol) or 7% (53 mmol/mol), sulfonylureas provide higher value than incretin. Indeed, use of incretins as second-line agents (treatment regimens T2 and T3) resulted in significantly higher cost but slightly less clinical benefit as measured by LYs and QALYs to first incident diabetes-related complication or death. However, ultimate value will depend on patient preference.

One strength of our study is the use of treatment effect estimates, based on real-world data. Since real world data accounts for potential lack of adherence and other behavioral factors that may not be present in the ideal setting of an RCT. We also recognize the merit of RCTs in limiting bias, and therefore we reported results based on treatment effects reported in RCTs. We find that using the treatment effect estimates from RCTs do not change our main conclusions.

Our study has several limitations. First, the results presented in this chapter are based on a Markov model rather than a clinical trial, and no model can provide a perfect representation of reality. Specifically, our model assumes that HbA1c varies among discrete states and at discrete 3-month time intervals rather than continuously; furthermore, transitions among states are
assumed to depend only on the most recent HbA1c state. For addressing these limitations, the assumptions were carefully validated based on real patient data. Treatment regimens were designed as sequential one-by-one additions of different classes of anti-hyperglycemic medications, while in clinical practice patients may start two or more drugs at the same time. We also assumed that insulin would replace the previously used second-line drug, as recommended by most clinical practice guidelines, but it is possible for patients to continue using two or more non-insulin agents in conjunction with insulin. We assumed that insulin will ultimately result in achievement of the glycemic goal; this is an idealized assumption that is based on the physiology of insulin action, and there is likely to be substantial variation among patients in whether they achieve and maintain their glycemic goal over time. We did not included the cost of hypoglycemia in our model. Finally, the model is based on data that represents a privately insured population. Therefore, it is possible that these results may not be generalizable to the Medicare and Medicaid populations.

Several features that were not incorporated into the current model are due to insufficient evidence in literature such as the potential variability in how medications influence HbA1c trajectory, the potential variability in the duration of observing the effect of medications, and the potential indirect pleiotropic effects of these medications not mediated by their glucose-lowering properties. Medication disutility values were based on limited empirical data because definitive evidence is not yet available. Our analyses were focused on primary prevention of the most common micro- and macrovascular complications of diabetes, and patients included were treatment naïve and newly diagnosed with diabetes. To the extent possible, we have used previously published data on the utility decrements for complications and treatments; however, utility estimates are limited in that they represent an average measure and do not reflect individual patients well-being. To address this, we performed sensitivity analysis on the utility estimates. Finally, not all known adverse medication effects were included in the model. We did not consider severe nausea and other gastrointestinal side effects of metformin or DPP-4 inhibitors [99], since these symptoms and availability of alternatives would likely cause
the medication to be discontinued. We did not consider pancreatitis risk from the new agents due to the uncertainty of this evidence [120, 121]. Ultimately, however, our proposed model is sufficiently versatile to allow for easy integration of newly acquired clinical knowledge and its continued refinement.

The cost of hypoglycemia was not included in the model. However, it would not change any of our conclusions. The estimate of the cost of hypoglycemia hospitalization of $17,564 per event for an inpatient admission with the incidence rate of 13.5 events per 10,000 person-years, $1387 for an emergency department (ED) visit with the incidence rate of 32.8 events per 10,000 person-years, and $394 for an outpatient visit with the incidence rate of 118.9 events per 10,000 person-years [122]. The estimated time frame for the newer medications, from Table 3.4 is 1.59 to 2.76 years. Thus, using the upper bound of 2.76 years as an estimate of the time on medication, the total average cost per patient due to hypoglycemia can be conservatively estimated at $89. Assuming all events are attributed to sulfonylurea, and none to the newer medications, this estimate would not change any of our conclusions.

Two key factors that were not explicitly incorporated into the model are medication adherence and lifestyle modifications, both of which are known to improve glycemic control, particularly in early stages of diabetes. However, this is alleviated by our use of real-world observational data for patients who adhere to their treatments and lifestyle recommendations with the frequency expected from any general population among which such therapies are to be deployed. This affords our model an aspect of generalizability and validity that makes it attractive and relevant to patients, health care providers, and payers.
Optimization of Glycemic Control for Type 2 Diabetes

4.1 Introduction

In this chapter, we formulate an MDP model to study the optimal sequence and time to initiate hyperglycemia lowering medications. As in Chapter 3, we focus on pharmacologic approaches for glycemic control for the U.S. population newly diagnosed with type 2 diabetes who have not yet initiated any treatment. In contrast to Chapter 3, which investigates simple HbA1c threshold treatment policies, in this chapter we investigate the optimal treatment policy which maximizes patients’ expected QALYs prior to the first event, where an event is defined as a micro- and/or macro-vascular complication (e.g., heart attack, stroke, kidney failure), or death from other causes. Our model optimally trades off the potential benefits from reducing HbA1c level (longer expected survival due to reduced risk of micro- and macro-vascular complications) with the disutility associated with the side effects of taking hyperglycemia lowering medications. The remainder of this chapter is structured as follows. In Section 4.2 we present the MDP model formulation. In Section 4.3 we explore model properties. In Section 4.4 we provide the numerical results and identify important clinical insights about optimal polices for the glycemic control.
Finally, in Section 4.5 we highlight the main conclusions and limitations.

4.2 Model Formulation

We present a finite-horizon MDP model to determine when and in what order hyperglycemia lowering medications should be initiated. We assume the treatment decisions are made at a discrete set of decision epochs indexed by \( t \in \mathcal{T} = \{1, 2, \ldots, T\} \) where decision epoch 1 refers to the age at diagnosis. Decisions are revisited periodically (e.g., quarterly visits to a clinician) over the patient’s lifetime up to some reasonable upper limit on a patient’s age (e.g., age 100). We refer to the period after decision horizon \( T \) as the *post decision horizon*, where patients are assumed to continue the same course of treatment as in the last decision epoch for the remainder of their life. Figure 4.1 shows a flow diagram of the treatment decision process. We will introduce each element in the figure in the following paragraphs.

![Figure 4.1: A flow diagram illustrating the entire treatment decision process.](image)

Our model includes *HbA1c states, medication states*, and *absorbing states*. The HbA1c state
at decision epoch $t \in T$ is denoted by

$$\ell_t \in \mathcal{L} = \{\ell(1), \ell(2), \ldots, \ell(k)\}, \quad (4.1)$$

where $\ell(i), \forall i \in K = \{1, 2, \ldots, k\}$ is defined by a certain range of HbA1c. HbA1c states are ordered according to HbA1c level, namely, a patient in HbA1c state $\ell(i)$ has lower HbA1c than a patient in HbA1c state $\ell(j)$ if $i < j$. If needed, we use the notation $\ell_t(i)$ to represent that a patient is in HbA1c state $\ell(i)$ at decision epoch $t$ for clarity. Medication states are used to keep track of a patient’s treatment regimen over time. The medication state during time interval $(t - 1, t]$ is denoted by

$$m_t \in \mathcal{M} = \{(m_{1,t}, m_{2,t}, \ldots, m_{n,t}) : m_{i,t} \in \{0, 1\}, \forall i \in N\} \quad (4.2)$$

where $N = \{1, 2, \ldots, n\}$ is the index set for medications, $n$ is the total number of medications. At decision epoch 1, $m_1$ represents the medication state at diagnosis. It is set to be 0 to represent that the patient is not on any hyperglycemia lowering medication before diagnosis. At decision epoch $t$ for $t > 1$, $m_{i,t} = 0$ represents that the patient is not on medication $i$ over the time interval $(t - 1, t]$, and $m_{i,t} = 1$ represents that the patient is on medication $i$ over the time interval $(t - 1, t]$. The absorbing state is represented by

$$d \in \mathcal{D} = \{D^O, E^{\text{macro}}, E^{\text{micro}}\}, \quad (4.3)$$

where $D^O$ denotes death from other causes, $E^{\text{macro}}$ denotes fatal or nonfatal macrovascular events, $E^{\text{micro}}$ denotes fatal or nonfatal microvascular events. The complete set of states in our model is given by $\mathcal{S} = \{\mathcal{L} \times \mathcal{M}\} \cup \mathcal{D}$.

At each decision epoch $t \in T$, the action is the selection of which medication to initiate, if any. Given the patient is in HbA1c state $\ell_t \in \mathcal{L}$, and medication state $m_t$ at decision epoch $t$, we denote $\alpha_t \in \mathcal{A}_t(\ell_t, m_t) = \{(A_{1,t}(\ell_t, m_t), A_{2,t}(\ell_t, m_t), \ldots, A_{n,t}(\ell_t, m_t))\}$ to be the action
taken at epoch $t$ to apply during the time interval $(t, t+1]$ where

$$A_{i,t}(\ell_t, m_t) = \begin{cases} I, & \text{to initiate or keep using medication } i \text{ during } (t, t+1], \\ D, & \text{otherwise,} \end{cases} \quad \forall i \in \mathcal{N}. \quad (4.4)$$

With this formulation, given a patient’s medication state, $m_t$, over time interval $(t-1, t]$, and action $\alpha_t$, at decision epoch $t$, the medication state, $m_{t+1}$, over the time interval $(t, t+1]$ can be updated as follows:

$$m_{i,t+1} = \begin{cases} 1, & \text{if } A_{i,t}(\ell_t, m_t) = I, \\ 0, & \text{if } A_{i,t}(\ell_t, m_t) = D, \end{cases} \quad \forall i \in \mathcal{N}, \quad (4.5)$$

for any $\ell_t \in \mathcal{L}, t \in \mathcal{T}$. For a patient in absorbing state $d \in \mathcal{D}$, there is no action to take, i.e., $\mathcal{A}_t(d) = \emptyset, \forall d \in \mathcal{D}, t \in \mathcal{T}$.

Our model includes two types of transition probabilities: transition probabilities among HbA1c states, and transition probabilities from HbA1c states to absorbing states. The transitions among HbA1c states and absorbing states are shown in Figure 4.2. Given that the patient is in the HbA1c state $\ell_t$ and medication state $m_t$, and take action $\alpha_t$ at decision epoch $t$, the probability of transiting into an absorbing state $d \in \mathcal{D}$ during the period $(t, t+1]$ is denoted by $p_t(d|\ell_t, m_t, \alpha_t)$, where

$$p_t(d|\ell_t, m_t, \alpha_t) = \begin{cases} \pi_t^O, & \text{if } d = D^O, \\ \pi_t^{\text{macro}}(\ell_t, m_t, \alpha_t), & \text{if } d = E^{\text{macro}}, \\ \pi_t^{\text{micro}}(\ell_t, m_t, \alpha_t), & \text{if } d = E^{\text{micro}}, \end{cases} \quad (4.6)$$

and $p_t(d|d) = 1$ for all $t \in \mathcal{T}$ and $d \in \mathcal{D}$. A macrovascular event (fatal and nonfatal) occurs with probability $\pi_t^{\text{macro}}(\ell_t, m_t, \alpha_t)$, and a microvascular event (fatal and nonfatal) occurs with probability $\pi_t^{\text{micro}}(\ell_t, m_t, \alpha_t)$. These transition probabilities depend on the HbA1c state, the medication state, the action, and other metabolic risk factors, such as TC, HDL, and blood
Figure 4.2: State transition diagram illustrating transitions among the HbA1c states, the absorbing states, and the medication states. The dashed box represents the set of absorbing states. The solid lines represent the HbA1c state transitions.

pressure which are age dependent in our model. Death from other causes occurs with probability $\pi_t^O$, and this probability is age-dependent in our model. Moreover, we define the total probability of moving into absorbing states by

$$p_t^D(\ell_t, m_t, \alpha_t) = \sum_{d \in D} p_t(d|\ell_t, m_t, \alpha_t) \quad (4.7)$$

for future reference. Note that in our model, the medication state and the action affect the transition probabilities of moving into an absorbing state by altering a patient’s HbA1c trajectory over time based on medication effects.

We denote the matrix of transition probabilities among HbA1c states conditional on not transitioning into an absorbing state by $Q_t = [q_{t,\ell_t}(\ell_{t+1})]_{k \times k}$ where $q_{t,\ell_t}(\ell_{t+1})$ represents the transition probability from state $\ell_t$ to state $\ell_{t+1}$ for any $\ell_t, \ell_{t+1} \in \mathcal{L}$. These probabilities do not depend on the medication state and the action since they are computed from the natural progression of HbA1c in the absence of medication.
We define the probability of a patient being in state $s_{t+1} \in S$ at epoch $t+1$ given the patient is in state $s_t \in S$ and takes action $\alpha_t \in A_t(s_t)$ at epoch $t$ as follows:

$$p_t(s_{t+1}|s_t, \alpha_t) = \begin{cases} (1 - p_t^D(\ell_t, m_t, \alpha_t))q_t, & \text{if } s_t = (\ell_t, m_t), s_{t+1} = (\ell_{t+1}, m_{t+1}) \in L \times M, \\ p_t(s_{t+1} | \ell_t, m_t, \alpha_t), & \text{if } s_t = (\ell_t, m_t) \in L \times M, s_{t+1} \in D, \\ 1, & \text{if } s_t, s_{t+1} \in D, \\ 0, & \text{otherwise.} \end{cases}$$ (4.8)

There are a number of possible choices for the reward function. Our choice is based on the QALYs utility model with multiplicative structure:

$$u(y_1, \ldots, y_n, t) = t \times u(y_1) \times u(y_2) \times \cdots \times u(y_n),$$

where $u(y_i)$ is the health utility associated with attribute $y_i$ [123], as noted in Chapter 3. Such models are commonly employed in the medical decision making literature [124]. The one-period reward function, $r_t(\ell_t, m_t, \alpha_t)$, can be written as follows:

$$r_t(\ell_t, m_t, \alpha_t) = \begin{cases} \delta (1 - D^\text{hyper}(\ell_t, m_t, \alpha_t)(1 - D^\text{med}(m_t, \alpha_t)), & \forall \ell_t \in L, \\ 0, & \text{otherwise.} \end{cases}$$ (4.9)

where $\delta$ is the length of time interval $(t, t+1]$ expressed in years, $D^\text{hyper}(\ell_t, m_t, \alpha_t)$ is the disutility of daily hyperglycemia symptoms when the patient is in HbA1c state $\ell_t$, medication state $m_t$, and takes action $\alpha_t$ during the period $(t, t+1]$, and $D^\text{med}(m_t, \alpha_t)$ is the disutility of taking medications over time interval $(t, t+1]$. If the patient is on more than one medication, $D^\text{med}(m_t, \alpha_t)$ is the sum of individual medication disutilities.

Given a patient is in HbA1c state $\ell_t \in L$, and medication state $m_t$ at decision epoch $t$, we let $v_t(\ell_t, m_t)$ denote the patient’s maximum total expected QALYs prior to the first event under
the optimal treatment policy. The optimality equations can be written as follows:

\[ v_t(\ell_t, m_t) = \begin{cases} \max_{\alpha_t \in A_t(\ell_t, m_t)} \left\{ r_t(\ell_t, m_t, \alpha_t) + [1 - p_t^D(\ell_t, m_t, \alpha_t)] \lambda \sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_t}(\ell_{t+1}) v_{t+1}(\ell_{t+1}, m_{t+1}) \right\} \\
R_T(\ell_T, m_T), \quad \forall \ell_T \in \mathcal{L}, m_T \in \mathcal{M}, t = T, \\
0, \quad \text{otherwise}, 
\end{cases} \quad (4.10) \]

for any \( \ell_t \in \mathcal{L}, t \in T \), where \( \lambda \in (0, 1] \) is the discount factor, \( R_T(\ell_T, m_T) \) denotes the end-of-horizon reward which represents the expected reward for a patient living past the decision horizon (e.g., past age 100). We keep the definition of the end-of-horizon reward in the general form here, and provide the specific estimations for numerical experiments in Section 4.4.1. For all \( t \in T \setminus \{T\} \), the optimal action set can be written as follows:

\[ \alpha_t^* = \arg\max_{\alpha_t \in A_t(\ell_t, m_t)} \left\{ r_t(\ell_t, m_t, \alpha_t) + [1 - p_t^D(\ell_t, m_t, \alpha_t)] \lambda \sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_t}(\ell_{t+1}) v_{t+1}(\ell_{t+1}, m_{t+1}) \right\}. \quad (4.11) \]

### 4.3 Model Properties and Insights

This section provides some properties of the MDP model. We provide and prove sufficient conditions under which the optimal value function is nonincreasing in a patient’s HbA1c state for any medication states and decision epochs, and sufficient conditions under which the optimal policy is of control limit type. We use the notation \( \ell_t(i) \) to represent that a patient is in HbA1c state \( \ell(i) \) at decision epoch \( t \) for clarity.

#### 4.3.1 Model Properties

The following Lemma 4.3.1 will be used to prove the monotonic property of the value function with respect to HbA1c states. Lemma 4.3.1 is similar to the Lemma 4.7.2 in Puterman’s book [22], however, Puterman’s lemma is for the infinite-horizon case, i.e. \( N = \infty \), and for nondecreasing sequence \( \{v_j\}_{j=0,1,...} \).
Lemma 4.3.1. Let \( \{x_j\}, \{x'_j\} \) be real-value non-negative sequences satisfying

\[
\sum_{j=n}^{N} x_j \geq \sum_{j=n}^{N} x'_j
\]

for all \( n \), with equality holding for \( n = 0 \). Suppose \( v_{j+1} \leq v_j \) for \( j = 0, 1, \ldots, N \), then

\[
\sum_{j=0}^{N} v_j x_j \leq \sum_{j=0}^{N} v_j x'_j
\]

Proof. Let \( n \) be arbitrary and \( v_{-1} = 0 \). Then

\[
\sum_{j=0}^{N} v_j x_j = \sum_{j=0}^{N} x_j \sum_{i=0}^{j} (v_i - v_{i-1}) = \sum_{j=0}^{N} (v_j - v_{j-1}) \sum_{i=j}^{N} x_i = \sum_{j=1}^{N} (v_j - v_{j-1}) \sum_{i=j}^{N} x_i + v_0 \sum_{i=0}^{N} x_i
\]

\[
\leq \sum_{j=1}^{N} (v_j - v_{j-1}) \sum_{i=j}^{N} x'_i + v_0 \sum_{i=0}^{N} x_i \quad (\because v_j - v_{j-1} \leq 0, \forall j = 1, \ldots, N)
\]

\[
= \sum_{j=1}^{N} (v_j - v_{j-1}) \sum_{i=j}^{N} x'_i + v_0 \sum_{i=0}^{N} x'_i \quad (\because \sum_{j=0}^{N} x_j = \sum_{j=0}^{N} x'_j)
\]

\[
= \sum_{j=0}^{N} (v_j - v_{j-1}) \sum_{i=j}^{N} x'_i = \sum_{j=0}^{N} x'_j \sum_{i=0}^{j} (v_i - v_{i-1}) = \sum_{j=0}^{N} v_j x'_j
\]

Therefore, \( \sum_{j=0}^{N} v_j x_j \leq \sum_{j=0}^{N} v_j x'_j \) holds. \( \square \)

The following proposition 4.3.1 states that the value function does not increase as the patient’s HbA1c becomes worse.

Proposition 4.3.1. If the following conditions hold:

(i): the transition probability matrix, \( Q_t \), has the following property:

\[
H_t(\ell_t(j)|\ell_t(i)) = \sum_{s=j}^{k} q_{t,\ell_t(i)}(\ell_{t+1}(s)), \forall t \in T \quad (4.12)
\]
is nondecreasing in $i$ for all $j = 1, \ldots, k$.

(ii): the one-period reward, $r_t(\ell_t(i), m_t, \alpha_t)$, is nonincreasing in $i$ for any $m_t \in \mathcal{M}, \alpha_t \in \mathcal{A}_t(\ell_t, m_t), t \in T$,

(iii): the end-of-horizon reward, $R_T(\ell_T(i), m_T)$, is nonincreasing in $i$ for any $m_T \in \mathcal{M}$,

(iv): the probability of transitioning into the absorbing state, $p^D_t(\ell_t(i), m_t, \alpha_t)$, is nondecreasing in $i$ for any $m_t \in \mathcal{M}, \alpha_t \in \mathcal{A}_t(\ell_t, m_t), t \in T$,

then the value function $v_t(\ell_t(i), m_t)$ is nonincreasing in $i$ for any $t \in T$ and $m_t \in \mathcal{M}$.

The property presented in assumption (i) is commonly called the increasing failure rate (IFR) property, and commonly used in proving the structural properties of MDPs [32, 34, 37]. It can be interpreted to mean that a patient in a high HbA1c state has higher probability of staying in the current or a worse HbA1c state than a patient in a low HbA1c state. Assumptions (ii) and (iii) state that the one-period reward and the end-of-horizon reward are nonincreasing in HbA1c state for any medication state. In other words, the rewards don’t improve as HbA1c increases. Finally, Assumption (iv) states that the probability of moving to the absorbing state is nondecreasing in HbA1c states for any medication state. This is a reasonable assumption because the probability of having macro- and/or micro-vascular events is nondecreasing as the HbA1c state increases according to the risk equations shown in Appendix A; the probability of death from other causes is only age dependent.

Proof. This proof is by induction. For the base case $t = T$, we know $v_T(\ell_T(i), m_T) = R_T(\ell_T(i), m_T)$ is nonincreasing in $i$ by assumption (iii). Now assuming that $v_\tau(\ell_\tau(i), m_\tau)$ is nonincreasing in $i$ for $\tau = t + 1, t + 2, \ldots, T - 1$, then we need to prove this for $\tau = t$. Assume $a, b \in \{1, 2, \ldots, k\}$ and $a < b$, the optimal value functions for state $\ell_t(a)$ and $\ell_t(b)$ for a given
According to the definition of action set, \( A \)

Thus \( v \)

Therefore, we have

satisfied, and the result follows.

\[ \text{medication state } m_t \text{ are} \]

\[ v_t(\ell_t(a), m_t) = \max_{\alpha \in A_t(\ell_t(a), m_t)} \{ r_t(\ell_t(a), m_t, \alpha_t) \]

\[ +[1 - p_t^P(\ell_t(a), m_t, \alpha)] \lambda \sum_{s=1}^{k} q_{t,\ell_t(a)}(\ell_{t+1}(s))v_{t+1}(\ell_{t+1}(s), m_{t+1}) \}, \]

(4.13)

and

\[ v_t(\ell_t(b), m_t) = \max_{\alpha \in A_t(\ell_t(b), m_t)} \{ r_t(\ell_t(b), m_t, \alpha_t) \]

\[ +[1 - p_t^P(\ell_t(b), m_t, \alpha)] \lambda \sum_{s=1}^{k} q_{t,\ell_t(b)}(\ell_{t+1}(s))v_{t+1}(\ell_{t+1}(s), m_{t+1}) \}, \]

(4.14)

Let \( \alpha_b^* \) denotes the optimal action such that

\[ v_t(\ell_t(b), m_t) = r_t(\ell_t(b), m_t, \alpha_b^*) \]

\[ +[1 - p_t^P(\ell_t(b), m_t, \alpha_b^*)] \lambda \sum_{s=1}^{k} q_{t,\ell_t(b)}(\ell_{t+1}(s))v_{t+1}(\ell_{t+1}(s), m_{t+1}). \]

(4.15)

According to the definition of action set, \( A_t(\ell_t(b), m_t) = A_t(\ell_t(a), m_t) \), so \( \alpha_b^* \in A_t(\ell_t(a), m_t) \).

Therefore, we have

\[ v_t(\ell_t(b), m_t) = r_t(\ell_t(b), m_t, \alpha_b^*) \]

\[ +[1 - p_t^P(\ell_t(b), m_t, \alpha_b^*)] \lambda \sum_{s=1}^{k} q_{t,\ell_t(b)}(\ell_{t+1}(s))v_{t+1}(\ell_{t+1}(s), m_{t+1}) \]

\[ \leq r_t(\ell_t(b), m_t, \alpha_b^*) \quad \text{(By Lemma 4.3.1)} \]

\[ +[1 - p_t^P(\ell_t(b), m_t, \alpha_b^*)] \lambda \sum_{s=1}^{k} q_{t,\ell_t(a)}(\ell_{t+1}(s))v_{t+1}(\ell_{t+1}(s), m_{t+1}) \]

(4.16)

\[ \leq r_t(\ell_t(a), m_t, \alpha_b^*) \quad \text{(By Assumption (ii))} \]

\[ +[1 - p_t^P(\ell_t(a), m_t, \alpha_b^*)] \lambda \sum_{s=1}^{k} q_{t,\ell_t(a)}(\ell_{t+1}(s))v_{t+1}(\ell_{t+1}(s), m_{t+1}) \]

\[ \leq v_t(\ell_t(a), m_t). \]

Thus \( v_t(\ell_t(i), m_t) \) is nonincreasing in \( i \) for any \( m_t \in M, t \in T \), the induction hypothesis is satisfied, and the result follows. \[ \square \]
Proposition 4.3.1 shows that the patient’s expected future reward does not decrease as HbA1c decreases. This fact is used to prove Theorem 4.3.1. In our numerical experiments, we find that the HbA1c state TPM, $Q_t$, is not exactly IFR; however, the patient’s expected future reward is still monotonic nonincreasing in HbA1c state for any medication states and decision epochs.

**Theorem 4.3.1.** If the following conditions hold:

1. the transition probability matrix, $Q_t$, has the IFR property,

2. the one-period reward, $r_t(\ell_t, m_t, \alpha_t)$, is nonincreasing in $\ell_t$ for any $m_t \in \mathcal{M}$, $\alpha_t \in \mathcal{A}_t(\ell_t, m_t)$, $t \in \mathcal{T}$,

3. the end-of-horizon reward, $R_T(\ell_T, m_T)$, is nonincreasing in $\ell_T$ for any $m_T \in \mathcal{M}$,

4. the probability of transitioning into the absorbing state, $p_D^t(\ell_t, m_t, \alpha_t)$, is nondecreasing in $\ell_t$ for any $m_t \in \mathcal{M}$, $\alpha_t \in \mathcal{A}_t(\ell_t, m_t)$, $t \in \mathcal{T}$,

5. the one-period reward, $r_t(\ell_t, m_t, \alpha_t)$, is superadditive function on $\mathcal{L} \times \mathcal{A}$ for any $m_t \in \mathcal{M}$, $t \in \mathcal{T}$, and

6. 

$$\sum_{\ell_{t+1} \in \mathcal{L}} [1 - p_D^t(\ell_t, m_{t+1})]q_{t,t+1}(\ell_{t+1})v_{t+1}(\ell_{t+1}, m_{t+1})$$

is superadditive function on $\mathcal{L} \times \mathcal{A}$, $\forall m_t \in \mathcal{M}$, $t \in \mathcal{T}$,

then there exists optimal decision rules which are nondecreasing in $\ell_t$ for any $m_t$, $t \in \mathcal{T}$.

The proof is based on applying Theorem 4.7.5 of Puterman [22]. Theorem 4.3.1 is important in that it provides conditions under which a simple threshold control policy is optimal for clinical applications in practice. However, it is worth noting that not all conditions necessarily hold. In particular, condition 1 is violated (to a small degree) in some cases, and conditions 5 and 6 may not hold in all cases. In the numerical experiments, we will show a case when the optimal policy is of control-limit type and a case when the optimal policy is not of control-limit type.
4.4 Results

In this section we present numerical results illustrating the optimal time and sequence in which to initiate the three most commonly used medications: metformin, sulfonylurea, and insulin for the U.S. population newly diagnosed with type 2 diabetes. We consider these three medications because they were shown to be cost-effective based on the conclusions from Chapter 3. Backward induction was used to compute the optimal treatment policy for all instance of the MDP. Our model and solution method were coded in C/C++, and experiments were performed on a Dell Windows workstation with an Intel Core i7 2.79GHz CPU and 8GB shared RAM.

The remainder of this section is organized as follows: In Section 4.4.1, we define the specific model parameter estimates and the data sources. In Section 4.4.2, we discuss numerical results illustrating the optimal policy for base case male and female patients. Finally, in Section 4.4.4, we compare the optimal policy with the actual policy obtained from our data set, and the current guideline.

4.4.1 Data Sources and Parameter Estimation

The data source we used to calibrate our model is the same retrospective administrative claims data set that we described in Section 3.2.4 in Chapter 3. In the numerical experiments, we look at the same population as in Chapter 3. Patients’ baseline characteristics are shown in Table 3.1. The initial HbA1c state distributions, mean HbA1c value at diagnosis, and HbA1c state transition probability matrices for men and women are shown in Tables 3.2 and 3.3. We consider 3-month decision epochs, which is consistent with the current recommendation for the frequency of HbA1c tests [5]. Decisions are made until age 100 and then the patients are assumed to continue the same course of treatment for the remainder of their life. We chose age 100 for two reasons. First, the average life expectancy after 100 years old is only 2.24 years for women and 2.05 years for men, as reported in [115]. Second, the probability of having no macro- and/or micro-vascular event or death occur until age 100 is very low. For example, it is only 0.0723 for females and 0.0204 for males using estimates based on the Centers for Disease Control and
Prevention (CDC) 2007 mortality table [117]. The discount factor $\lambda$ is set to be 1.

In contrast to Chapter 3, we also consider the disutility of hyperglycemia symptoms associated with high HbA1c. The disutility of hyperglycemia daily symptoms, $D^{\text{hyper}}(\ell_t, m_t, \alpha_t), \forall t \in T$ was defined to be a stepwise function as follows:

$$D^{\text{hyper}}(\ell_t, m_t, \alpha_t) = \begin{cases} 
0, & \text{if } \bar{h}(\ell_t, m_t, \alpha_t) \leq 8\%, \\
0.035, & \text{otherwise} 
\end{cases}$$

(4.17)

where $\bar{h}(\ell_t, m_t, \alpha_t)$ represents the mean HbA1c value when patients are in HbA1c state $\ell_t$, medication state $m_t$, and take action $\alpha_t$ at decision epoch $t$. For any $m_t \in M$, and $\alpha_t \in A_t(\ell_t, m_t)$, $\bar{h}(\ell_t, m_t, \alpha_t)$ is nondecreasing in $\ell_t$, therefore, $D^{\text{hyper}}(\ell_t, m_t, \alpha_t)$ is nondecreasing in $\ell_t$. The 8% threshold is consistent with clinical HbA1c threshold for observing hyperglycemia daily symptoms, and 0.035 is the disutility of these symptoms [125].

The end-of-horizon reward, $R(\ell_T, m_T)$, was estimated as follows:

$$R(\ell_T, m_T) = \begin{cases} 
L \times (1 - D^{\text{hyper}}(\ell_T, m_T))(1 - D^{\text{med}}(m_T)), & \text{if } \ell_T \in \mathcal{L}, m_T \in M \\
0, & \text{otherwise} 
\end{cases}$$

(4.18)

where $L$ is the life expectancy after 100 years which are assumed to be 2.24 years for women and 2.05 years for men based on a 2008 U.S. life table [115]. For any $m_T \in M$, $L$ and $1 - D^{\text{med}}(m_T)$ are constants, and $D^{\text{hyper}}(\ell_T, m_T)$ is nondecreasing in $\ell_T$, thus the end-of-horizon reward, $R(\ell_T, m_T)$, is nonincreasing in $\ell_T$.

The medications’ effects on HbA1c and their disutilities are shown in Table 3.1. The probabilities of diabetes complications are determined by a patient’s age, sex, ethnicity (Afro-Caribbean or not), smoking status, BMI, HbA1c, systolic blood pressure, total cholesterol, and HDL cholesterol; history of peripheral vascular disease, atrial fibrillation, ischemic heart disease, and congestive heart failure; and blindness at diagnosis using the UKPDS outcomes model [116]. The probabilities of death from other causes were obtained from the CDC mortality tables [117] (see Appendix B for a summary).
A complete list of model input sources can be found in Table 4.1.

Table 4.1: Summary of sources for model inputs.

<table>
<thead>
<tr>
<th>Model Input</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probabilities among HbA1c States</td>
<td>A claims data set with linked laboratory date</td>
</tr>
<tr>
<td>Probabilities of micro- and macro-vascular events</td>
<td>UKPDS outcome model [116]</td>
</tr>
<tr>
<td>Probability of death from other causes</td>
<td>CDC mortality tables [117]</td>
</tr>
<tr>
<td>Probability of hypoglycemia</td>
<td>UKPDS [126]</td>
</tr>
<tr>
<td>End-of-horizon reward</td>
<td>CDC life-expectancy tables [115]</td>
</tr>
<tr>
<td>Utility of medications</td>
<td>Sinha et al. [105]</td>
</tr>
<tr>
<td>Utility of hyperglycemia</td>
<td>Kahn et al. [125]</td>
</tr>
</tbody>
</table>

4.4.2 Base Case Analysis

In the base case analysis, we use a utility decrement as part of the medication utility to model hypoglycemia associated with sulfonylurea and insulin, and discontinue the use of sulfonylurea if insulin is initiated. The optimal expected QALYs prior to the first event is 68.50 QALYs for women and 64.46 QALYs for men. Compared with no treatment, the expected QALYs prior to first event are increased by 4.03 QALYs and 3.07 QALYs for women and men, respectively.

The optimal policy for patients who are on insulin is to keep using current medication(s). The optimal policies for patients who are not on insulin, including patients not on any medications, patients on metformin only, patients on sulfonylurea only, and patients on metformin and sulfonylurea, are shown in Figure 4.3. We find that the optimal policies are of control-limit type although the HbA1c TPMs do not satisfy the IFR property exactly. The optimal sequence to initiate medications between women and men are the same, but the time to initiate each medication is different. At the time of diagnosis when patients are not on any medication, the optimal policy for those who with HbA1c less than 10% is to use metformin and sulfonylurea together; and for those who with HbA1c ≥ 10% is to initiate insulin immediately. All patients will eventually start insulin due to the setup of the model where the glycemic control is assumed
to deteriorate over time.

Figure 4.3: The first 6 years optimal policy from diagnosis for patients who are not on insulin, including patients not on any medications, patients on metformin only, patient’s on metformin and sulfonylurea, when hypoglycemia is modeled as a transient symptom (base case). The following medication states are never optimal in the figure: no medication, metformin only, and sulfonylurea only.

4.4.3 Impact of Hypoglycemia on Optimal Policy

The results from the ACCORD trial have recently drawn attention to the potential risks of frequent hypoglycemia. In this section, we consider a conservative way to model hypoglycemia in order to investigate how it affects the optimal policy and the optimal expected QALYs prior to first event. To be consistent with clinical goal of avoiding hospitalization and/or possible mortality associated with hypoglycemia, we consider the hypoglycemia as an event (an absorbing state) that physicians/patients try to prevent. Based on the results from the intensive glycemia control arm in UKPDS study, the quarterly probability of hypoglycemia associated
with insulin, sulfonylurea, and metformin are set to 0.0096, 0.003, and 0.0008, respectively, where hypoglycemia includes cases where the patient is temporarily incapacitated but able to control symptoms without help, incapacitated and requires assistance to control symptoms, and requires medical attention or a glucagon injection [11].

The optimal expected QALYs prior to the first event is 65.71 QALYs for women and 62.40 QALYs for men. Compared with no treatment, the expected QALYs prior to first event are increased by 1.24 QALYs and 1.01 QALYs for women and men, respectively.

The optimal policy for patients who are not on insulin is shown in Figure 4.4. We find that the optimal policy is not of control-limit type when we consider hypoglycemia as an event. This is because the assumptions (1, 2, 5) are violated. The optimal sequence and time to initiate medications are different for women and men. At the time of diagnosis when patients are not on
any medication, the optimal policy for women with HbA1c greater than 8% or men with HbA1c greater than 6.5% is to use metformin monotherapy; and for women with HbA1c less than 8% or men with HbA1c less than 6.5% the optimal action is not to initiate any medication. Same as the base case, patients will eventually start insulin due to deterioration of glycemic control.

4.4.4 Comparison of the optimal policy and HbA1c threshold treatment policies

In this section, we compare the expected QALYs prior to the first event with different treatment policies including the optimal policy and HbA1c threshold treatment policies with various HbA1c thresholds from published treatment guidelines and estimated from the retrospective administrative data set. The HbA1c threshold treatment policies, with HbA1c thresholds obtained from published treatment guidelines are referred to as treatment guidelines. The HbA1c threshold treatment policy, with HbA1c thresholds estimated from the retrospective administrative data set, is referred to as the observed average policy. In the observed average policy, patients initiate metformin when HbA1c reaches 7.66%, on average. Sulfonylurea is used to intensify treatment when HbA1c reaches 8.38%, on average. Insulin is initiated to substitute sulfonylurea when HbA1c reaches 9.26%, on average. In the observed average policy and the ADA’s treatment guideline, the time to initiate each medication only depends on the HbA1c level, therefore, the policies are the same regardless of how the hypoglycemia is modeled.

As shown in Table 4.2, where hypoglycemia is considered as a transient symptom, a 6.75% HbA1c threshold leads to the largest expected total QALYs from birth to the first event, among all six treatment guidelines. The treatment guideline with HbA1c threshold of 7% (the current ADA’s treatment goal) results in 0.38 additional QALYs from birth to the first event for women, and 0.31 additional QALYs from birth to the first event for men, on average, compared to the observed average policy. A comparison between the treatment guideline with HbA1c threshold of 7% and the optimal policy shows that the optimal policy leads to 0.09 additional QALYs from birth to the first event for women, and 0.08 more QALYs from birth to the first event for
Table 4.2: Comparison of the expected total QALYs from birth to the first event where patients are assumed to have one QALY per year before diagnosis. Results were generated by evaluating the optimal policy, the observed average policy, and treatment guidelines with 6 different treatment goals under the base case model where hypoglycemia is considered as a transient symptom, and the alternative model where the hypoglycemia is modeled as an event.

<table>
<thead>
<tr>
<th>Hypoglycemia as a transient symptom</th>
<th>Hypoglycemia as an event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>Men</td>
</tr>
<tr>
<td>Optimal policy</td>
<td>68.50 64.46</td>
</tr>
<tr>
<td>Observed average policy</td>
<td>68.03 64.07</td>
</tr>
<tr>
<td>Threshold treatment guideline with different HbA1c thresholds</td>
<td></td>
</tr>
<tr>
<td>6.5%</td>
<td>68.41 64.37</td>
</tr>
<tr>
<td>6.75%</td>
<td>68.42 64.38</td>
</tr>
<tr>
<td>7%</td>
<td>68.41 64.37</td>
</tr>
<tr>
<td>7.25%</td>
<td>68.40 64.36</td>
</tr>
<tr>
<td>7.50%</td>
<td>68.37 64.34</td>
</tr>
<tr>
<td>7.75%</td>
<td>68.34 64.31</td>
</tr>
<tr>
<td>8%</td>
<td>68.28 64.26</td>
</tr>
</tbody>
</table>

In the alternative model, where hypoglycemia is considered as an event, among all six HbA1c threshold treatment policies, an 8% HbA1c threshold leads to the largest expected total QALYs from birth to the first event. The treatment guideline with HbA1c threshold of 7% results in 0.63 fewer QALYs from birth to the first event for women, and 0.43 fewer QALYs from birth to the first event for men, on average, compared to the observed average policy. A comparison between the treatment guideline with HbA1c threshold of 7% and the optimal policy shows that the optimal policy leads to 0.84 additional QALYs from birth to the first event for women, and 0.54 more QALYs from birth to the first event for men, on average.

4.5 Conclusions

In this chapter, we presented an MDP for optimizing the treatment decisions, including the sequence and start times to initiate hyperglycemia lowering medications, for patients with type 2 diabetes. Based on different decision maker’s opinions about hypoglycemia, which is an important
side effect associated with some hyperglycemia lowering medications such as sulfonylurea and
insulin, we provide two approaches to model it. The first approach considers hypoglycemia as a
transient symptom, and the second approach considers the hypoglycemia as an event.

From our numerical experiments, we found that the optimal policies between males and
females are similar when hypoglycemia is modeled as a transient symptom, and are different
when hypoglycemia is considered as an event. Compared with the treatment guideline where the
action of initiating medications is only determined by a patient’s HbA1c level, the optimal policy,
which considers a patient’s gender, age, HbA1c, and the medication’s effect and side effects
simultaneously, results in higher expected QALYs prior to the first event, and the magnitude of
the improvement depends on how the hypoglycemia is modeled. For example, we found that the
use of the optimal policy over the current ADA’s treatment goal can provide 0.09 additional
QALYs for a female patient and 0.08 additional QALYs for a male patient when hypoglycemia
is considered as a transient symptom; and can provide 0.84 additional QALYs for for a female
patient and 0.54 additional QALYs for for a male patient when hypoglycemia is considered as
an event.

We provided sufficient conditions under which the optimal value function is monotonic and
the optimal policy is of control-limit type in our context. From our numerical experiments, we
found that the HbA1c TPM does not satisfy the IFR property exactly but optimal policies
are nevertheless of control-limit type for both genders when hypoglycemia is modeled as a
transient symptom. However, when hypoglycemia is modeled as an event, optimal policies are
not necessarily of control-limit type.

One of the limitations in our model is that we only consider decisions prior to the first event.
A more comprehensive model could be built based on the current model to maximize expected
QALYs over a patient’s entire life. Instead of considering micro- and macro-vascular events as
absorbing states, we consider them as event states as a part of the set of health states. Then
QALYs would be calculated based on the disutility of using medications and the disutility of
being in the event state. The health state transition probabilities would be calculated based
on HbA1c state transition probabilities and event state transition probabilities. The challenges associated with this include the difficulty of estimating the probability of a secondary event (e.g. CHD or stroke) given that a first event is a risk factor. Furthermore, the implied history dependence of events results in a significant increase in the size of the state space, which could present computational challenges.
Chapter 5

Robust Optimization of Medical Treatment Decisions

Many MDP models, such as the model in Chapter 4, have been proposed to optimize medical treatment decisions (see the detailed review in Section 2.4 for additional examples). A key component of these models is the transition probability matrix (TPM), which mathematically describes the stochastic change of a patient’s health status over time. The TPM is usually estimated from observational data and are, therefore, subject to uncertainty. Such uncertainty arises from missing data, measurement errors, and estimation errors. In this chapter, we generalize the MDP model presented in Chapter 4, referred to as a Nominal MDP (NMDP) in this chapter, to a robust MDP treatment model (RMDP-TM), which considers the uncertainty in TPMs, for medical treatment decisions. Next, we analyze the model to identify properties that provide insights into the optimal policies and properties of the RMDP-TM that can be exploited to achieve computational advantages, and insights into the optimal solution. Finally, we present a case study of glycemic control to prevent micro- and macro-vascular events for patients with type 2 diabetes.
5.1 An RMDP Formulation for Medical Treatment Decisions

We formulate the medical treatment decisions as a finite-horizon, discrete-state, and finite-action RMDP with uncertain transition probabilities. Model notation is similar to Chapter 4. We assume that treatment decisions are made at a finite and discrete set of decision epochs indexed by \( t \in \mathcal{T} = \{1, 2, \ldots, T\} \). Thus, treatment decisions are revisited periodically (e.g., 3-month visits to a clinician) over a patient’s lifetime up to some reasonable upper limit on a patient’s age (e.g., age 100). The period after the last decision epoch \( T \) is the post decision horizon where patients are assumed to continue the same course of treatment as the last decision epoch for the remainder of their life.

States in our model include health states and absorbing states. The health states are defined by health and medication status, both of which influence the risk of future disease complications that the treatment aims to prevent. We define the health status by a set of metabolic states, \( \mathcal{L} \), and the medication status by a finite set of medications the patient has been prescribed, \( \mathcal{M} \). An absorbing state represents an occurrence of a major complication or death from any causes. The set of absorbing states is denoted by \( \mathcal{D} \). Thus, the complete set of states in the model is given by \( \mathcal{S} = \{\mathcal{L} \times \mathcal{M}\} \cup \mathcal{D} \). Note that although \( \mathcal{L} \) and \( \mathcal{M} \) are defined independently they are dependent, with their dependency defined through the actions that define the treatment policy.

At each decision epoch \( t \in \mathcal{T} \), the action is to determine which medication(s) to initiate, if any, at the beginning of the current decision epoch. For a patient in health state \( (\ell_t, m_t) \in \mathcal{L} \times \mathcal{M} \), the action space, \( \mathcal{A}_t(\ell_t, m_t) \) is dependent on the medications that have been initiated in the previous epoch. For a patient in absorbing state \( d \in \mathcal{D} \), there is no action, i.e., \( \mathcal{A}_t(d) = \emptyset \), \( \forall d \in \mathcal{D}, t \in \mathcal{T} \). For ease of exposition, and where it is obvious from the context, we will denote the action set by \( \mathcal{A}_t(s_t), \forall s_t \in \mathcal{S}, t \in \mathcal{T} \).

There are two types of transition probabilities in our model: probabilities of transitioning to the absorbing states, and probabilities of transitioning among health states. Given a patient is in health state \( (\ell_t, m_t) \) and takes action \( \alpha_t \in \mathcal{A}_t(\ell_t, m_t) \) at epoch \( t \), the probability of entering
the absorbing states, \( p^E_t(\ell_t, m_t, \alpha_t) \), is:

\[
p^E_t(\ell_t, m_t, \alpha_t) = \sum_{d \in D} p_t(d|\ell_t, m_t, \alpha_t) \tag{5.1}
\]

where \( p_t(d|\ell_t, m_t, \alpha_t), \forall d \in D \) represents probability of having event \( d \) occur during the current period. We denote the transition probability among metabolic states in the absence of transitioning into absorbing states during period \((t, t+1]\) by \( q_{t,\ell_t}(\ell_{t+1}), \forall t \in T, \ell_t, \ell_{t+1} \in \mathcal{L}. \) The transition probability of a patient being in state \( s_{t+1} \in \mathcal{S} \) at epoch \( t+1 \) given the patient is in state \( s_t \in \mathcal{S}, \) and takes action \( \alpha_t \in A_t(s_t) \) at epoch \( t, \) is defined as follows:

\[
p_t(s_{t+1}|s_t, \alpha_t) = \begin{cases} 
(1 - p^E_t(\ell_t, m_t, \alpha_t))q_{t,\ell_t}(\ell_{t+1}), & \text{if } s_t, s_{t+1} \in \mathcal{L} \times \mathcal{M}, \\
p_t(s_{t+1}|\ell_t, m_t, \alpha_t), & \text{if } s_t \in \mathcal{L} \times \mathcal{M}, s_{t+1} \in \mathcal{D}, \\
1, & \text{if } s_t, s_{t+1} \in \mathcal{D}, \\
0, & \text{otherwise}. 
\end{cases} \tag{5.2}
\]

The metabolic state TPM, \( \mathcal{Q}_t \equiv \{ q_{t,\ell_t}(\ell_{t+1}) \}_{\mathcal{L} \times |\mathcal{L}|}, \forall t \in T, \) is often estimated from a longitudinal data set as in Chapter 3 and 4, and other studies \([32, 33, 34, 35, 36, 37, 38]\); therefore, it is subject to uncertainty caused by sample variation. We assume that it is known to lie in a given uncertainty set \( \mathcal{Q}_t, \forall t \in T. \) We will discuss the construction and estimation of the uncertainty set, \( \mathcal{Q}_t, \) in more detail in Section 5.2 and Section 5.5.1.

The optimal value function, \( v_t(s_t), \) for a patient in state \( s_t \in \mathcal{S} \) at epoch \( t \) can be computed by solving the following optimality equations:

\[
v_t(s_t) = \max_{\pi_t \in \Pi_t} \min_{\theta_t \in \Theta_t} \mathbb{E}_{s_{t}} \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k) + \lambda^{T-t} r_T(s_T) \right\}, \quad \forall s_t \in \mathcal{L} \times \mathcal{M}, \tag{5.3}
\]

\[
0, \quad \text{otherwise},
\]

\( \forall t \in T, \) where \( \pi_t \) represents a policy from epoch \( t \) to \( T - 1, \) \( \Pi_t \) represents the set of all possible policies from epoch \( t \) to \( T - 1, \) \( r_k(s_k, \alpha_k), \forall k \in T, \) represents the one-period reward of being
in state $s_k$, given action $\alpha_k$ is taken at epoch $k$, and $\theta^*_t \triangleq (Q_t, Q_{t+1}, \ldots, Q_{T-1})$ is a stationary policy of nature which is a collection of metabolic state TPMs which are the same for all decision epochs. The set $\Theta_t \triangleq \prod_{k=t}^{T-1} Q_k$, $\forall t = 1, 2, \ldots, T - 1$, represents the set of all admissible policies of nature since epoch $t$, $\lambda \in (0, 1]$ denotes the discount factor, and $r_T(s_T)$ denotes the total expected future reward during the post decision horizon given that the patient is in state $s_T$ at epoch $T$. The optimal value function, $v_t(s_t)$, can be interpreted as the maximum expected total worst-case reward from epoch $t$ to the first major complications or death given that the patient is in state $s_t \in S$ at epoch $t$. Note that Equation (5.3) assumes a stationary policy of nature. In general this is a significant complicating factor that makes many RMDPs difficult or impossible to solve in practice. Moreover, there is no guarantee that the corresponding RMDP-optimal policy is deterministic.

An easy to solve relaxation of the RMDP-TM is obtained by relaxing the stationary requirement for nature’s policy. The optimality equations of the relaxed RMDP-TM can be written as follows:

$$\tilde{v}_t(s_t) = \begin{cases} \max_{\pi_t} \min_{\theta_t} \mathbb{E}_{s_t} \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k) + \lambda^{T-t} r_T(s_T) \right\}, & \forall s_t \in \mathcal{L} \times \mathcal{M}, \\ 0, & \text{otherwise}, \end{cases} \quad (5.4)$$

$\forall t \in \mathcal{T}$, where nature can choose any policy $\theta_t \triangleq (Q_t, Q_{t+1}, \ldots, Q_{T-1}) \in \Theta_t$, $t = 1, 2, \ldots, T - 1$. This non-stationary behavior results in the optimality equations, (5.4), being much easier to solve; however, this assumption is not realistic given that the source of uncertainty is typically attributed to differences among patients, rather than significant changes from one decision epoch to the next.

Given that there are various risk equations available in the literature that can be used to estimate the probability of having a major complication, the probabilities of transitioning into absorbing states are also subject to uncertainty. Taking this uncertainty into account, the
optimal equations (5.3) become:

\[
v'_t(s_t) = \begin{cases} 
\max_{\pi_t \in \Pi_t} \min_{(p_t^E, \theta_t) \in (\mathcal{P}_t^E, \Theta_t)} \mathbb{E}_{s_t} \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k) + \lambda^{T-t} r_T(s_T) \right\}, & \forall s_t \in \mathcal{L} \times \mathcal{M}, \\
0, & \text{otherwise},
\end{cases}
\tag{5.5}
\]

\(\forall t \in \mathcal{T}\), where \(p_t^E = (p_t^E(s_t, \alpha_t), p_{t+1}^E(s_{t+1}, \alpha_{t+1}), \ldots, p_{T-1}^E(s_{T-1}, \alpha_{T-1}))\) is a collection of the probabilities of transitioning into absorbing states from epoch \(t\) to \(T-1\), and \(\mathcal{P}_t^E = \prod_{k=t}^{T-1} \mathcal{P}_k^E(s_k, \alpha_k)\) is all possible collections of the probabilities of transitioning into absorbing states from epoch \(t\) to \(T-1\), and \(\mathcal{P}_k^E(s_k, \alpha_k), \forall k \in \mathcal{T}\), is the uncertainty set of the probabilities of transitioning into absorbing states while being in state \(s_k\), and taking action \(\alpha_k\) at epoch \(k\). Since the source of the uncertainty in the probabilities of transitioning into absorbing states and the source of the uncertainty in the probabilities of transitioning among metabolic states are different, it is reasonable to assume that the selection of \(p_t^E\) over its uncertainty set \(\mathcal{P}_t^E\) and the selection of \(\theta_t^s\) over its uncertainty set \(\Theta_t\) are independent. Under this assumption, the problem in (5.5) is equivalent to solving the RMDP-TM problem in (5.3) where the probabilities of transitioning into absorbing states are all set to be the lowest possible values in their uncertainty sets. Therefore, in this chapter, we focus on the much more difficult problem of solving the RMDP-TM problem in (5.3).

### 5.2 A Polyhedral Uncertainty Set

The specific choice of the uncertainty set \(\mathcal{Q}_t, \forall t \in \mathcal{T}\) of TPMs plays an important role in determining the computational tractability of RMDPs and the conservativeness of the robust optimal solution. The interval matrix (IM) model is one way to model the uncertainty in TPM. It describes the uncertainty on the rows of the TPM independently [45]. The IM model is common when statistical estimates of confidence intervals on the components of the TPM are available.
The uncertainty set for row $\ell_t$ of matrix $Q_t$ is written as follows:

$$Q_{IM}^{t,\ell_{t}} = \left\{ q_{t,\ell_{t}} \in \mathbb{R}_{+}^{L}: \sum_{\ell_{t+1} \in L} q_{t,\ell_{t}}(\ell_{t+1}) = 1, q_{t,\ell_{t}}(\ell_{t+1}) \leq q_{t,\ell_{t}}^{l}(\ell_{t+1}) \leq q_{t,\ell_{t}}^{u}(\ell_{t+1}), \forall \ell_{t+1} \in L \right\}, \quad (5.6)$$

$\forall \ell_t \in L$, $\forall t \in T$, where $q_{t,\ell_{t}}^{l}(\ell_{t+1})$ and $q_{t,\ell_{t}}^{u}(\ell_{t+1})$ denote the lower and upper bounds of the transition probability, $q_{t,\ell_{t}}(\ell_{t+1})$, respectively, where $0 \leq q_{t,\ell_{t}}^{l}(\ell_{t+1}), q_{t,\ell_{t}}^{u}(\ell_{t+1}) \leq 1$. Thus, the uncertainty set of the TPM, $Q_t$, can be written as follows:

$$Q_{IM}^{t} = \prod_{\ell_{t} \in L} Q_{IM}^{t,\ell_{t}}. \quad (5.7)$$

The advantage of using the IM model is that its polyhedral structure will lead to a computationally tractable relaxation of RMDP-TM [45]. However, the resulting robust optimal solution is often too conservative, since it is based on the assumption that all transition probabilities take the worst possible values simultaneously, which is an event that rarely happens in practice [85].

We combine the interval matrix model with an additional uncertainty budget constraint to fully control the size of the uncertainty set, and therefore control the conservativeness of the robust optimal solution. We refer to our model as the interval model with uncertainty budget (IMUB). We denote $\hat{q}_{t,\ell_{t}}(\ell_{t+1})$, $\forall t \in T$, $\ell_t, \ell_{t+1} \in L$, to be the mean value of the transition probability $q_{t,\ell_{t}}(\ell_{t+1})$ (It is set to be the maximum likelihood estimation (MLE) of $q_{t,\ell_{t}}(\ell_{t+1})$ in our numerical experiments). For any $\ell_t, \ell_{t+1} \in L$ such that $\hat{q}_{t,\ell_{t}}(\ell_{t+1}) \neq 0$, we define the maximal left/right-hand-side variation associated with probability $q_{t,\ell_{t}}(\ell_{t+1})$ as the absolute difference between the lower/upper bound of the transition probability, $q_{t,\ell_{t}}(\ell_{t+1})$, and its mean value, $\hat{q}_{t,\ell_{t}}(\ell_{t+1})$; and denote them by $\delta_{t,\ell_{t}}^{l}(\ell_{t+1})$ and $\delta_{t,\ell_{t}}^{u}(\ell_{t+1})$, respectively. In addition, we define the degree of left/right-hand-side variation as the proportion of variation from the lower/upper bound of $q_{t,\ell_{t}}(\ell_{t+1})$ to its mean value, and denote them by $z_{t,\ell_{t}}^{l}(\ell_{t+1})$ and $z_{t,\ell_{t}}^{u}(\ell_{t+1})$, respectively.
Therefore, for any $\ell_t \in \mathcal{L}$, $t \in \mathcal{T}$, we have the following relationships:

\[
\begin{align*}
\delta^l_{t,\ell_t} (\ell_{t+1}) &\triangleq \hat{q}_{t,\ell_t} (\ell_{t+1}) - q^l_{t,\ell_t} (\ell_{t+1}), \\
\delta^u_{t,\ell_t} (\ell_{t+1}) &\triangleq q^u_{t,\ell_t} (\ell_{t+1}) - \hat{q}_{t,\ell_t} (\ell_{t+1}), \\
 z^l_{t,\ell_t} (\ell_{t+1}) &\triangleq \frac{\hat{q}_{t,\ell_t} (\ell_{t+1}) - q^l_{t,\ell_t} (\ell_{t+1})}{\delta^l_{t,\ell_t} (\ell_{t+1})}, \quad z^u_{t,\ell_t} (\ell_{t+1}) = 0, \quad \text{if } q^l_{t,\ell_t} (\ell_{t+1}) \leq \hat{q}_{t,\ell_t} (\ell_{t+1}), \\
 z^u_{t,\ell_t} (\ell_{t+1}) &\triangleq \frac{\hat{q}_{t,\ell_t} (\ell_{t+1}) - q^u_{t,\ell_t} (\ell_{t+1})}{\delta^u_{t,\ell_t} (\ell_{t+1})}, \quad \text{if } q^u_{t,\ell_t} (\ell_{t+1}) > \hat{q}_{t,\ell_t} (\ell_{t+1}). \quad (5.8)
\end{align*}
\]

For any $\ell_t, \ell_{t+1} \in \mathcal{L}$ such that $\hat{q}_{t,\ell_t} (\ell_{t+1}) = 0$, we assume that there is no variation associated with $q_{t,\ell_t} (\ell_{t+1})$. Therefore, $q_{t,\ell_t} (\ell_{t+1}) = 0$ and $z^l_{t,\ell_t} (\ell_{t+1}) = z^u_{t,\ell_t} (\ell_{t+1}) = 0$.

The uncertainty budget on row $\ell_t$ of matrix $Q_t$ is denoted by $\Gamma_{t,\ell_t}$. It defines a limit on the total allowable variation of the probabilities, $q_{t,\ell_t} (\ell_{t+1})$, from its mean values as measured by the sum of $z^l_{t,\ell_t} (\ell_{t+1})$ and $z^u_{t,\ell_t} (\ell_{t+1})$ for all $\ell_{t+1}$. Thus $\Gamma_{t,\ell_t}$ is limited to the range from 0 to $|\mathcal{L}|$.

The complete IMUB for row $\ell_t$ of the TPM, $Q_t$, can be written as follows:

\[
\begin{align*}
Q^\text{IMUB}_{\ell_t,\ell_t} (\Gamma_{t,\ell_t}) = \left\{ q_{t,\ell_t} \in \mathbb{R}^{|\mathcal{L}|}_+ \right\} \\
&\quad \left| q_{t,\ell_t} (\ell_{t+1}) = \hat{q}_{t,\ell_t} (\ell_{t+1}) - \delta^l_{t,\ell_t} (\ell_{t+1}) z^l_{t,\ell_t} (\ell_{t+1}) \right. \\
&\quad \left. + \delta^u_{t,\ell_t} (\ell_{t+1}) z^u_{t,\ell_t} (\ell_{t+1}), \quad \forall \ell_{t+1} \in \mathcal{L}, \right. \\
&\quad z^l_{t,\ell_t} (\ell_{t+1}) \cdot z^u_{t,\ell_t} (\ell_{t+1}) = 0, \quad \forall \ell_{t+1} \in \mathcal{L}, \\
&\quad 0 \leq z^l_{t,\ell_t} (\ell_{t+1}), z^u_{t,\ell_t} (\ell_{t+1}) \leq 1, \quad \forall \ell_{t+1} \in \mathcal{L}, \\
&\quad 0 \leq q_{t,\ell_t} (\ell_{t+1}) \leq 1, \quad \forall \ell_{t+1} \in \mathcal{L}, \\
&\quad \sum_{\ell_{t+1} \in \mathcal{L}} q_{t,\ell_t} (\ell_{t+1}) = 1, \\
&\quad \sum_{\ell_{t+1} \in \mathcal{L}} \left[ z^l_{t,\ell_t} (\ell_{t+1}) + z^u_{t,\ell_t} (\ell_{t+1}) \right] \leq \Gamma_{t,\ell_t}. \quad (5.9)
\end{align*}
\]

The first equation in (5.9) represents the variable $q_{t,\ell_t} (\ell_{t+1})$ in terms of its mean value, and the degree of left-hand-side and right-hand-side variation. Since $q_{t,\ell_t} (\ell_{t+1})$ can only be on one side of its mean value, $z^l_{t,\ell_t} (\ell_{t+1})$ and $z^u_{t,\ell_t} (\ell_{t+1})$ cannot be positive simultaneously. Therefore, the second constraint guarantees that either $z^l_{t,\ell_t} (\ell_{t+1})$ or $z^u_{t,\ell_t} (\ell_{t+1})$ or both are zero. The third constraint provides the lower and upper bounds for the variables $z^l_{t,\ell_t} (\ell_{t+1})$ and $z^u_{t,\ell_t} (\ell_{t+1})$. The fourth and fifth constraints guarantee that the variables, $q_{t,\ell_t} (\ell_{t+1})$, $\forall \ell_{t+1} \in \mathcal{L}$, are confined to
the probability simplex. The last constraint is the uncertainty budget constraint, which requires
the total degree of uncertainty on the row \( \ell_t \) of matrix \( Q_t \) to be less than or equal to \( \Gamma_{t,\ell_t} \). The
uncertainty set of the TPM, \( Q_t, \forall t \in \mathcal{T} \), can be written as follows:

\[
Q_t^{\text{IMUB}}(\Gamma_{t,\ell_t}) = \prod_{\ell_t \in \mathcal{L}} Q_{t,\ell_t}^{\text{IMUB}}(\Gamma_{t,\ell_t}), \forall t \in \mathcal{T}.
\] (5.10)

**Remark 1:** Notice that when \( \Gamma_{t,\ell_t} = 0 \), \( \forall \ell_t \in \mathcal{L}, t \in \mathcal{T} \), we have \( q_{t,\ell_t}(\ell_{t+1}) = \hat{q}_{t,\ell_t}(\ell_{t+1}) \),
\( \forall \ell_t, \ell_{t+1} \in \mathcal{L}, t \in \mathcal{T} \), and the RMDP is equivalent to the NMDP. Likewise, when \( \Gamma_{t,\ell_t} = |\mathcal{L}| \),
\( \forall \ell_t \in \mathcal{L}, t \in \mathcal{T} \), the uncertainty set is the same as IM model. Therefore, \( \Gamma_{t,\ell_t} \) controls the
conservativeness of the RMDP policy.

There is a large body of literature on robust optimization typically in the context of robust
counterparts of linear programs and stochastic programs [61]. Our model has some similarities
to previously proposed budgeted uncertainty models. The budget constraint in the cardinality
constrained model proposed by Bertsimas (2004) [65] requires that the number of coefficients,
which can actually change in the interval, be less than or equal to the budget. Similarly, the
budget constraint in the “Lighting does not strike twice” model proposed by Mannor (2012) [85]
required that in the TPM of an RMDP model the number of states whose parameters deviate from
their mean values be bounded by the given budget. Although their models have the capability
of controlling the conservativeness of the robust solution; however, it is hard to interpret the
meaning of the budget parameter in practice. In our model on the other hand, it is the total
degree of variation from the mean value that is limited by the budget; therefore, intuitively, the
budget parameter provides a measurement of how far away the uncertain parameter could be
away from the mean value. Our model is most closely related to the D-Norm model proposed by
Bertsimas (2011) [61]; however, our model has an additional probability constraint, namely, the
sum of the probabilities, \( \{q_{t,\ell_t}(\ell_{t+1})\}_{\ell_{t+1} \in \mathcal{L}} \), has to be equal to 1. As we will show this constraint
significantly increases the difficulty in solving the inner problem.
5.3 Bi-criteria RMDP-TM

Another way to control the conservativeness of the robust optimal solution is to consider the objectives of the NMDP and the RMDP-TM simultaneously, which we refer to as a bi-criteria RMDP-TM (BCRMDP-TM). Based on the weighted sum method [127], the optimal value functions of the BCRMDP-TM can be written as follows:

\[
\begin{align*}
\nu_t^{BCRMDP-TM}(s_t) &= \max_{\pi_t \in \Pi_t} \left\{ \beta \mathbb{E}_{\hat{\theta}_t} \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k) + \lambda^{T-t} r_{T}(s_T) \right\} \right. \\
& \quad + \left(1 - \beta \right) \min_{\theta_t \in \Theta_t} \mathbb{E}_{\theta_t} \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k) + \lambda^{T-t} r_{T}(s_T) \right\} \right\}, 
\end{align*}
\]

for \( s_t \in \mathcal{L} \times \mathcal{M} \), \( t \in \mathcal{T} \), where \( \beta \in [0, 1] \) weights the two competing criteria, and \( \hat{\theta}_t = (\hat{Q}_t, \hat{Q}_{t+1}, \ldots, \hat{Q}_{T-1}) \), \( t = 1, 2, \ldots, T - 1 \), is a collection of MLEs of the metabolic state TPMs.

The BCRMDP-TM is difficult or impossible to solve in practice. Similar to the relaxed RMDP-TM shown in Equation 5.4, a relaxed BCRMDP can obtained by relaxing the stationary requirement for nature’s policy. The optimal value functions of the relaxed BCRMDP-TM can be written as follows:

\[
\tilde{v}_t^{BCRMDP-TM}(s_t) = \max_{\pi_t \in \Pi_t} \left\{ \beta \mathbb{E}_{\hat{\theta}_t} \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k) + \lambda^{T-t} r_{T}(s_T) \right\} \right. \\
& \quad + \left(1 - \beta \right) \min_{\theta_t \in \Theta_t} \mathbb{E}_{\theta_t} \left\{ \sum_{k=t}^{T-1} \lambda^{k-t} r_k(s_k, \alpha_k) + \lambda^{T-t} r_{T}(s_T) \right\} \right\},
\]

for \( s_t \in \mathcal{L} \times \mathcal{M} \), \( t \in \mathcal{T} \), where \( \beta \in [0, 1] \) weights the two competing criteria, and \( \hat{\theta}_t = (\hat{Q}_t, \hat{Q}_{t+1}, \ldots, \hat{Q}_{T-1}) \), \( t = 1, 2, \ldots, T - 1 \), is a collection of MLEs of the metabolic state TPMs.

The following Proposition 5.3.1 provides sufficient conditions under which the relaxed BCRMDP-TM is equivalent to an relaxed RMDP-TM.
Proposition 5.3.1. If the uncertainty set, $Q_t$, of the TPM, $Q_t$, $\forall t \in T$ satisfies the state rectangular property [45], then

$$\hat{v}_t^{BCRMDP-TM}(s_t) = \max_{\pi_t \in \Pi_t} \left\{ r_t(s_t, \alpha_t) + (1 - p_t^E(\ell_t, m_t, \alpha_t)) \sum_{l_{t+1} \in \mathcal{L}} q_{t, l_t}(\ell_{t+1}) \hat{v}_{t+1}^{BCRMDP-TM}(s_{t+1}) \beta \hat{q}_{t, l_t}(\ell_{t+1}) \right\}$$

(5.13)

Proof. When the uncertainty set $Q_t$ satisfies the state rectangular property [45], the relaxed BCRMDP-TM can be written into the following recursive form:

$$\hat{v}_t^{BCRMDP-TM}(s_t) = \max_{\pi_t \in \Pi_t} \left\{ r_t(s_t, \alpha_t) + (1 - p_t^E(\ell_t, m_t, \alpha_t)) \lambda \sum_{l_{t+1} \in \mathcal{L}} q_{t, l_t}(\ell_{t+1}) \hat{v}_{t+1}^{BCRMDP-TM}(s_{t+1}) \right\}$$

$$+ (1 - \beta) \left[ r_t(s_t, \alpha_t) + (1 - p_t^E(\ell_t, m_t, \alpha_t)) \lambda \min_{q_{t, l_t} \in Q_t} \sum_{l_{t+1} \in \mathcal{L}} q_{t, l_t}(\ell_{t+1}) \hat{v}_{t+1}^{BCRMDP-TM}(s_{t+1}) \right]$$

$$= \max_{\pi_t \in \Pi_t} \left\{ r_t(s_t, \alpha_t) + (1 - p_t^E(\ell_t, m_t, \alpha_t)) \lambda \sum_{l_{t+1} \in \mathcal{L}} q_{t, l_t}(\ell_{t+1}) \hat{v}_{t+1}^{BCRMDP-TM}(s_{t+1}) \right\}$$

This completes the proof. \(\square\)

Since the relaxation of BCRMDP-TM is equivalent to the relaxation of RMDP-TM when the uncertainty set satisfies the state rectangular property, we focus our the state rectangular property and the solution methods for solving the relaxed RMDP-TM in the next section.
5.4 Analysis of the RMDP-TM

In this section, we analyze the properties and solution methods for RMDP-TM and relaxed RMDP-TM assuming that the uncertainty set of the TPM is defined by the IMUB shown in Equation (5.10).

The definition of the uncertainty set of the TPM, \( \mathcal{Q}^{\text{IMUB}}_t(\Gamma_t, \ell_t) \), \( \forall t \in T \), is a cartesian product of \( \mathcal{Q}^{\text{IMUB}}_{t, \ell_t}(\Gamma_t, \ell_t) \), therefore, it satisfies the state rectangular property [83]. The state rectangular property is a form of an independence property. It can be interpreted as follows: the choice of particular distribution \( q_{t, \ell_t} \in \mathcal{Q}^{\text{IMUB}}_{t, \ell_t}(\Gamma_t, \ell_t) \) when the system is in state \( \ell_t \) does not limit the choice of the nature’s decision \( q_{t, \ell_t'} \in \mathcal{Q}^{\text{IMUB}}_{t, \ell_t'}(\Gamma_t, \ell_t') \), when the system is in state \( \ell_t', \forall \ell_t' \in \mathcal{L} \). The nature’s decision based on our definition may or may not have the IFR property. If we force the nature’s decision to have the IFR property, then the state rectangular property is violated because it imposes a correlation among states into the definition of the uncertainty set.

Based on Theorem 1 presented by Nilim and El Ghaoui (2005) in [45] and also Theorem 2.2 presented by Iyengar (2005) [83], the optimality equations of the relaxed RMDP-TM shown in (5.4) can be written into the following recursive form:

\[
\tilde{v}(s_t) = \begin{cases} 
\max_{\alpha_t \in A_t(\ell_t, m_t)} \left\{ r_t(\ell_t, m_t, \alpha_t) + (1 - p_t^E(\ell_t, m_t, \alpha_t)) \lambda \min_{q_{t, \ell_t} \in \mathcal{Q}^{\text{IMUB}}_{t, \ell_t}} \sum_{\ell_t+1 \in \mathcal{L}} q_{t, \ell_t}(\ell_t+1) \tilde{v}_{t+1}(\ell_t+1, m_{t+1}) \right\}, \\
\forall s_t = (\ell_t, m_t) \in \mathcal{L} \times \mathcal{M}, t \in T \setminus \{T\}, \ \ \ \ \ \ \ \ \ \ \ \ r_T(s_T), \ \ \ \ \ \ \ \ \ \ \ \ 0, \end{cases}
\]

(5.14)

The minimization problem presented in Equation (5.14) is often referred to as the inner problem:

\[
\sigma_t(\ell_t, m_t, \alpha_t, \Gamma_t, \ell_t) = \min_{q_{t, \ell_t} \in \mathcal{Q}^{\text{IMUB}}_{t, \ell_t}} \sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_t}(\ell_{t+1}) \tilde{v}_{t+1}(\ell_{t+1}, m_{t+1})
\]

(5.15)

\forall t \in T, (\ell_t, m_t) \in \mathcal{L} \times \mathcal{M}, \alpha_t \in A_t(\ell_t, m_t). \text{ Notice that the medication state } m_t \text{ and the action } \alpha_t \text{ determine the medication state, } m_{t+1}, \text{ at epoch } t + 1. \text{ Therefore, we also denote the inner problem } \sigma_t(\ell_t, m_t, \alpha_t, \Gamma_t, \ell_t) \text{ by } \sigma_t(\ell_t, m_{t+1}, \Gamma, \ell_t) \text{ for simplicity.
Algorithm 2 An exact method to solve the relaxed RMDP-TM

1: Step 1. Initialize the value function at stage $T$:

$$v_T(\ell_T, m_T) \leftarrow R_T(\ell_T, m_T), \forall \ell_T \in \mathcal{L}, m_T \in \mathcal{M}$$

2: Step 2. Recursion
3: for $t = T - 1 \rightarrow 0$ do
4:   for $\ell_t \in \mathcal{L}$ do
5:     for $m_t \in \mathcal{M}$ do
6:       for $\alpha_t \in \mathcal{A}_t(\ell_t, m_t)$ do
7:         Solve the inner problem to get $\sigma_t^*(\ell_t, m_t, \alpha_t, \Gamma_{t,\ell_t})$
8:         Calculate $w_t(\ell_t, m_t, \alpha_t)$:

$$w_t(\ell_t, m_t, \alpha_t) = r_t(\ell_t, m_t, \alpha_t) + \lambda[1 - p_t^E(\ell_t, m_t, \alpha_t)]\sigma_t^*(\ell_t, m_t, \alpha_t, \Gamma_{t,\ell_t})$$
9:       end for
10:  Update Value Function:
11:    $$v_t(\ell_t, m_t) \leftarrow \max_{\alpha_t \in \mathcal{A}_t(\ell_t, m_t)} \{w_t(\ell_t, m_t, \alpha_t)\}$$
12:  Update Optimal Action Set:
13:    $$\mathcal{A}_t^*(\ell_t, m_t) \leftarrow \arg\max_{\alpha_t \in \mathcal{A}_t(\ell_t, m_t)} \{w_t(\ell_t, m_t, \alpha_t)\}$$
14: end for

end for
The relaxed RMDP-TM, assuming the uncertainty set defined by IMUB, can be solved by the following backward induction-based robust dynamic programming Algorithm 2 presented in Nilim and El Ghaoui [45]. For each decision epoch, each state, and each action, the inner problem is solved to get the expected worst-case future reward. From the Algorithm 2 we can see that the inner problem must be solved $|S| \cdot |T\backslash\{T\}|$ times. Therefore, the computational effort of solving the relaxed RMDP-TM depends on how efficiently the inner problem can be solved.

The inner problem (5.15) is a nonlinear program (NLP) which can be written as follows:

$$
\min_{\ell, t} d_{t}^{\text{IMUB-NLP}}(\ell_t, m_t, \alpha_t, \Gamma_t, \ell_{t+1}) = \sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_{t+1}}(\ell_{t+1}) v_{t+1}(\ell_{t+1}, m_{t+1})
$$

s.t.  
$$
q_{t, \ell_{t+1}}(\ell_{t+1}) = \hat{q}_{t, \ell}(\ell_{t+1}) - \delta_{t, \ell}(\ell_{t+1}) z^l_{t, \ell}(\ell_{t+1}) + \delta_{t, \ell}(\ell_{t+1}) z^u_{t, \ell}(\ell_{t+1}), \forall \ell_{t+1} \in \mathcal{L},
$$

(IMUB-NLP)

$$
\sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_{t+1}}(\ell_{t+1}) = 1,
$$

$$
\sum_{\ell_{t+1} \in \mathcal{L}} [z^l_{t, \ell_{t+1}}(\ell_{t+1}) + z^u_{t, \ell_{t+1}}(\ell_{t+1})] \leq \Gamma_{t, \ell},
$$

$$
z^l_{t, \ell_{t+1}}(\ell_{t+1}) \cdot z^u_{t, \ell_{t+1}}(\ell_{t+1}) = 0, \forall \ell_{t+1} \in \mathcal{L},
$$

$$
0 \leq z^l_{t, \ell_{t+1}}(\ell_{t+1}), z^u_{t, \ell_{t+1}}(\ell_{t+1}) \leq 1, \forall \ell_{t+1} \in \mathcal{L}.
$$

The following Proposition 5.4.1 provides a linear reformulation of the inner problem that can be exploited to achieve significant computational advantages in computing the RMDP-optimal policy of the relaxed RMDP-TM (5.14). As we will shown in our numerical experiments that the LP reformulation of the inner problem presented in Proposition 5.4.1 can significantly reduce the computational time of solving the relaxed RMDP-TM with uncertainty set defined by IMUB.
Proposition 5.4.1. The IMUB-NLP has the same optimal objective function value as the following linear program (LP):

\[
\begin{align*}
\min_{\ell_{t+1} \in \mathcal{L}} & \quad \sigma_{t}^{IMUB-LP}(\ell_{t}, m_{t}, \alpha_{t}, \Gamma_{t, \ell}) = \sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_{t}}(\ell_{t+1}) v_{t+1}(\ell_{t+1}, m_{t+1}) \\
\text{s.t.} & \quad \sum_{\ell_{t+1} \in \mathcal{L}} q_{t, \ell_{t}}(\ell_{t+1}) = 1, \quad (5.21) \\
& \quad \sum_{\ell_{t+1} \in \mathcal{L}} [x_{t, \ell_{t}}^{l}(\ell_{t+1}) + x_{t, \ell_{t}}^{u}(\ell_{t+1})] \leq \Gamma_{t, \ell_{t}}, \quad (5.22) \\
& \quad x_{t, \ell_{t}}^{u}(\ell_{t+1}) \geq \frac{q_{t, \ell_{t}}(\ell_{t+1}) - \hat{q}_{t, \ell_{t}}(\ell_{t+1})}{\delta_{t, \ell_{t}}^{u}(\ell_{t+1})}, \forall \ell_{t+1} \in \mathcal{L}, \quad (5.23) \\
& \quad x_{t, \ell_{t}}^{l}(\ell_{t+1}) \geq \frac{\hat{q}_{t, \ell_{t}}(\ell_{t+1}) - q_{t, \ell_{t}}(\ell_{t+1})}{\delta_{t, \ell_{t}}^{l}(\ell_{t+1})}, \forall \ell_{t+1} \in \mathcal{L}, \quad (5.24) \\
& \quad 0 \leq x_{t, \ell_{t}}^{l}(\ell_{t+1}), x_{t, \ell_{t}}^{u}(\ell_{t+1}) \leq 1, \forall \ell_{t+1} \in \mathcal{L}. \quad (5.25)
\end{align*}
\]

Proof. First, we prove that if \((z_{t, \ell_{t}}^{l}(\ell_{t+1}), z_{t, \ell_{t}}^{u}(\ell_{t+1}), q_{t, \ell_{t}}^{*}(\ell_{t+1}))\) is an optimal solution to the (IMUB-NLP), then we can construct a feasible solution, \((x_{t, \ell_{t}}^{l}(\ell_{t+1}), x_{t, \ell_{t}}^{u}(\ell_{t+1}), q_{t, \ell_{t}}^{*}(\ell_{t+1}))\), to the (IMUB-LP) by setting \(x_{t, \ell_{t}}^{l}(\ell_{t+1}) = z_{t, \ell_{t}}^{l}(\ell_{t+1}), x_{t, \ell_{t}}^{u}(\ell_{t+1}) = z_{t, \ell_{t}}^{u}(\ell_{t+1})\) and \(q_{t, \ell_{t}}^{*}(\ell_{t+1}) = q_{t, \ell_{t}}^{*}(\ell_{t+1})\). Based on constraints (5.16) and (5.19) in the (IMUB-NLP), we have the following relationship. If \(q_{t, \ell_{t}}^{*}(\ell_{t+1}) \leq \hat{q}_{t, \ell_{t}}(\ell_{t+1})\), then

\[
\begin{align*}
x_{t, \ell_{t}}^{l}(\ell_{t+1}) & = z_{t, \ell_{t}}^{l}(\ell_{t+1}) \\
& = \frac{\hat{q}_{t, \ell_{t}}(\ell_{t+1}) - q_{t, \ell_{t}}^{*}(\ell_{t+1})}{\delta_{t, \ell_{t}}^{l}(\ell_{t+1})} \\
& \geq 0, \quad (5.26) \\
x_{t, \ell_{t}}^{u}(\ell_{t+1}) & = z_{t, \ell_{t}}^{u}(\ell_{t+1}) = 0.
\end{align*}
\]
If \( q_{t,\ell}^{*}(t+1) > \hat{q}_{t,\ell}(t+1) \) then
\[
x_{t,\ell}^{u,t}(t+1) = z_{t,\ell}^{u,t}(t+1) \\
= q_{t,\ell}^{*}(t+1) - \hat{q}_{t,\ell}(t+1) \\
= \frac{\delta_{t,\ell}(t+1)}{\delta_{t,\ell}(t+1)} \geq 0, \\
x_{t,\ell}^{l,t}(t+1) = z_{t,\ell}^{l,t}(t+1) = 0.
\] (5.27)

It follows that (5.26, 5.27) imply constraints (5.23) and (5.24) are satisfied. Satisfaction of constraints (5.21), (5.22), and (5.25) is implied by the constraints (5.17), (5.18), and (5.20) in the (IMUB-NLP), respectively. Therefore, \((x_{t,\ell}^{l,t}(t+1), x_{t,\ell}^{u,t}(t+1), q_{t,\ell}^{l,t}(t+1))\) is a feasible solution to the (IMUB-LP). Since the (IMUB-NLP) and (IMUB-LP) have the same objective function, it follows that:
\[
\sigma_{t}^{\text{IMUB-NLP},*}(\ell, m_{t}, \alpha_{t}, \Gamma_{t,\ell}) \geq \sigma_{t}^{\text{IMUB-LP},*}(\ell, m_{t}, \alpha_{t}, \Gamma_{t,\ell})
\] (5.28)

Next we prove that if \((x_{t,\ell}^{l,t}(t+1), x_{t,\ell}^{u,t}(t+1), q_{t,\ell}^{l,t}(t+1))\) is the optimal solution to the (IMUB-LP), then we can construct a feasible solution \((x_{t,\ell}^{l,t}(t+1), z_{t,\ell}^{u,t}(t+1), q_{t,\ell}^{l,t}(t+1))\) to the (IMUB-NLP). For any \( \ell+1 \in \mathcal{L} \), we define \( z_{t,\ell}^{l,t}(t+1), z_{t,\ell}^{u,t}(t+1) \) and \( q_{t,\ell}^{l,t}(t+1) \) as follows:
\[
q_{t,\ell}^{l,t}(t+1) = q_{t,\ell}^{*}(t+1), \\
\begin{cases}
q_{t,\ell}^{u,t}(t+1) = \frac{q_{t,\ell}^{*}(t+1) - q_{t,\ell}^{l,t}(t+1)}{\delta_{t,\ell}(t+1)} \geq 0, & \text{if } q_{t,\ell}^{l,t}(t+1) \leq \hat{q}_{t,\ell}(t+1), \\
z_{t,\ell}^{l,t}(t+1) = 0, \\
z_{t,\ell}^{u,t}(t+1) = \frac{q_{t,\ell}^{*}(t+1) - q_{t,\ell}^{l,t}(t+1)}{\delta_{t,\ell}(t+1)} \geq 0, & \text{if } q_{t,\ell}^{l,t}(t+1) > \hat{q}_{t,\ell}(t+1). \\
z_{t,\ell}^{l,t}(t+1) = 0.
\end{cases}
\] (5.29)

It follows that (5.29) implies that constraints (5.16) and (5.19) are satisfied. Based on the constraint (5.21) in the (IMUB-LP), we have \( \sum_{t+1 \in \mathcal{L}} q_{t,\ell}^{*}(t+1) = 1 \). Therefore, (5.29) implies that \( \sum_{t+1 \in \mathcal{L}} q_{t,\ell}^{l,t}(t+1) = 1 \), namely, the constraint (5.17) is satisfied. Based on constraints (5.23) and
(5.24) in the (IMUB-LP) and (5.29), we know that

\[ x_{t,\ell_t}(\ell_{t+1}) \geq \frac{\hat{q}_{t,\ell_t}(\ell_{t+1}) - q_0^*(\ell_{t+1})}{\delta_{t,\ell_t}(\ell_{t+1})} = z_{t,\ell_t}(\ell_{t+1}), \tag{5.30} \]

and

\[ x_{t,\ell_t}(\ell_{t+1}) \geq \frac{q_0^*(\ell_{t+1}) - \hat{q}_{t,\ell_t}(\ell_{t+1})}{\delta_{t,\ell_t}(\ell_{t+1})} = z_{t,\ell_t}(\ell_{t+1}). \tag{5.31} \]

Therefore,

\[ \sum_{\ell_{t+1}} [z_{t,\ell_t}(\ell_{t+1}) + z_{t,\ell_t}(\ell_{t+1})] \leq \sum_{\ell_{t+1}} [x_{t,\ell_t}(\ell_{t+1}) + x_{t,\ell_t}(\ell_{t+1})] \leq \Gamma_{t,\ell_t}, \tag{5.32} \]

\[ z_{t,\ell_t}(\ell_{t+1}) \leq x_{t,\ell_t}(\ell_{t+1}) \leq 1, \tag{5.33} \]

and

\[ z_{t,\ell_t}(\ell_{t+1}) \leq x_{t,\ell_t}(\ell_{t+1}) \leq 1. \tag{5.34} \]

Inequalities (5.32), (5.33), and (5.34) imply constraints (5.18) and (5.20) in the (IMUB-NLP) are satisfied. Thus, \((z_{t,\ell_t}(\ell_{t+1}), z_{t,\ell_t}(\ell_{t+1}), q_{t,\ell_t}(\ell_{t+1}))\) is a feasible solution to the (IMUB-NLP).

Since the (IMUB-NLP) and (IMUB-LP) have the same objective function, it follows that:

\[ \sigma_{t}^{\text{IMUB-NLP}}(\ell, m_t, \alpha_t \Gamma_{t,\ell_t}) \leq \sigma_{t}^{\text{IMUB-LP}}(\ell, m_t, \alpha_t, \Gamma_{t,\ell_t}). \tag{5.35} \]

Together (5.28) and (5.35) imply

\[ \sigma_{t}^{\text{IMUB-NLP}}(\ell, m_t, \alpha_t, \Gamma_{t,\ell_t}) = \sigma_{t}^{\text{IMUB-LP}}(\ell, m_t, \alpha_t, \Gamma_{t,\ell_t}), \]

and the proof is complete.
Next, we show that when the uncertainty set is defined without an uncertainty budget constraint, the inner problem (5.15) of the relaxed RMDP-TM can be solved by a fast algorithm with complexity of $O(|L|^2)$. First, we rewrite the inner problem (5.15) in the following general form:

$$\min \psi(c, y^l, y^u) = \sum_{i=1}^{n} c_i y_i$$ \hspace{1cm} (5.36)

s.t. \hspace{1cm} 

$$y_i \geq y_i^l, \hspace{1cm} i = 1, 2, \ldots, n,$$ \hspace{1cm} (5.37)

$$y_i \leq y_i^u, \hspace{1cm} i = 1, 2, \ldots, n,$$ \hspace{1cm} (5.38)

$$\sum_{i=1}^{n} y_i = 1,$$ \hspace{1cm} (5.39)

$$y_i \geq 0, \hspace{1cm} i = 1, 2, \ldots, n.$$ \hspace{1cm} (5.40)

where $y_i, i = 1, 2, \ldots, n$, are the decision variables, $c = (c_1, c_2, \ldots, c_n)$, where $c_i \geq 0, i = 1, 2, \ldots, n$, is the vector of objective function coefficients, $y^l = (y_1^l, y_2^l, \ldots, y_n^l)$ and $y^u = (y_1^u, y_2^u, \ldots, y_n^u)$ are vectors of lower and upper bounds of the decision variables, $y_i, i = 1, 2, \ldots, n$, respectively. We assume that $0 \leq y_i^l < y_i^u$ without loss of generality.

**Proposition 5.4.2.** If (LP.1) has at least one feasible solution, then Algorithm (3) generates an optimal solution, $(y_1^*, \ldots, y_n^*)$, to (LP.1).
Algorithm 3 A fast algorithm to solve (LP.1)

1: Step 1. Sort the coefficients of the objective function in nonincreasing order, and relabel index $i$ according to this order.
2: Step 2. For all $i = 1, \ldots, n$, initialize $y^l_i$ and $y^u_i$
3: Step 3. Find an optimal solution
4: for $\tau = 1 \rightarrow n$ do
5: \hspace{1em} $I_1 \leftarrow \{i \in \{1, \ldots, n\} : c_i > c_\tau\}$
6: \hspace{1em} $I_2 \leftarrow \{i \in \{1, \ldots, n\} : c_i \leq c_\tau\}$
7: \hspace{1em} if $1 - [\sum_{j \in I_1} y^l_j + \sum_{k \in I_2} y^u_k] \in [y^l_\tau, y^u_\tau]$ then
8: \hspace{2em} $y^l_j \leftarrow y^l_j, \quad \forall j \in I_1$
9: \hspace{2em} $y^\tau_\tau \leftarrow 1 - [\sum_{j=1}^{\tau-1} y^l_j + \sum_{k=\tau+1}^{n} y^u_k]$
10: \hspace{2em} $y^u_k \leftarrow y^u_k, \quad \forall k \in I_2$
11: \hspace{1em} else
12: \hspace{2em} Next $\tau$
13: \hspace{1em} end if
14: end for

Proof. The assumption that (LP.1) has at least one feasible solution implies that the feasible region of (LP.1) is nonempty. It is also bounded since $0 \leq y_i \leq 1, \forall i = 1, \cdots, n$. Since the coefficients of the objective function are nonnegative, the value of the objective function is also bounded. Therefore, based on Corollary 2.3 in Bertsimas [128], we know that there exists at least one basic feasible solution which is an optimal solution to (LP.1).

Now we analyze the structure of the optimal solution to (LP.1). The standard form of (LP.1) can be written as follows:
\[
\begin{align*}
\text{min} & \quad \sum_{i=1}^{n} c_i y_i \\
\text{s.t.} & \quad y_i - a_i = y_i^l, \quad i = 1, 2, \ldots, n, \quad (5.42) \\
& \quad y_i + b_i = y_i^u, \quad i = 1, 2, \ldots, n, \quad (5.43) \\
& \quad \sum_{i=1}^{n} y_i = 1, \quad (5.44) \\
& \quad y_i, a_i, b_i \geq 0, \quad i = 1, 2, \ldots, n. \quad (5.45)
\end{align*}
\]

This formulation includes \(3n\) decision variables and \(2n + 1\) constraints. Thus, a basic feasible solution has \(n - 1\) non-basic variables. For any \(i \in \{1, \ldots, n\}\), consider the following eight cases for a basic feasible solution \((y'_1, \ldots, y'_n, a'_1, \ldots, a'_n, b'_1, \ldots, b'_n)\):

1: \(y_i, a_i, b_i\) are all basic variables. In this case, \(y'_i\) is in the range \([y_i^l, y_i^u]\).

2: \(a_i\) is a non-basic variable, and \(y_i, b_i\) are basic variables. In this case, \(y'_i = y_i^l\).

3: \(b_i\) is a non-basic variable, and \(y_i, a_i\) are basic variables. In this case, \(y'_i = y_i^u\).

4: \(y_i\) is a non-basic variable, and \(a_i, b_i\) are basic variables. This is possible if and only if \(y_i^l = 0\). In this case, \(y_i^*\) is at its lower bound 0.

5: \(a_i, y_i\) are both non-basic variables, and \(b_i\) is a basic variable. In this case, the basis matrix include a row of zeros and this case does not correspond to a basic feasible solution.

6: \(a_i, b_i\) are both non-basic variables, and \(y_i\) is a basic variable. This case corresponds to an infeasible solution since \(y_i\) cannot be simultaneously equal to \(y_i^l\) and \(y_i^u\) by assumption.

7: \(y_i, b_i\) are both non-basic variables, \(a_i\) is basic variable. This case corresponds to an infeasible solution since the constraint \(y_i + b_i = y_i^u > 0\) will be violated.

8: \(y_i, a_i, b_i\) are all non-basic variables. This is infeasible since the constraint \(y_i + b_i = y_i^u > 0\) will be violated.
Therefore, cases (1–4) are the only possible cases. For any $i \in \{1, \ldots, n\}$, among the three decision variables $y_i$, $a_i$, and $b_i$, at most one can be a non-basic variable. Since there are $n - 1$ non-basic variables, there is at most one $k$ that is of case 1, and all others are of case 2, 3, or 4. Therefore, there exists one optimal solution of (LP.1), which is a basic feasible solution, is of the following structure: at most one $y_i \in (y^l_i, y^u_i)$, and all others are either at the lower bound or upper bound.

For basic feasible solution $(y'_1, y'_2, \ldots, y'_n)$, assume variable $y'_r$ corresponding to the case 1. If there is a variable $y'_k$ such that $y'_k = y^l_k$ and $c_k < c_r$, then we can pivot to a better basic feasible solution via the following pivot scheme A.

1: If $y'_r - y'_r \leq y^u_k - y^l_k$, then

$$y'_r \leftarrow y'_r, \quad y'_k \leftarrow y'_k + (y'_r - y'_r), \quad \tau \leftarrow k.$$

2: If $y'_r - y'_r > y^u_k - y^l_k$, then

$$y'_r \leftarrow y'_r - (y^u_k - y^l_k), \quad y'_k \leftarrow y^u_k.$$

On the other hand, if there is a $k$ such that $y'_k = y^u_k$ and $c_k > c_r$, then we can pivot to a better basic feasible solution via the following pivot scheme B.

1: If $y^u_r - y'_r \leq y^u_k - y^l_k$, then

$$y'_r \leftarrow y^u_r, \quad y'_k \leftarrow y^u_k - (y^u_r - y'_r), \quad \tau \leftarrow k.$$

2: If $y^u_r - y'_r > y^u_k - y^l_k$, then

$$y'_r \leftarrow y'_r + (y^u_k - y^l_k), \quad y'_k \leftarrow y^l_k.$$
Therefore, there exists a basic feasible solution with basic variable $y_\tau$ such that

$$y_\tau = 1 - \sum_{j \in I_1} y_j^l - \sum_{k \in I_2} y_k^u \in [y_\tau^l, y_\tau^u].$$  \tag{5.46}$$

where $I_1 = \{i : c_i > c_\tau\}$, and $I_2 = \{i : c_i \leq c_\tau\}$. Thus the existence of the index $\tau$ in the Algorithm 3 is guaranteed.

Next, we prove that the solution, $(y_1^*, \ldots, y_n^*)$, generated by Algorithm 3 is an optimal solution of (LP.1). The feasibility of $(y_1^*, \ldots, y_n^*)$ is guaranteed by the Algorithm 3. To prove that is optimal, we first consider the dual problem of (LP.1):

$$\begin{align*}
\text{max} & \quad \sum_{i=1}^n (\mu_i y_i^l - \rho_i y_i^u) + w \\
\text{s.t.} & \quad \mu_i - \rho_i + w \leq c_i, \quad i = 1, 2, \ldots, n, \\
\text{(Dual of LP.1)} & \quad \mu_i, \rho_i \geq 0 \quad i = 1, 2, \ldots, n, \\
& \quad w \text{ u.r.s.}
\end{align*}$$

Based on the primal feasible solution, $(y_1^*, \ldots, y_n^*)$, we can construct a dual solution as follows:

$$\begin{align*}
w^* & = c_\tau, \\
\mu_i^* & = c_i - c_\tau, \quad \text{if } i \in I_1, \\
\rho_i^* & = 0, \\
\mu_\tau^* & = 0, \\
\rho_\tau^* & = 0, \\
\mu_i^* & = 0, \quad \text{if } i \in I_2, \\
\rho_i^* & = c_\tau - c_i.
\end{align*}$$  \tag{5.51}$$

The dual solution satisfies the constraints (5.48), and the nonnegativity constraints for $\mu_i$ and $\rho_i$ for all $i$ are guaranteed by the fact that $c_i$ are in nonincreasing order in $i$. Therefore, $(\mu_1^*, \ldots, \mu_n^*)$ is an optimal solution of (Dual of LP.1).
\( \cdots, \mu^*_n, \rho^*_1, \cdots, \rho^*_n, w^* \) is a feasible solution of (Dual of LP.1). For the pair of primal and dual problems shown in (LP.1) and (Dual of LP.1), the complementary slackness condition can be written as follows:

\[
\begin{align*}
(c_i - \mu_i - \rho_i - w) y_i &= 0, & \forall i = 1, \ldots, n \quad (5.52) \\
\mu_i (y_i - y^*_i) &= 0, & \forall i = 1, \ldots, n \quad (5.53) \\
\rho_i (y^*_i - y_i) &= 0, & \forall i = 1, \ldots, n \quad (5.54)
\end{align*}
\]

For any \( i \in \{1, 2, \ldots, n\} \), plugging (5.51) into the left-hand-size of equation (5.52), condition (5.52) satisfies. For any \( i \in \mathcal{I}_1 \), \( y^*_i = y^*_j \) and therefore \( \mu^*_i (y^*_i - y^*_j) = 0 \). For any \( i \in \mathcal{I}_2 \cap \{\tau\} \), \( \mu^*_i = 0 \) and therefore \( \mu^*_i (y_j - y^*_j) = 0 \). Thus, the condition (5.53) holds. For any \( i \in \mathcal{I}_1 \cap \{\tau\} \), \( \rho^*_i = 0 \) and therefore \( \rho^*_i (y^*_i - y^*_j) = 0 \). For any \( i \in \mathcal{I}_2 \), \( y^*_i = y^*_u \) and therefore \( \rho^*_i (y^*_u - y^*_i) = 0 \). Thus, the condition (5.54) holds. Since \( (\mu^*_1, \cdots, \mu^*_n, \rho^*_1, \cdots, \rho^*_n, w^*) \) and \( (y^*_1, \cdots, y^*_n) \) satisfy the complementary slackness conditions (5.52–5.54), \( (y^*_1, \cdots, y^*_n) \) is an optimal solution to (LP.1). 

Proposition 5.4.2 establishes an \( O(|\mathcal{L}|^2) \) algorithm for solving the relaxed RMDP-TM under the condition that there is no uncertainty budget contraint. Next, we provide sufficient conditions under which the relaxed RMDP-TM (5.4) and the RMDP-TM (5.3) have zero optimality gap, namely, \( v_t(s) = \tilde{v}_t(s), \forall t \in \mathcal{T}, s \in \mathcal{L} \times \mathcal{M} \). This provides sufficient conditions under which the RMDP-TM (5.3) can be solved via Algorithm 2. We begin with a definition of nonincreasing worst-case (NIWC) TPM that is relevant to our theorem.

**Definition 5.4.1.** The one-step transition probability matrix, \( Q^\text{NIWC}_t(\Gamma) \), at epoch \( t \) is called NIWC TPM in the uncertainty set \( Q^\text{IMUB}_t = \prod_{\ell_t \in \mathcal{L}} Q^\text{IMUB}_{t,\ell_t}(\Gamma) \), if for any \( \ell_t \in \mathcal{L} \), row \( \ell_t \) of \( Q^\text{NIWC}_t(\Gamma) \) is an optimal solution of the inner problem (5.15) with nonincreasing objective function coefficients.

The following theorem establishes an important relationship between the RMDP-TM and
the relaxed RMDP-TM in the absence of an uncertainty budget.

**Theorem 5.4.1.** If the following conditions hold:

(I): The uncertainty set \( Q_{t,t_t} \) of the relaxed RMDP-TM (5.4) is the IMUB model for all \( t \in T, m_t \in M, \) and \( \Gamma_{t,t_t} = |L|, \forall \ell_t \in L, t \in T, \)

(II): \( Q_{t,t_t}(\Gamma_{t,t_t}) = Q_{t',t_t}(\Gamma'_{t,t_t}), \forall \ell_t \in L, t, t' \in T, \)

(III): \( r_t(\ell_t, m_t, \alpha_t) \) is nonincreasing in \( \ell_t \) for any \( m_t \in M, \alpha_t \in \mathcal{A}_t(\ell_t, m_t), \) and \( t \in T \setminus \{T\}, \)

and \( r_T(\ell_T, m_T) \) is nonincreasing in \( \ell_T \) for any \( m_T \in M, \)

(IV): \( p^E_t(\ell_t, m_t, \alpha_t) \) is nondecreasing in \( \ell_t \) for any \( m_t \in M, \alpha_t \in \mathcal{A}_t(\ell_t, m_t), \) and \( t \in T, \)

(V): \( Q^\text{NIWC}_t(|L|), \forall t \in T \) have the increasing failure rate (IFR) property [129],

then the following holds:

(a): the optimal value function of the relaxed RMDP-TM, \( \bar{v}_t(\ell_t, m_t) \), is nonincreasing in \( \ell_t \) for any \( m_t \in M, \)

(b): the optimal policy of nature obtained by solving (5.4) is stationary, namely, \( Q^*_1 = Q^*_2 = \cdots = Q^*_{T-1}, \)

and

(c): the relaxed RMDP-TM (5.4) and the RMDP-TM (5.3) have zero optimality gap.

**Proof.** This proof is by induction. For the base case \( t = T \), we know \( \bar{v}_T(\ell_T, m_T) \) is nonincreasing in \( \ell_T \) by assumption (III). Now assume that \( \bar{v}_k(\ell_k, m_k) \) is nonincreasing in \( \ell_k \in L \) for \( k = t + 1, t + 2, \ldots, T - 1 \), then we need to prove this for \( k = t \). For \( k = t + 1, t + 2, \ldots, T - 1 \), based on assumptions (I), (II), and the induction assumption that \( \bar{v}_k(\ell_k, m_k) \) is nonincreasing in \( \ell_k \), the optimal solution to the inner problem (5.15) is an NIWC TPM, namely,

\[
Q^*_k = Q^\text{NIWC}_k(|L|), k = t, t + 1, \ldots, T - 1. \tag{5.55}
\]
In addition, \( Q^*_k, \ k = t + 1, \ldots, T - 1 \) have the IFR property based on assumption (V). by assumptions (III), and (IV), \( \tilde{v}_t(\ell, m_t) \) is nonincreasing in \( \ell_t \) for any \( m_t \in M \) by Lemma 4.3.1. The induction hypothesis is satisfied, and the result follows. Moreover, under the conditions (I), and (II), the optimal policy of nature is \( Q^*_1 = Q^*_2 = \cdots = Q^*_T = Q_{NIWC}(|L|) \), and therefore, it is stationary. Therefore, the RMDP-optimal solution obtained by solving the relaxed RMDP model is also an optimal solution to the RMDP-TM, and (5.4) and (5.3) have zero optimality gap.

Based on the Theorem 5.4.1, we develop an approximate algorithm to solve the RMDP-TM as shown in Algorithm 4. In the approximate algorithm, we first re-order metabolic states \( \ell_t \in \mathcal{L} \) so that the end-of-horizon rewards are non-increasing in \( \ell_t \). Then we obtain the NIWC TPM \( Q_{NIWC}(\Gamma) \), by solving the inner problem for the decision epoch \( T - 1 \). Finally, we solve the RMDP-TM as an NMDP where the TPM is set to be the NIWC TPM. We denote \( v_{NIWC}^t(s_t) \), \( \forall t \in mcT, s_t \in S \), to be the optimal value functions of the approximate RMDP-TM, and it can be written as follows:

\[
v_{NIWC}^t(s_t) = \begin{cases}
\max_{\alpha_t \in A_t(\ell_t, m_t)} \left\{ r_t(\ell_t, m_t, \alpha_t) + (1 - p^E_t(\ell_t, m_t, \alpha_t))\lambda \sum_{\ell_{t+1} \in \mathcal{L}} q_{NIWC}^{\ell_t, \ell_{t+1}} v_{NIWC}^{t+1}(\ell_{t+1}, m_{t+1}) \right\}, \\
\quad \forall s_t = (\ell_t, m_t) \in \mathcal{L} \times \mathcal{M}, t \in T \setminus \{T\}, \\
r_T(s_T), \\
0,
\end{cases}
\]

(5.56)

**Proposition 5.4.3.** The optimal value functions of the approximate RMDP-TM and the optimal value functions of the relaxed RMDP-TM provide the upper bound and lower bound of the optimal value functions of the RMDP-TM, respectively.

\[
v_{NIWC}^t(s_t) \geq v_t(s_t) \geq \tilde{v}_t(s_t), \forall s_t \in S, \ t \in T.
\]

(5.57)

Equalities hold when conditions in Theorem 5.4.1 hold.
Algorithm 4: An approximate algorithm to solve the RMDP-TM

1: Step 1. Initialize the value function at stage $T$:

\[ v_T(\ell_T, m_T) \leftarrow R_T(\ell_T, m_T), \forall \ell_T \in \mathcal{L}, m_T \in \mathcal{M} \]

2: Step 2. Order $v_T$ from the largest to the smallest, relabel state so that

\[ v_T(\ell(1)) \geq v_T(\ell(2)) \geq \cdots \geq v_T(\ell(|\mathcal{L}|)) \]

3: Step 3. Compute NIWC TMP: $Q^{NIWC}(\Gamma)$

4: Step 4. Recursion

5: for $t = T - 1 \rightarrow 0$ do

6: for $\ell_t \in \mathcal{L}$ do

7: for $m_t \in \mathcal{M}$ do

8: for $\alpha_t \in A_t(\ell_t, m_t)$ do

9: Calculate

\[ w_t(\ell_t, m_t, \alpha_t) = r_t(\ell_t, m_t, \alpha_t) + \lambda [1 - p_t^E(\ell_t, m_t, \alpha_t)] \sum_{\ell_{t+1} \in \mathcal{L}} q_{t+1}^{NIWC, \Gamma}(\ell_{t+1}) v_{t+1}(\ell_{t+1}, m_{t+1}) \]

10: end for

11: Update Value Function:

\[ v_t(\ell_t, m_t) \leftarrow \max_{\alpha_t \in A_t(\ell_t, m_t)} \{ w_t(\ell_t, m_t, \alpha_t) \} \]

12: Update Optimal Action Set:

\[ A^*_t(\ell_t, m_t) \leftarrow \arg\max_{\alpha_t \in A_t(\ell_t, m_t)} \{ w_t(\ell_t, m_t, \alpha_t) \} \]

13: end for

14: end for

15: end for
Proof. The collection of NIWC TMPs,

$\theta^NIWC_t = (Q^{NIWC}, Q^{NIWC}, \ldots, Q^{NIWC})$,

is a stationary nature’s policy, therefore,

$v^{NIWC}_t(s_t) \geq v_t(s_t)$.

Since the relaxed RMDP-TM is defined by relaxing the stationary requirement for nature’s policy,

$v_t(s_t) \geq \tilde{v}_t(s_t)$.

Theroem 5.4.1 provides sufficient conditions under which the optimal nature’s policy for the relaxed RMDP-TM is stationary and equals to the collection of NIWC TMPs, therefore, when those conditions hold,

$v^{NIWC}_t(s_t) = v_t(s_t) = \tilde{v}_t(s_t), \forall s_t \in S, t \in T$. \hfill (5.58)

The remainder of this section discusses the influence of the uncertainty budget parameter on the RMDP-TM. As mentioned before, the uncertainty budget parameter controls the size of the uncertainty set and therefore controls the conservativeness of the optimal solution. We define an effective budget, $\Gamma^*$, such that beyond the effective budget any increase in the uncertainty budget will not affect the optimal solution of the inner problem and therefore will not change the optimal decision of nature.

**Definition 5.4.2.** We define the effective budget of RMDP-TM as:

$$\Gamma^*(v_{t+1}(m_{t+1})) = \max_{t \in T} \max_{m_{t+1} \in M} \Gamma^*_t(Q^{IMUB}_t(|L|), v_{t+1}(m_{t+1}))$$ \hfill (5.59)
where

\[ v_{t+1}(m_{t+1}) = (v_{t+1}(l_t(1), m_{t+1}), v_{t+1}(l_t(2), m_{t+1}), \ldots, v_{t+1}(l_t(|L|), m_{t+1})) \]  \hspace{1cm} (5.60)

\[ \Gamma^*_t(Q_{t, l_t}^{\text{IMUB}}(|L|), v_{t+1}(m_{t+1})) = \max_{l_t \in L} q_{t, l_t}^* \min_{l_t \in L} q_{t, l_t}^* \left[ \sum_{i \in I} \frac{q_{t, l_t}^*(i) - \hat{q}_{t, l_t}(i)}{q_{t, l_t}^*(i) - \hat{q}_{t, l_t}(i)} + \sum_{j \in J} \frac{q_{t, l_t}^*(j) - \hat{q}_{t, l_t}(j)}{q_{t, l_t}^*(j) - \hat{q}_{t, l_t}(j)} \right] \]  \hspace{1cm} (5.61)

for all \( t \in T \), \( m_{t+1} \in M \), and \( Q_{t, l_t}^* \) is the set of optimal solutions of the inner problem (5.15) over the uncertainty set, \( Q_{t}^{\text{IMUB}}(|L|) \). \( I = \{ i \in L : q_{t, l_t}^*(i) \leq \hat{q}_{t, l_t}(i), \hat{q}_{t, l_t}(i) \neq 0 \} \), and \( J = \{ j \in L : q_{t, l_t}^*(j) > \hat{q}_{t, l_t}(j) \} \).

The effective budget \( \Gamma^* \) is closely related to the optimal RMDP-TM policy and the optimal value of RMDP-TM. We define the uncertainty budget of the RMDP-TM as follows:

\[ \Gamma = \max_{t \in T, l_t \in L} \Gamma_{t, l_t}. \]

For any RMDP-TM with uncertainty budget greater than or equals to the effective budget, the optimal RMDP-TM policies and the optimal value of RMDP-TMs are the same. If we denote \( \pi^*(\Gamma) \) and \( v^*_{\Gamma}(s_t) \), \( \forall s_t \in S \), to be the optimal policy and the optimal values of the RMDP-TM with uncertainty budget \( \Gamma \), respectively, we have

\[ \pi^*(\Gamma) = \pi^*(\Gamma^*), \ \forall \Gamma \geq \Gamma^*, \]  \hspace{1cm} (5.62)

and

\[ v^*_{\Gamma}(s_t) = v^*_{\Gamma^*}(s_t), \ \forall s_t \in S, \Gamma \geq \Gamma^*. \]  \hspace{1cm} (5.63)

On the other hand, for RMDP-TMs with uncertainty budget less than the effective budget, the optimal RMDP-TM policies may not be the same, even if the optimal policies are the same, the optimal values of the RMDP-TMs may be different.

The effective budget is not easy to compute since the calculation is based on solutions of all inner problems as shown in Equations (5.59) and (5.61). Therefore, to obtain the effective budget
of an RMDP-TM is as hard as solving this RMDP-TM. The following proposition provides conditions under which the effective budget of an RMDP-TM can be easily calculated.

**Proposition 5.4.4.** If the following conditions hold

1. $Q_t^{\text{IMUB}}(\mid L \mid) = Q_{t'}^{\text{IMUB}}(\mid L \mid), \forall t, t' \in T$, and

2. $v_t(\ell_t, m_t)$ is strictly decreasing in $\ell_t$ for any $m_t \in M, t \in T$,

then

$$\max_{\ell \in \mathcal{L}} K_\ell - 1 \leq \Gamma^* \leq \max_{\ell \in \mathcal{L}} K_\ell,$$  \hspace{1cm} (5.64)

where $K_\ell, \forall \ell \in \mathcal{L}$ is the number of nonzero elements in row $\ell$ of $\hat{Q}$, where $\hat{Q}$ is the center of the uncertainty set $Q_t^{\text{IMUB}}(\mid L \mid), \forall t \in T$.

**Proof.** Based on assumption (2), for all $t \in T, m_{t+1} \in M$, the inner problem (5.15) with the uncertainty set $Q_t^{\text{IMUB}}(\mid L \mid)$ has an unique optimal solution $q^*_{t, \ell_t}$. Thus,

$$\Gamma^*_t(Q_t^{\text{IMUB}}(\mid L \mid), v_{t+1}(m_{t+1})) = \max_{\ell \in \mathcal{L}} \left[ \sum_{i \in \mathcal{I}} \hat{q}_{t, \ell_t}(i) - q^*_{t, \ell_t}(i) + \sum_{j \in \mathcal{J}} q^*_{t, \ell_t}(j) - \hat{q}_{t, \ell_t}(j) \right]$$

$$= \max_{\ell \in \mathcal{L}} \left[ \sum_{i \in \mathcal{I}} \hat{q}_{t, \ell_t}(i) - q^\text{NIWC}_{\ell_t}(i) + \sum_{j \in \mathcal{J}} q^\text{NIWC}_{\ell_t}(j) - \hat{q}_{t, \ell_t}(j) \right]$$  \hspace{1cm} (5.65)

for all $m_{t+1} \in M$ where $q^\text{NIWC}_{\ell_t}(\cdot), \forall \ell_t$ denotes the element of $Q_t^{\text{NIWC}}(\mid L \mid)$. It implies that $\Gamma^*_t(Q_t^{\text{IMUB}}(\mid L \mid), v_{t+1}(m_{t+1}))$ is independent of $m_{t+1}$. Moreover, based on assumption (1), we have $Q_t^{\text{NIWC}}(\mid L \mid) = Q_{t'}^{\text{NIWC}}(\mid L \mid) \triangleq Q^{\text{NIWC}}(\mid L \mid)$. Therefore,

$$\Gamma^* = \max_{\ell \in \mathcal{L}} \left[ \sum_{i \in \mathcal{I}} \hat{q}_{\ell}(i) - q^\text{NIWC}_{\ell}(i) + \sum_{j \in \mathcal{J}} q^\text{NIWC}_{\ell}(j) - \hat{q}_{\ell}(j) \right],$$  \hspace{1cm} (5.66)

where $q^\text{NIWC}_{\ell}(\cdot), \forall \ell \in \mathcal{L}$ denotes the element in $\hat{Q}$, and $q^\text{NIWC}_{\ell}(\cdot), \forall \ell \in \mathcal{L}$ denotes the element in $Q^{\text{NIWC}}(\mid L \mid)$. Based on Proposition 5.4.2 and from the Algorithm 3, there exists at most one
state $\ell(\tau)$ such that

$$q^d_\ell(\ell(\tau)) < q^\text{NIWC}_\ell(\ell(\tau)) < q^u_\ell(\ell(\tau)).$$

For all other states $\ell' \in \mathcal{L}\setminus\{\ell(\tau)\}$, $q^\text{NIWC}_\ell(\ell')$ equals $q^d_\ell(\ell')$ or $q^u_\ell(\ell')$. Therefore,

$$\Gamma^* = \left\{ \begin{array}{ll}
\max_{\ell \in \mathcal{L}} [K_\ell - 1 + \frac{q^d_\ell(\ell(\tau)) - q^\text{NIWC}_\ell(\ell(\tau))}{q^d_\ell(\ell(\tau)) - q^u_\ell(\ell(\tau))}], & \text{if } q^d_\ell(\ell') < q^\text{NIWC}_\ell(\ell(\tau)) < q^u_\ell(\ell'), \\
\max_{\ell \in \mathcal{L}} [K_\ell - 1 + \frac{q^\text{NIWC}_\ell(\ell(\tau)) - q^u_\ell(\ell(\tau))}{q^d_\ell(\ell(\tau)) - q^u_\ell(\ell(\tau))}], & \text{if } q^\text{NIWC}_\ell(\ell(\tau)) > q^u_\ell(\ell') > q^d_\ell(\ell(\tau)), \\
\max K_\ell - 1 & \text{if } q^\text{NIWC}_\ell(\ell(\tau)) = q^d_\ell(\ell(\tau)), \\
\max K_\ell & \text{if } q^\text{NIWC}_\ell(\ell(\tau)) = q^u_\ell(\ell'), \text{ or } q^\text{NIWC}_\ell(\ell(\tau)) = q^d_\ell(\ell').
\end{array} \right.$$

Finally, from Equation (5.67) we have

$$\max_{\ell \in \mathcal{L}} K_\ell - 1 \leq \Gamma^* \leq \max_{\ell \in \mathcal{L}} K_\ell.$$

\[\square\]

**Remark 3:** The assumption (2) can be relaxed. The conclusion in Equation (5.64) holds as long as $v_t(\ell_t, m_t)$ has consistent ordering in $\ell_t$ for any $m_t \in \mathcal{M}, t \in \mathcal{T}$ and for any $\ell, \ell' \in \mathcal{L}$, $v_t(\ell', m_t) \neq v_t(\ell', m_t), \forall t \in \mathcal{T}, m_t \in \mathcal{M}$.

For some MDP models in which the transitions among states are restricted (see for example [130]), the corresponding RMDP models will have same effective budget based on the Proposition 5.4.4. Therefore, even with big uncertainty budget, the difference in the optimal values between the NMDP and the RMDP may be small.

### 5.5 Numerical Experiments

In this section, we present a case study to illustrate the application of the proposed RMDP glycemc control for patients with type 2 diabetes. As in Chapter 4, we consider the three most commonly used medications: metformin, sulfonylurea, and insulin for the U.S. population newly diagnosed with type 2 diabetes. The objective is to maximize QALYs prior to the patient’s first
macro- or micro-vascular event or death from other causes (base-case in Chapter 4). All model parameters are the same as Chapter 4 except the TPMs. In order to investigate the variation of the uncertainty budget in the RMDP-optimal policy, we use a small sample to estimate HbA1c transition probability matrices and their uncertainty set. This sample was generated by extracting patients who had at least 15 HbA1c records within 5 years. The set of patients comprised 272 female patients and 377 male patients.

The reminder of this section is organized as follows: in Section 5.5.1 we describe a method to estimate the uncertainty set for the HbA1c state transition probabilities. In Section 5.5.2 we present and compare the base-case results obtained by solving the relaxed RMDP-TM exactly, and by solving the RMDP-TM approximately. In Section 5.5.3 we compare the computation time to solve the relaxed RMDP-TM with NLP formulation and LP formulation exactly, and the computation time to solve the RMDP-TM approximately. Finally, in Section 5.5.4 we evaluate the performance of the RMDP-optimal policy for the worst-case and by random sampling of TPMs in the uncertainty set.

5.5.1 Uncertainty Set Estimation

We assumed that uncertainty sets are the same for all decision epochs, namely, \( Q_1 = Q_2 = \cdots = Q_T \). We first calculate the \( 100 \times (1 - \alpha/|L|) \)% simultaneous confidence intervals for each row of matrix \( Q_t \); and then based on the Bonferroni inequality, all \( |L| \) confidence intervals will form the \( 100 \times (1 - \alpha) \)% confidence region for matrix \( Q_t \). For any \( \ell_t, \ell_{t+1} \in L \), the following Equation (5.68), proposed by Gold (1963) [131], was used to calculate the confidence interval for the true transition probability \( q_{\ell_t}(\ell_{t+1}) \):

\[
\left( \hat{q}_{\ell_t}(\ell_{t+1}) - \left[ \chi^2_{|L|-1,\alpha/(2|L|)} \frac{\hat{q}_{\ell_t}(\ell_{t+1})(1-\hat{q}_{\ell_t}(\ell_{t+1}))}{N_{\ell_t}} \right]^\frac{1}{2} \hat{q}_{\ell_t}(\ell_{t+1}) + \left[ \chi^2_{|L|-1,1-\alpha/(2|L|)} \frac{\hat{q}_{\ell_t}(\ell_{t+1})(1-\hat{q}_{\ell_t}(\ell_{t+1}))}{N_{\ell_t}} \right]^\frac{1}{2} \right)
\]

(5.68)

where \( \hat{q}_{\ell_t}(\ell_{t+1}) = \frac{n_{\ell_t,\ell_{t+1}}}{N_{\ell_t}} \), and \( n_{\ell_t,\ell_{t+1}} \) denotes the observed number of patients who transition from state \( \ell_t \) to state \( \ell_{t+1} \), and \( N_{\ell_t} = \sum_{\ell_{t+1} \in L} n_{\ell_t,\ell_{t+1}} \) denotes the total number of patients in state
The maximal left-hand-side and right-hand-side variation for transition probability \( q_{\ell t}(\ell_{t+1}) \) can be calculated as follows:

\[
\delta^L_{\ell t}(\ell_{t+1}) = \min \left\{ \hat{q}_{\ell t}(\ell_{t+1}), \left[ \chi^2_{|L|-1,\alpha/(2|L|)} \frac{\hat{q}_{\ell t}(\ell_{t+1})(1-\hat{q}_{\ell t}(\ell_{t+1}))}{N_{\ell t}} \right]^{\frac{1}{2}} \right\} \\
= \begin{cases} 
\hat{q}_{\ell t}(\ell_{t+1}), & \text{if } \hat{q}_{\ell t}(\ell_{t+1}) \leq \frac{\chi^2_{|L|-1,\alpha/(2|L|)}}{1 + \chi^2_{|L|-1,\alpha/(2|L|)}} \\
\left[ \chi^2_{|L|-1,\alpha/(2|L|)} \frac{\hat{q}_{\ell t}(\ell_{t+1})(1-\hat{q}_{\ell t}(\ell_{t+1}))}{N_{\ell t}} \right]^{\frac{1}{2}}, & \text{otherwise.}
\end{cases}
\] (5.69)

and

\[
\delta^U_{\ell t}(\ell_{t+1}) = \min \left\{ 1 - \hat{q}_{\ell t}(\ell_{t+1}), \left[ \chi^2_{|L|-1,1-\alpha/(2|L|)} \frac{\hat{q}_{\ell t}(\ell_{t+1})(1-\hat{q}_{\ell t}(\ell_{t+1}))}{N_{\ell t}} \right]^{\frac{1}{2}} \right\} \\
= \begin{cases} 
1 - \hat{q}_{\ell t}(\ell_{t+1}), & \text{if } \hat{q}_{\ell t}(\ell_{t+1}) \geq \frac{1}{1 + \chi^2_{|L|-1,1-\alpha/(2|L|)}} \\
\left[ \chi^2_{|L|-1,1-\alpha/(2|L|)} \frac{\hat{q}_{\ell t}(\ell_{t+1})(1-\hat{q}_{\ell t}(\ell_{t+1}))}{N_{\ell t}} \right]^{\frac{1}{2}}, & \text{otherwise.}
\end{cases}
\] (5.70)

### 5.5.2 Base Case Analysis

Tables 3, 5, and 6 in the Appendix contain the MLE, left-hand-side variation, and right-hand-side variation of the HbA1c state TPM for female patients, and Tables 9, 11, and 12 in the Appendix contain the MLE, left-hand-side variation, and right-hand-side variation of the HbA1c state TPM for male patients. These estimates were used for our numerical experiments. We found that the MLE of the TPMs for both male and female patients did not satisfy the IFR assumption exactly. Based on the following worst-case violation measurement [32]:

\[
\epsilon = \max_{t \in T} \max_{j \in \{1,\ldots,|L|-1\}} \max_{i \in \{1,\ldots,|L|\}} \sum_{s=1}^{|L|} \left[ q_{t,\ell t(j)}(\ell_{t+1}(s)) - q_{t,\ell t(j+1)}(\ell_{t+1}(s)) \right].
\] (5.71)

The violation is 0.125 for female patients and 0.1191 for male patients. We also calculated the NIWC TPM for female and male patients as shown in Tables 7 and 13 in the Appendix, respectively. Again, we found the NIWC TPMs for both genders do not satisfy the IFR assumption exactly. The violation is 0.0898 for female patients and 0.1129 for male patients. However, under
this violation, we still found that the optimal value function of the relaxed RMDP-TM with any uncertainty budget level is nonincreasing in the HbA1c state. We used the definition of effective budget shown in Equation (5.59) to calculate the effective budget for the RMDP-TM. The effective budget is 6.88 for females and 8.53 for males. Therefore, before solving the relaxed RMDP-TM, we know that the RMDP-optimal solution, and the RMDP-optimal policies of the relaxed RMDP-TM with uncertainty budgets $\geq 6.88$ are the same for females, and $\geq 8.53$ are the same for males. After solving the relaxed RMDP-TM for $\Gamma = 0, 1 \cdots, |\mathcal{L}|$, we found that the RMDP-optimal policies with budgets $\geq 6$ are the same for female and $\geq 5$ are the same for males.

We use Algorithm 2 to solve the relaxed RMDP-TM exactly, and Algorithm 4 to solve the RMDP-TM approximately. From Table 5.1, we can see that the exact method used to solve the relaxed RMDP-TM and the approximate method used to solve the RMDP-TM generate the same solutions for all uncertainty budget levels range from 1 to 10. Therefore, the optimal solutions in Table 5.1 are indeed the optimal solutions of the RMDP-TM.

Table 5.1: Comparison of the expected worst-case QALYs prior to first event for glycemic control case study under base case. Upper bounds are obtained by solving RMDP-TMs approximately using Algorithm 4, and lower bounds are obtained by solving the relaxed RMDP-TMs exactly using Algorithm 2.

<table>
<thead>
<tr>
<th align="left">Uncertainty budget</th>
<th align="left">Model</th>
<th align="left">Men</th>
<th align="left"></th>
<th align="left"></th>
<th align="left">Women</th>
<th align="left"></th>
<th align="left"></th>
</tr>
</thead>
<tbody>
<tr>
<td align="left"></td>
<td align="left"></td>
<td align="left">Upper bound</td>
<td align="left">Lower bound</td>
<td align="left"></td>
<td align="left">Upper bound</td>
<td align="left">Lower bound</td>
<td align="left"></td>
</tr>
<tr>
<td align="left">0</td>
<td align="left">64.4854</td>
<td align="left">64.4854</td>
<td align="left"></td>
<td align="left">68.5096</td>
<td align="left">68.5096</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">1</td>
<td align="left">64.4502</td>
<td align="left">64.4502</td>
<td align="left"></td>
<td align="left">68.4731</td>
<td align="left">68.4731</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">2</td>
<td align="left">64.4264</td>
<td align="left">64.4264</td>
<td align="left"></td>
<td align="left">68.4509</td>
<td align="left">68.4509</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">3</td>
<td align="left">64.4119</td>
<td align="left">64.4119</td>
<td align="left"></td>
<td align="left">68.4349</td>
<td align="left">68.4349</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">4</td>
<td align="left">64.4032</td>
<td align="left">64.4032</td>
<td align="left"></td>
<td align="left">68.4228</td>
<td align="left">68.4228</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">5</td>
<td align="left">64.3986</td>
<td align="left">64.3986</td>
<td align="left"></td>
<td align="left">68.4178</td>
<td align="left">68.4178</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">6</td>
<td align="left">64.3981</td>
<td align="left">64.3981</td>
<td align="left"></td>
<td align="left">68.4158</td>
<td align="left">68.4158</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">7</td>
<td align="left">64.3980</td>
<td align="left">64.3980</td>
<td align="left"></td>
<td align="left">68.4152</td>
<td align="left">68.4152</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">8</td>
<td align="left">64.3979</td>
<td align="left">64.3979</td>
<td align="left"></td>
<td align="left">68.4152</td>
<td align="left">68.4152</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">9</td>
<td align="left">64.3979</td>
<td align="left">64.3979</td>
<td align="left"></td>
<td align="left">68.4152</td>
<td align="left">68.4152</td>
<td align="left"></td>
<td align="left"></td>
</tr>
<tr>
<td align="left">10</td>
<td align="left">64.3979</td>
<td align="left">64.3979</td>
<td align="left"></td>
<td align="left">68.4152</td>
<td align="left">68.4152</td>
<td align="left"></td>
<td align="left"></td>
</tr>
</tbody>
</table>
Figure 5.1: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently not on any medications when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin.

Figure 5.2: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently on metformin only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only.
Figure 5.3: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently on sulfonylurea only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin.

Figure 5.4: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for females who are currently on metformin and sulfonylurea together when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only.
Figure 5.5: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently not on any medications. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin.

Figure 5.6: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently on metformin only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only.
Figure 5.7: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently on sulfonylurea only when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and metformin plus insulin.

Figure 5.8: A comparison of NMDP-optimal policy and RMDP-optimal policies for the first 6 years after diagnosis for males who are currently on metformin and sulfonylurea together when hypoglycemia is considered as a transient symptom. The following medication states are never optimal in the figure: no medication, metformin only, sulfonylurea only, and insulin only.
The optimal action for patients who are on insulin is to keep using current medication(s) for the NMDP-optimal policy and all RMDP-optimal policies with various choices of $\Gamma$. Figures 5.1, 5.2, 5.3, and 5.4 compare the NMDP-optimal policy and RMDP-optimal policies for females who are not on any medications, on metformin only, on sulfonylurea only, and on metformin and sulfonylurea together, respectively. Figures 5.5, 5.6, 5.7, and 5.8 compare the NMDP-optimal policy and RMDP-optimal policies for males who are not on any medications, on metformin only, on sulfonylurea only, and on metformin and sulfonylurea together, respectively. RMDP-optimal policies and NMDP-optimal policy differ in the optimal start time to initiate medication.

### 5.5.3 Algorithm Performance Comparison

We present computational results for all numerical experiments for NMDPs, relaxed RMDP-TMs, and RMDP-TMs. NMDPs were solved using backward induction, relaxed RMDP-TMs were solved exactly using Algorithm 2, and RMDP-TMs were solved approximately using Algorithm 4. We present the computation time of solving relaxed RMDP-TMs with the NLP formulation and the LP formulation of inner problems to show the difference between the two formulations in terms of computational efficiency. The computational (CPU) times are summarized in Table 5.2.

From Table 5.2 we can see that NMDPs can be solved much faster than relaxed RMDP-TMs regardless of the formulation of the inner problem. It is because solving the relaxed RMDP-TM requires solving $|S| \cdot |T\setminus\{T\}|$ inner problems need to be solved. Moreover, the computation time of solving the relaxed RMDP-TM exactly does not change significantly as the uncertainty budget changes. The CPU time used to solve the relaxed RMDP-TM with LP formulation of the inner problems is about one tenth of the time used to solve the relaxed RMDP-TM with NLP formulation of the inner problems. This CPU time difference demonstrates the importance of Proposition 5.4.1. The approximate algorithm significantly improved the CPU time for solving the RMDP-TM; and in our experiments, it generates the exact optimal solution for the RMDP-TM.
Table 5.2: Comparison of the computation time (in seconds) of solving the NMDP and the relaxed RMDP-TM exactly, and solving the RMDP-TM approximately.

<table>
<thead>
<tr>
<th>Uncertainty budget</th>
<th>Relaxed RMDP-TMs solved by Algorithm 2</th>
<th>Male LP  NLP</th>
<th>Female LP  NLP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1149 10205</td>
<td>1285 10680</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1141 10417</td>
<td>1291 10613</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1185 10639</td>
<td>1273 10970</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1127 10848</td>
<td>1278 10196</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1124 10444</td>
<td>1248 10668</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1185 10104</td>
<td>1293 10643</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1170 10959</td>
<td>1269 10184</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1154 10915</td>
<td>1253 10857</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1143 10856</td>
<td>1243 10089</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1169 10209</td>
<td>1212 10342</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1126 10891</td>
<td>1260 10068</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RMDP-TMs solved by Algorithm 4</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

NMDP Model | – | 1 | 1

103
5.5.4 RMDP-optimal Policy Performance Comparison

In this section, we consider to evaluate and compare the performance of the NMDP-optimal policy, and RMDP-optimal policies for the base-case glycemic control model and an alternative RMDP-TM glycemic control model in which hypoglycemia is considered as an event (an absorbing state). Results of RMDP-optimal policy were obtained by evaluating the NMDP-optimal policy, and RMDP-optimal policies based on random sampling of TPMs in the uncertainty set and based on the worst-case TPMs.

For the purpose of evaluating and comparing the NMDP-optimal and RMDP-optimal policies under various choices of TPMs, a modified version of Smith’s mixing algorithm, shown in Algorithm 5, was used to randomly sample TPMs over the uncertainty set [132]. Detailed description was provided in Wu (2014) [133]. Each RMDP-optimal policy was evaluated under the sampled TPM to generate the expected QALYs prior to the first event. In the worst-case approach, the expected QALYs prior to first event were generated by evaluating each pre-generated RMDP-optimal policy under the worst case TPM.

Table 5.3: Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for men in base case when hypoglycemia is considered as a transient symptom.

<table>
<thead>
<tr>
<th>NMDP-optimal policy</th>
<th>Worst-case approach</th>
<th>Sampling approach Minimum</th>
<th>Mean</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>64.38980</td>
<td>64.4374</td>
<td>64.4862</td>
<td>64.5401</td>
</tr>
</tbody>
</table>

| RMDP-optimal policy | Γ = 1 | 64.39323 | 64.4383 | 64.4852 | 64.5374 |
|                     | Γ = 2 | 64.39635 | 64.4367 | 64.4814 | 64.5312 |
|                     | Γ = 3 | 64.39780 | 64.4344 | 64.4765 | 64.5235 |
|                     | Γ = 4 | 64.39789 | 64.4341 | 64.4758 | 64.5224 |
|                     | Γ = 5 | 64.39794 | 64.4335 | 64.4748 | 64.5212 |
|                     | Γ = 6 | 64.39794 | 64.4335 | 64.4748 | 64.5212 |
|                     | Γ = 7 | 64.39794 | 64.4335 | 64.4748 | 64.5212 |
|                     | Γ = 8 | 64.39794 | 64.4335 | 64.4748 | 64.5212 |
|                     | Γ = 9 | 64.39794 | 64.4335 | 64.4748 | 64.5212 |
|                     | Γ = 10| 64.39794| 64.4335 | 64.4748 | 64.5212 |
Algorithm 5 Modified Smith’s Mixing Algorithm

1: Initialize $Q_0 \leftarrow \hat{Q}$ \hspace{1cm} \triangleright \hat{Q}: \text{MLE of the HbA1c TPM}

2: for $n = 1 : N$ do \hspace{1cm} \triangleright N \text{ is the total number of samples}

3: \hspace{1em} for $h = 1 : |\mathcal{L}|$ do \hspace{1cm} \triangleright Loop through all HbA1c states

4: \hspace{2em} $X_0 \leftarrow Q_{n-1}(h)$ \hspace{1cm} \triangleright $Q_{n-1}(h)$: $h^{th}$ row of sampled TPM $Q_{n-1}$

5: \hspace{2em} while $j \leq M$ do

6: \hspace{3em} Find a random direction $d$ from $X_{j-1}$

7: \hspace{3em} Set $\hat{\lambda} \leftarrow \max \{ \alpha^+ : X_{j-1} + \alpha^+ d \in Q^M_{l(h)} \}$ as follows:

8: \hspace{4em} for $l = 1 : |\mathcal{L}|$ do

9: \hspace{5em} if $d_l > 0$ then

10: \hspace{6em} $\lambda^+_l \leftarrow \frac{1-x_{j-1,l}}{d_l}$

11: \hspace{5em} else if $d_l < 0$ then

12: \hspace{6em} $\lambda^+_l \leftarrow \frac{-x_{j-1,l}}{d_l}$

13: \hspace{5em} else

14: \hspace{6em} $\lambda^+_l \leftarrow 0$

15: \hspace{4em} end if

16: \hspace{3em} end for

17: \hspace{3em} $\bar{\lambda} \leftarrow \min \{ \lambda^+_l : 0 < l \leq |\mathcal{L}| \}$

18: \hspace{3em} Set $\lambda \leftarrow \min \{ \alpha^- : X_{j-1} + \alpha^- d \in Q^M_{l(h)} \}$ as follows:

19: \hspace{4em} for $l = 1 : |\mathcal{L}|$ do

20: \hspace{5em} if $d_l > 0$ then

21: \hspace{6em} $\lambda^-_l \leftarrow \frac{-x_{j-1,l}}{d_l}$

22: \hspace{5em} else if $d_l < 0$ then

23: \hspace{6em} $\lambda^-_l \leftarrow \frac{-x_{j-1,l}}{d_l}$

24: \hspace{5em} else

25: \hspace{6em} $\lambda^-_l \leftarrow 0$

26: \hspace{5em} end if

27: \hspace{4em} end for

28: \hspace{3em} $\Lambda \leftarrow \max \{ \lambda^-_l : 0 < l \leq |\mathcal{L}| \}$

29: \hspace{3em} Set $\lambda \leftarrow \text{a random sample from Uniform } [\Lambda, \bar{\lambda}] \text{ distribution}$

30: \hspace{3em} while $X_{j-1} + \lambda d \notin Q^M_{l(h)}$ do

31: \hspace{4em} if $\lambda \geq 0$ then

32: \hspace{5em} $\bar{\lambda} \leftarrow \lambda$

33: \hspace{4em} else

34: \hspace{5em} $\Lambda \leftarrow \lambda$

35: \hspace{5em} end if

36: \hspace{4em} $\lambda \leftarrow \text{a random sample from Uniform } [\Lambda, \bar{\lambda}] \text{ distribution}$

37: \hspace{4em} end while

38: \hspace{3em} $X_j \leftarrow X_{j-1} + \lambda d$

39: \hspace{3em} $j \leftarrow j + 1$

40: \hspace{3em} end while

41: $Q_n(h) \leftarrow X_M$

42: end for

43: return $Q_n$ \hspace{1cm} \triangleright $Q_n$: sampled TPM for the $n^{th}$ experiment

44: end for
Table 5.4: Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for women in base case when hypoglycemia is considered as a transient symptom.

<table>
<thead>
<tr>
<th>NMDP-optimal policy</th>
<th>Sampling approach</th>
<th>Worst-case approach</th>
<th>Minimum</th>
<th>Mean</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Γ = 1</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 2</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 3</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 4</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 5</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 6</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 7</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 8</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 9</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
<tr>
<td>Γ = 10</td>
<td>68.41111</td>
<td>68.4450</td>
<td>68.5083</td>
<td>68.5759</td>
<td></td>
</tr>
</tbody>
</table>

Tables 5.3 and 5.4 shows the performance of RMDP-optimal policies for the case when hypoglycemia is modeled as a transient symptom for males and females under the sampling approach based on 10,000 sampled TPMs for each gender, and worst-case approach. We provide maximum, minimum, and mean values for each RMDP-optimal policy obtained by the sampling approach.

We found that for each RMDP-optimal policy, the minimum expected QALYs prior to the first event generated by the sampling approach is significantly greater than the value function found by solving the RMDP-TM. Therefore, the worst-case solution appears to be very conservative. Furthermore, the NMDP-optimal policy generated the largest mean expected QALYs prior to first event as compared to other RMDP-optimal policies.

From the worst-case approach we found that the RMDP-optimal policy can save as much as 0.008 QALYs per patient for males, and 0.004 QALYs per patient for females as compared to the NMDP-optimal policy. To put this in practice, this result is comparable to the scale of health benefits provided by a number of highly effective population-based preventive and
Figure 5.9: The NMDP-optimal and RMDP-optimal policies performance comparisons based on the sampling approach and the worst-case approach when hypoglycemia is modeled as a transient symptom.
treatment strategies. For example, the use of the 23-valent pneumococcal vaccine to prevent
disease in elderly patients has been associated with a QALY gain of 0.003 per patient [134]. The
vaccination against measles and rubella, which has an estimated benefit per person of 0.008
QALYs [135].

Next, we investigated the case where hypoglycemia is considered as an event (an absorbing
state). We found that the RMDP-TM can not be solved exactly by solving the relaxed RMDP-
TM because the optimal policy of nature was not stationary. Based on Proposition 5.4.3 we
generated the upper and lower bounds of the optimal solution of this alternative case, and
results are shown in Table 5.5. The gap between the lower and upper bounds is less than 0.0075
QALYs per patient for males and 0.0019 QALYs per patient for females which suggests that the
approximate method provides a good approximate RMDP-optimal policies.

Table 5.5: Comparison of the expected worst-case QALYs prior to first event for glycemic
control case study when hypoglycemia is considered as an event. Upper bounds are obtained
by solving RMDP-TMs approximately using Algorithm 4, and lower bounds are obtained by
solving the relaxed RMDP-TMs exactly using Algorithm 2.

| Uncertainty budget | Model | Men | | | Women |
|---------------------|-------|-----|-----|-----|
|                     |       | Upper bound | Lower bound | Upper bound | Lower bound |
| 0                   | 0     | 62.5299     | 62.5299     | 65.7371     | 65.7371     |
|                     | 1     | 62.3556     | 62.3543     | 65.4433     | 65.4418     |
|                     | 3     | 62.0661     | 62.0604     | 65.0561     | 65.0544     |
|                     | 4     | 61.9805     | 61.9730     | 64.9747     | 64.9743     |
|                     | 5     | 61.9384     | 61.9383     | 64.9354     | 64.9353     |
|                     | 6     | 61.9266     | 61.9265     | 64.9171     | 64.9171     |
|                     | 7     | 61.9210     | 61.9210     | 64.9104     | 64.9104     |
|                     | 8     | 61.9190     | 61.9190     | 64.9104     | 64.9104     |
|                     | 9     | 61.9187     | 61.9187     | 64.9104     | 64.9104     |
|                     | 10    | 61.9187     | 61.9187     | 64.9104     | 64.9104     |
Table 5.6: Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for men when hypoglycemia is considered as an event.

<table>
<thead>
<tr>
<th></th>
<th>Worst-case approach</th>
<th>Sampling approach</th>
<th>Minimum</th>
<th>Mean</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMDP-optimal policy</td>
<td>61.7318</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>62.3002</td>
<td>62.5229</td>
<td>62.7069</td>
</tr>
<tr>
<td>( \Gamma = 1 )</td>
<td>61.7708</td>
<td>62.3016</td>
<td>62.5212</td>
<td>62.7042</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 2 )</td>
<td>61.8209</td>
<td>62.2982</td>
<td>62.5105</td>
<td>62.6918</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 3 )</td>
<td>61.8677</td>
<td>62.2832</td>
<td>62.4972</td>
<td>62.6818</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 4 )</td>
<td>61.9073</td>
<td>62.2484</td>
<td>62.4658</td>
<td>62.6560</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 5 )</td>
<td>61.9162</td>
<td>62.2271</td>
<td>62.4416</td>
<td>62.6298</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 6 )</td>
<td>61.9183</td>
<td>62.2179</td>
<td>62.4360</td>
<td>62.6256</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 7 )</td>
<td>61.9187</td>
<td>62.2063</td>
<td>62.4271</td>
<td>62.6178</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 8 )</td>
<td>61.9187</td>
<td>62.2062</td>
<td>62.4270</td>
<td>62.6177</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 9 )</td>
<td>61.9187</td>
<td>62.2062</td>
<td>62.4270</td>
<td>62.6177</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 10 )</td>
<td>61.9187</td>
<td>62.2062</td>
<td>62.4270</td>
<td>62.6177</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.7: Comparison of NMDP-optimal and RMDP-optimal policies on expected QALYs prior to first event based on sampling approach and worst-case approach for women when hypoglycemia is considered as an event.

<table>
<thead>
<tr>
<th></th>
<th>Worst-case approach</th>
<th>Sampling approach</th>
<th>Minimum</th>
<th>Mean</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMDP-optimal policy</td>
<td>64.5147</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65.3679</td>
<td>65.7126</td>
<td>66.0572</td>
</tr>
<tr>
<td>( \Gamma = 1 )</td>
<td>64.6675</td>
<td>65.3800</td>
<td>65.7066</td>
<td>66.0526</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 2 )</td>
<td>64.8600</td>
<td>65.3635</td>
<td>65.6777</td>
<td>66.0345</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 3 )</td>
<td>64.9052</td>
<td>65.3312</td>
<td>65.6385</td>
<td>65.9974</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 4 )</td>
<td>64.9095</td>
<td>65.3207</td>
<td>65.6210</td>
<td>65.9767</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 5 )</td>
<td>64.9104</td>
<td>65.3011</td>
<td>65.5916</td>
<td>65.9394</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 6 )</td>
<td>64.9104</td>
<td>65.3011</td>
<td>65.5916</td>
<td>65.9394</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 7 )</td>
<td>64.9104</td>
<td>65.3011</td>
<td>65.5916</td>
<td>65.9394</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 8 )</td>
<td>64.9104</td>
<td>65.3011</td>
<td>65.5916</td>
<td>65.9394</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 9 )</td>
<td>64.9104</td>
<td>65.3011</td>
<td>65.5916</td>
<td>65.9394</td>
<td></td>
</tr>
<tr>
<td>( \Gamma = 10 )</td>
<td>64.9104</td>
<td>65.3011</td>
<td>65.5916</td>
<td>65.9394</td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.10: The NMDP-optimal and RMDP-optimal policies performance comparisons based on the sampling approach and the worst-case approach when hypoglycemia is modeled as an event.
We compared the performance of the NMDP-optimal policy and approximate RMDP-optimal policies which were obtained from solving the RMDP-TM using the approximate algorithm. Tables 5.6 and 5.7 show the results by using the sampling approach based on 10,000 sampled TPMs for each gender, and worst-case approach. We still found that the NMDP-optimal policy generated the largest mean expected QALYs prior to first event as compared to all RMDP-optimal policies in the sampling approach. However, the RMDP-optimal policy can save as much as 0.18 QALYs per patient for males, and 0.40 QALYs per patient for females as compared to the NMDP-optimal policy under the worst-case approach. This result is two orders of magnitude greater than the benefit per person of the previous mentioned vaccination prevention programs, and also an order of magnitude greater than the use of aspirin for secondary prevention of myocardial infarction in 45-year-old men, which has been estimated to provide a QALY gain of 0.04 per patient [136]. These benefits are on the same scale as the use of statins in the secondary prevention of coronary artery disease, estimated at 0.25 QALYs per patient [137].

By comparing results obtained by sampling approach shown in Figures 5.9 and 5.10, we found that variations on the value function are similar for all the RMDP-optimal policies and the NMDP-optimal policy, and RMDP-optimal policies perform better than NMDP-optimal policy under the worst case regardless of how the hypoglycemia is modeled. This is reasonable because the objective of the robust formulation is to maximize the worst case instead of minimizing the variance in the optimal value function when TPMs vary in the uncertainty set.

5.5.5 Sensitivity of Worst-Case Performance to Model Parameters

Based on the results from Figures 5.9 and 5.10, we note that the difference in the worst-case value functions using the NMDP-optimal policy and RMDP-optimal policies can vary significantly. The worst-case expected QALYs prior to the first event for each state \( s_t \in \mathcal{L} \times \mathcal{M} \) under the NMDP-optimal policy and an RMDP-optimal policy were obtained using the following two
equations:

\[ v_{\text{NMDP-WC}}^N(s_t) = r_t(s_t, \alpha_{\text{NMDP}}^t) + (1 - p_{\text{E}}^t(s_t, \alpha_{\text{NMDP}}^t)) \lambda \times \min_{q_{t, \ell_t} \in Q_{\text{IMUB}}} \sum_{\ell_{t+1} \in L} q_{t, \ell_t}(\ell_{t+1}) v_{\text{NMDP-WC}}^{\ell_{t+1}}(m_{t+1}(\alpha_{\text{NMDP}}^t)), \]  

(5.72)

\[ v_{\text{RMDP-WC}}^R(s_t) = r_t(s_t, \alpha_{\text{RMDP}}^t) + (1 - p_{\text{E}}^t(s_t, \alpha_{\text{RMDP}}^t)) \lambda \times \min_{q_{t, \ell_t} \in Q_{\text{IMUB}}} \sum_{\ell_{t+1} \in L} q_{t, \ell_t}(\ell_{t+1}) v_{\text{RMDP-WC}}^{\ell_{t+1}}(m_{t+1}(\alpha_{\text{RMDP}}^t)). \]  

(5.73)

If the optimal solutions of the minimization problems in Equations (5.72) and (5.73) are \( q_{t, \ell_t}^{\text{NMDP-WC}} \) and \( q_{t, \ell_t}^{\text{RMDP-WC}} \), respectively, then the difference in the worst-case expected QALYs prior to the first event is given by:

\[ \Delta v_t(s_t) = v_{\text{RMDP-WC}}^t(s_t) - v_{\text{NMDP-WC}}^t(s_t) \]

\[ = [r_t(s_t, \alpha_{\text{RMDP}}^t) - r_t(s_t, \alpha_{\text{NMDP}}^t)] \]

\[ + [(1 - p_{\text{E}}^t(s_t, \alpha_{\text{RMDP}}^t)) \lambda \sum_{\ell_{t+1} \in L} q_{t, \ell_t}^{\text{RMDP-WC}}(\ell_{t+1}) v_{\text{RMDP-WC}}^{\ell_{t+1}}(m_{t+1}(\alpha_{\text{RMDP}}^t))] \]

\[ - (1 - p_{\text{E}}^t(s_t, \alpha_{\text{NMDP}}^t)) \lambda \sum_{\ell_{t+1} \in L} q_{t, \ell_t}^{\text{NMDP-WC}}(\ell_{t+1}) v_{\text{NMDP-WC}}^{\ell_{t+1}}(m_{t+1}(\alpha_{\text{NMDP}}^t))], \]  

(5.74)

for \( t = 1, 2, \ldots, T - 1 \) and \( \Delta v_T(s_T) = 0 \). Equation 5.74 shows that the difference between \( v_{\text{RMDP-WC}}^t(s_t) \) and \( v_{\text{NMDP-WC}}^t(s_t) \) depends on differences between the NMDP-optimal policy and the RMDP-optimal policy. If there is no difference between the two policies, then \( \Delta v_t(s_t) \) equals 0. If there are differences between the two policies, then \( \Delta v_t(s_t) \) depends on the difference in the one-period reward and the event probability under the NMDP-optimal policy and the RMDP-optimal policy. To evaluate this empirically we performed additional experiments to study how the hypoglycemia probability influences the variation in the value function under the worst case.

For these experiments we solve the RMDP-TM without the uncertainty budget to estimate the maximum variation. We set the risk of having hypoglycemia to be the base case parameter times an adjustment factor which ranges from 0 to 1. Therefore, when the adjustment factor equals 0, the model is equivalent to the base case where hypoglycemia is modeled as a transient
Figure 5.11: The difference in the expected worst-case QALYs prior to the first event obtained from the NMDP-optimal policy and the RMDP-optimal policy without the uncertainty budget constraint. The adjustment factor is a parameter ranging from 0 to 1. When the adjustment factor equals zero, the model is equivalent to the base case where hypoglycemia is modeled as a transient symptom, and when the adjustment factor equals 1, the model is equivalent to the alternative case when hypoglycemia is modeled as an event.
symptom, and when the adjustment factor equals 1, the model is equivalent to the alternative case when hypoglycemia is modeled as an event. Figure 5.11 shows an increasing trend in the difference in the expected worst-case QALYs prior to the first event obtained using the NMDP-optimal policy and the RMDP-optimal policy without the uncertainty budget constraint. Therefore, in the glycemic control case study, the hypoglycemia probability is one factor which will affect the difference in the worst-case value functions.

5.6 Conclusions

We presented a general RMDP-TM for medical treatment decisions in which the uncertainty in transition probabilities was taken into account. In particular, we focused on the uncertainties in transition probabilities among the metabolic states. This RMDP-TM with a general uncertainty set was difficult to solve. Therefore, we relaxed the RMDP-TM and incorporated a new uncertainty set model named IMUB so that the relaxed RMDP-TM could be solved exactly using the robust dynamic programming algorithm. A more important feature of the IMUB is that the size of the uncertainty set could be controlled so that the conservativeness of the RMDP-optimal solution could be controlled. We provided an LP formulation of the inner problem which could be used to solve the relaxed RMDP-TM efficiently. The numerical experiments showed that the computation time could be reduced by 90%. Moreover, we provided a fast algorithm to solve the inner problem. In addition, we proposed upper and lower bounds for the RMDP-TM in Proposition 5.4.3, where the upper bound was obtained by solving the RMDP-TM using an approximate algorithm, and the lower bound was obtained by solving the relaxed RMDP-TM exactly. The approximate algorithm was as efficient as solving the MDP model using backward induction. In Theorem 5.4.1, we provided and proved sufficient conditions under which the upper bound and the lower bound are the same, namely, the RMDP-TM could be solved by solving the relaxed RMDP-TM. Our numerical experiments showed that in the base case model, where hypoglycemia is considered as a transient symptom, the optimal solution of the relaxed RMDP-TM is also the optimal solution of the RMDP-TM.
Moreover, we provided properties of the uncertainty budget. Proposition 5.4.4 provided an upper bound on the uncertainty budget for the RMDP-TM above which any increase in the uncertainty budget will not affect the RMDP-optimal solution. This bound is one of the factors that can be used to predict the sensitivity of the optimal solution when the TPM is varied in the uncertainty set.

We also proposed to use a bi-criteria optimization model, called BCRMDP, based on the weighted sum average method to address the issue of conservativeness of the robust solution. We proved that when the uncertainty set of the TPM satisfies the state rectangular property, the relaxed BCRMDP is equivalent to a relaxed RMDP-TM. Thus the methods we discussed apply equally well to the bi-criteria version of the problem.

We present a new framework to compare RMDP-optimal policies based on both the sampling approach and the worst-case approach. We found that the variations in the value function are similar for all the RMDP-optimal policies and the NMDP-optimal policy regardless of how the hypoglycemia is modeled. The worst-case results are very conservative as compared to the results generated by the sampling approach. The differences in the worst-case value function under the NMDP-optimal policy and RMDP-optimal policies are affected by policy differences and differences in the one-period reward and the event probability. Based on the numerical experiments for the glycemic control problem, we found that the RMDP-optimal policies can provide benefits to a proportion of patients who experience the worst HbA1c transitions, and the magnitude of these benefits is comparable to some well-known prevention programs. However, for the general population, the NMDP-optimal policy performs better than the RMDP-optimal policies. This observation holds true for both genders under the base case as well as the alternative RMDP-TM where hypoglycemia is considered as an event.
Conclusions

In this dissertation we first developed and validated a new population-based glycemic control model based on an HbA1c Markov chain. We used this model to evaluate and compare different treatment policies. Then, we developed an MDP model, based on the glycemic control model, to optimize the treatment decisions for glycemic control. Numerical results provided insights into the difference between the optimal policy, the treatment plan used in practice, and the current treatment guidelines. Finally, we extended the MDP model to an RMDP-TM where uncertainty in TPMs is taken into account. We also extended the application of this RMDP-TM to general medical treatment decisions. The following is a summary of some of the most significant findings presented in Chapters 3, 4, and 5.

In Chapter 3, we developed and validated a new population-based glycemic control Markov model that simulates the natural variation in HbA1c progression. The model was calibrated using a U.S. data set of privately insured individuals diagnosed with type 2 diabetes. We compared treatment intensification of metformin monotherapy with sulfonylurea, dipeptidyl peptidase-4 inhibitor, glucagon-like peptide-1 receptor agonist, or insulin from diagnosis to first diabetes complication (ischemic heart disease, myocardial infarction, congestive heart failure, stroke, blindness, renal failure, amputation) or death. According to our model, all regimens resulted in similar LYs and QALYs regardless of glycemic control goal, but the regimen with
sulfonylurea incurred significantly lower cost per QALY and resulted in the longest time to insulin dependence. These findings hold for both observed effects of medications from real-world data and RCT results.

In Chapter 4, we presented a finite-horizon, discrete-state MDP model to determine when and in what order hyperglycemia lowering medications should be initiated from diagnosis to the first event, based on the Markov chain developed in Chapter 3. The goal of our model is to maximize a patient’s expected QALYs prior to the first event. Based on the different opinions on hypoglycemia, which is an important side effect associated with some hyperglycemia lowering medications such as sulfonylurea and insulin, we provide two approaches to model it. The first approach considers hypoglycemia as a transient symptom, and the second approach considers the hypoglycemia as an event. We found that the optimal policies between the males and females are similar when hypoglycemia is modeled as a transient symptom, and are different when hypoglycemia is considered as an event. Compared with the treatment guideline, which is only determined by a patient’s HbA1c level, the optimal policy, which considers a patient’s gender, age, HbA1c, and medication effects and side effects simultaneously, results in higher expected QALYs prior to first event, and the magnitude of the improvement depends on how the hypoglycemia is modeled. In the numerical experiment, we showed the use of optimal policy over the current ADA’s treatment goal can provide 0.09 additional QALYs per patient for women and 0.08 additional QALYs per patient for men when hypoglycemia was considered as a transient symptom; and can provide 0.84 additional QALYs per patient for women and 0.54 additional QALYs per patient for men when hypoglycemia was considered as an event. From the population point of view, these per patient savings in QALYs can result in significant QALYs savings when apply the optimal policy to all type 2 diabetes patients in the United States.

We also proved structural properties related the optimal value function of the MDP model, and sufficient conditions under which the optimal policy of the MDP model is of control-limit type. From our numerical experiments, we found that the HbA1c TPM does not satisfy the IFR property exactly but optimal policies are nevertheless of control-limit type for both genders.
when hypoglycemia is modeled as a transient symptom. However, when hypoglycemia is modeled as an event, optimal policies are not necessarily of control-limit type.

In Chapter 5, we extended the MDP presented in Chapter 4 to a robust MDP model referred to as the RMDP-TM, which considers the uncertainty in TPMs, for medical treatment decisions. We considered the same criteria of maximizing expected worst-case QALYs prior to first event as in Chapter 4. The RMDP-TM with a general uncertainty set is difficult to solve. We relaxed the RMDP-TM and incorporated a new uncertainty set model named IMUB so that the relaxed RMDP-TM can be solved exactly using the robust dynamic programming algorithm. We provided and proved an LP reformulation of the original NLP formulation of the inner problem to improve the computational efficiency in solving the relaxed RMDP-TM. We also provided a fast approximation algorithm to solve the RMDP-TM. We proved that an upper bound for the optimal solution of the RMDP-TM can be obtained by solving the RMDP-TM approximately by using the approximation algorithm, and a lower bound for the optimal solution of the RMDP-TM can be obtained by solving the relaxed RMDP-TM exactly. Moreover, we provided and proved sufficient conditions under which the relaxed RMDP-TM and the RMDP-TM have zero optimality gap. Computational experiments were used to demonstrate the efficiency of algorithms used to solve the relaxed RMDP-TM and the RMDP-TM.

In the case study, the RMDP-TM was used to find robust optimal treatment decisions for glycemic control. We used the sampling approach and the worst-case approach to evaluate the performance of RMDP-optimal policies. We found that the worst-case results are very conservative, however, under the worst case, the RMDP-optimal policy could save up to 0.18 QALYs per patient for males, and 0.40 QALYs per patient for females as compared to the NMDP-optimal policy when hypoglycemia is considered as an event, and the RMDP-optimal policy could save up to 0.008 QALYs per patient for males, and 0.004 QALYs per patient for females when hypoglycemia was considered as a transient symptom. These results were comparable to the scale of health benefits provided by many well-known treatment prevention programs, and will result in a significant increase in QALYs prior to first event at the population
level.

There are a number of opportunities to extend the research in this dissertation. From an application perspective, it would be valuable to combine the previously published blood pressure and cholesterol control models with the glycemic control model [138] so that the new model will have the capability of generating optimal treatment decisions to control all three risk factors simultaneously for type 2 diabetes patients. In addition, it would be useful to test the generalizability of the RMDP-TM by applying it to other medical treatment decisions reported in the literature. From a theoretical perspective, it would be interesting to provide and prove sufficient conditions under which the optimality gap between the RMDP-TM and the relaxed RMDP-TM is zero for any uncertainty budget. Furthermore, it is interesting to identify a class of RMDPs which are sensitive to uncertainties in TPMs. This can be done by investigating the impact of the reward function and/or the event probabilities on the optimal value functions.

In conclusion, many insights have been developed about the use of stochastic models, and robust MDP models in particular, for the study of medical decision making problems. The models and methods discussed in this dissertation provide a framework within which the these future research directions can be studied.
REFERENCES


APPENDICES
A  Risk Equations of Microvascular and Macrovascular complications for Patients with Type 2 Diabetes

UKPDS outcome model [116] provides the risk equations for both fatal and non fatal diabetes-related complications in the form of the proportional hazard Weibull regression model. The Weibull hazard function, cumulative hazard function and the probability of having an events between time \( t \) to \( t + 1 \) giving no events before time \( t \) are given in Equations (1), (2) and (3), respectively. The outcome model considers the end stage of these three types of microvascular complications and the renal disease is independent from the HbA1c value.

\[
h(t) = \rho t^{\rho-1}e^{\lambda+\sum \beta_i x_i} \tag{1}
\]

\[
H(t) = t^\rho e^{\lambda+\sum \beta_i x_i} \tag{2}
\]

\[
P(t < E < t + 1|E > t) = \frac{S(t) - S(t+1)}{S(t)} = 1 - e^{H(t) - H(t+1)} = 1 - e^{[t^\rho-(t+1)^\rho]e^{\lambda+\sum \beta_i x_i}} \tag{3}
\]

\( S(t) \) is the survival function and \( S(t) = e^{-H(t)} = e^{-\int_0^t h(u)du} \). The component \( e^{\beta_i} \) can be interpreted as hazard ratio for the covariate \( x_i \) (some of the covariates are time-varying such as HbA1c, total cholesterol and some are not such as gender and race). The parameters for the risk equations of each complication is shown in Table 1.

Over 6 years of follow-up of patients with type 2 diabetes in the U.K. Prospective Diabetes Study (UKPDS), 2.4% of those using metformin, 3.3% of those using a sulfonylurea, and 11.2% of those using insulin reported major hypoglycemia (requiring medical attention or admission to hospital) [126].
**Table 1:** Parameters and $\beta$ coefficients for risk equations to estimate the probability of diabetes-related complications.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Transformation</th>
<th>IHD$^1$</th>
<th>MI$^2$</th>
<th>CHF$^2$</th>
<th>STROKE</th>
<th>AMP</th>
<th>BLIND</th>
<th>RENAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td></td>
<td>-5.31</td>
<td>-4.977</td>
<td>-8.018</td>
<td>-7.163</td>
<td>-8.718</td>
<td>-6.464</td>
<td>-10.016</td>
</tr>
<tr>
<td>$\rho$</td>
<td></td>
<td>1.150</td>
<td>1.257</td>
<td>1.711</td>
<td>1.497</td>
<td>1.451</td>
<td>1.154</td>
<td>1.865</td>
</tr>
<tr>
<td>AGE$^2$</td>
<td>AGE-52.59</td>
<td>0.031</td>
<td>0.055</td>
<td>0.093</td>
<td>0.085</td>
<td>0.069</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENDER*</td>
<td>1 for female</td>
<td>-0.471</td>
<td>-0.826</td>
<td>-0.516</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RACE$^3*$</td>
<td></td>
<td>-1.312</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMOKE*</td>
<td>1 for smoker</td>
<td>0.346</td>
<td>0.355</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI$^2$</td>
<td>BMI-27.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBA1C (%)</td>
<td>HBA1C-7.09</td>
<td>0.125</td>
<td>0.118</td>
<td>0.157</td>
<td>0.128</td>
<td>0.435</td>
<td>0.221</td>
<td></td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>(SBP-135.09)/10</td>
<td>0.098</td>
<td>0.101</td>
<td>0.114</td>
<td>0.276</td>
<td>0.228</td>
<td>0.404</td>
<td></td>
</tr>
<tr>
<td>Lipid Ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln(Lipid Ratio)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PVD$^2*$</td>
<td>1 for having PVD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATRFIB$^1$</td>
<td>1 for having PVD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IHD$^*$</td>
<td>1 for history of IHD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHF$^*$</td>
<td>1 for history of CHF</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BLIND$^*$</td>
<td>1 for history of one eye blindness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2:value at diagnosis. 3:1 for Afro-Caribbean, 0 for Caucasian and Asian Indian.*: binary covariate.

**B CDC Mortality Table**

The 3-month all cause mortalities for both genders shown in Table 2 are estimated based on the CDC 2007 mortality table using the Equation (4).

\[
(1 - P(\text{die in 3-month}))^4 = (1 - P(\text{die in 1 year}))
\]

(4)

**Table 2:** The estimated 3-month all cause mortality table

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age</th>
<th>Prob of death (all cause)</th>
<th>Gender</th>
<th>Age</th>
<th>Prob of death (all cause)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>35-44</td>
<td>0.000579</td>
<td>Female</td>
<td>35-44</td>
<td>0.000342</td>
</tr>
<tr>
<td></td>
<td>45-54</td>
<td>0.001342</td>
<td></td>
<td>45-54</td>
<td>0.000788</td>
</tr>
<tr>
<td></td>
<td>55-64</td>
<td>0.002748</td>
<td></td>
<td>55-64</td>
<td>0.001674</td>
</tr>
<tr>
<td></td>
<td>65-74</td>
<td>0.006123</td>
<td></td>
<td>65-74</td>
<td>0.004074</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
<td>0.014983</td>
<td></td>
<td>75-84</td>
<td>0.010702</td>
</tr>
<tr>
<td>≥ 85</td>
<td></td>
<td>0.03441</td>
<td>≥ 85</td>
<td></td>
<td>0.030627</td>
</tr>
</tbody>
</table>

133
### Table 3: Glycosylated hemoglobin (HbA1c) used in the RMDP-TM for women

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c Range</td>
<td>&lt;6</td>
<td>[6.6,6.5)</td>
<td>[6.5,6.7)</td>
<td>[7.7,7.5)</td>
<td>[7.5,8.5)</td>
<td>[8.8,8.5)</td>
<td>[8.5,9)</td>
<td>[9.9,9.5)</td>
<td>[9.5,10)</td>
<td>10</td>
</tr>
<tr>
<td>Mean HbA1c value (%)</td>
<td>5.7</td>
<td>6.25</td>
<td>6.74</td>
<td>7.24</td>
<td>7.73</td>
<td>8.23</td>
<td>8.73</td>
<td>9.22</td>
<td>9.72</td>
<td>11.73</td>
</tr>
<tr>
<td>Initial HbA1c Distribution</td>
<td>0.0771</td>
<td>0.1543</td>
<td>0.2125</td>
<td>0.18</td>
<td>0.1105</td>
<td>0.0848</td>
<td>0.0502</td>
<td>0.035</td>
<td>0.0273</td>
<td>0.0683</td>
</tr>
<tr>
<td>TPM</td>
<td>HbA1c state 1</td>
<td>0.6471</td>
<td>0.3529</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>0.1800</td>
<td>0.5200</td>
<td>0.2200</td>
<td>0.0600</td>
<td>0.0200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>0.0435</td>
<td>0.1957</td>
<td>0.4783</td>
<td>0.2174</td>
<td>0.0652</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>0.0192</td>
<td>0.0577</td>
<td>0.2500</td>
<td>0.3846</td>
<td>0.1923</td>
<td>0.0769</td>
<td>0.0192</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>0.0323</td>
<td>0</td>
<td>0.1935</td>
<td>0.2903</td>
<td>0.2258</td>
<td>0.1935</td>
<td>0.0323</td>
<td>0</td>
<td>0.0323</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>0</td>
<td>0</td>
<td>0.0370</td>
<td>0.1852</td>
<td>0.1852</td>
<td>0.2963</td>
<td>0.2222</td>
<td>0.0370</td>
<td>0.0370</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>0</td>
<td>0.0588</td>
<td>0</td>
<td>0.0588</td>
<td>0.2333</td>
<td>0.2333</td>
<td>0.1765</td>
<td>0.1176</td>
<td>0.1176</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2500</td>
<td>0.2500</td>
<td>0</td>
<td>0.5000</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1250</td>
<td>0.1250</td>
<td>0.2500</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0500</td>
<td>0.1500</td>
<td>0.0500</td>
<td>0.1000</td>
<td>0</td>
<td>0.2000</td>
</tr>
</tbody>
</table>

### Table 4: Test TPM for IFR property of the MLE of the TPM for women

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>1</td>
<td>0.3529</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>1</td>
<td>0.8200</td>
<td>0.3000</td>
<td>0.0800</td>
<td>0.0200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>1</td>
<td>0.9565</td>
<td>0.7608</td>
<td>0.2826</td>
<td>0.0652</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>1</td>
<td>0.9808</td>
<td>0.9231</td>
<td>0.6731</td>
<td>0.2884</td>
<td>0.0961</td>
<td>0.0192</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>1</td>
<td>0.9677</td>
<td>0.9677</td>
<td>0.7742</td>
<td>0.4839</td>
<td>0.2581</td>
<td>0.0646</td>
<td>0.0323</td>
<td>0.0323</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9630</td>
<td>0.7778</td>
<td>0.5926</td>
<td>0.2962</td>
<td>0.0740</td>
<td>0.0370</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>1</td>
<td>1</td>
<td>0.9412</td>
<td>0.9412</td>
<td>0.8824</td>
<td>0.6471</td>
<td>0.4117</td>
<td>0.2352</td>
<td>0.1176</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.7500</td>
<td>0.5000</td>
<td>0.5000</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.8750</td>
<td>0.7500</td>
<td>0.5000</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9500</td>
<td>0.8000</td>
<td>0.7500</td>
<td>0.6500</td>
<td>0.6500</td>
<td>0.4500</td>
</tr>
</tbody>
</table>
Table 5: Left-hand-side maximum deviation of the TPM in RMDP-TM for women.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>0.1143</td>
<td>0.1143</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>0.0536</td>
<td>0.0696</td>
<td>0.0577</td>
<td>0.0331</td>
<td>0.0195</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>0.0296</td>
<td>0.0577</td>
<td>0.0726</td>
<td>0.0599</td>
<td>0.0359</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>0.0188</td>
<td>0.0319</td>
<td>0.0592</td>
<td>0.0665</td>
<td>0.0539</td>
<td>0.0364</td>
<td>0.0188</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>0.0313</td>
<td>0</td>
<td>0.0699</td>
<td>0.0804</td>
<td>0.0740</td>
<td>0.0699</td>
<td>0.0313</td>
<td>0</td>
<td>0.0313</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>0</td>
<td>0</td>
<td>0.0358</td>
<td>0.0737</td>
<td>0.0737</td>
<td>0.0866</td>
<td>0.0789</td>
<td>0.0358</td>
<td>0.0358</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>0</td>
<td>0.0563</td>
<td>0</td>
<td>0.0563</td>
<td>0.1014</td>
<td>0.1014</td>
<td>0.0911</td>
<td>0.0770</td>
<td>0</td>
<td>0.0770</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2134</td>
<td>0.2134</td>
<td>0</td>
<td>0.2464</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1153</td>
<td>0.1153</td>
<td>0.1509</td>
<td>0.1687</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0480</td>
<td>0.0787</td>
<td>0.0480</td>
<td>0.0661</td>
<td>0.0882</td>
</tr>
</tbody>
</table>

Table 6: Right-hand-side maximum deviation of the TPM in RMDP-TM for women.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>0.3529</td>
<td>0.6313</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>0.2959</td>
<td>0.3848</td>
<td>0.3191</td>
<td>0.1829</td>
<td>0.1078</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>0.1638</td>
<td>0.3186</td>
<td>0.4012</td>
<td>0.3312</td>
<td>0.1983</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>0.1037</td>
<td>0.1761</td>
<td>0.3271</td>
<td>0.3675</td>
<td>0.2977</td>
<td>0.2013</td>
<td>0.1037</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>0.1728</td>
<td>0.1761</td>
<td>0.3271</td>
<td>0.3675</td>
<td>0.2977</td>
<td>0.2013</td>
<td>0.1037</td>
<td>0</td>
<td>0</td>
<td>0.1728</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>0</td>
<td>0</td>
<td>0.1980</td>
<td>0.4072</td>
<td>0.4072</td>
<td>0.4786</td>
<td>0.4358</td>
<td>0.1980</td>
<td>0.1980</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>0</td>
<td>0.3108</td>
<td>0</td>
<td>0.3108</td>
<td>0.5603</td>
<td>0.5603</td>
<td>0.5036</td>
<td>0.4256</td>
<td>0.4256</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.7500</td>
<td>0.7500</td>
<td>0</td>
<td>0</td>
<td>0.5000</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.6369</td>
<td>0.6369</td>
<td>0.7500</td>
<td>0.6250</td>
<td>0.6369</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2654</td>
<td>0.4349</td>
<td>0.2654</td>
<td>0.3654</td>
<td>0</td>
<td>0.4872</td>
</tr>
</tbody>
</table>

135
Table 7: NIWC TPM for women.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>No. Nonzero Element</th>
<th>Effective Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>0.5328</td>
<td>0.4672</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1.18</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>0.1264</td>
<td>0.4504</td>
<td>0.1623</td>
<td>0.1331</td>
<td>0.1278</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4.40</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>0.0139</td>
<td>0.1380</td>
<td>0.4057</td>
<td>0.1790</td>
<td>0.2635</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>4.64</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>0.0004</td>
<td>0.0258</td>
<td>0.1908</td>
<td>0.3181</td>
<td>0.1384</td>
<td>0.2035</td>
<td>0.1229</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>6.63</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>0.0010</td>
<td>0</td>
<td>0.1236</td>
<td>0.2099</td>
<td>0.1518</td>
<td>0.1236</td>
<td>0.1830</td>
<td>0</td>
<td>0.2051</td>
<td>0</td>
<td>7</td>
<td>6.88</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>0</td>
<td>0</td>
<td>0.0012</td>
<td>0.1115</td>
<td>0.1115</td>
<td>0.2097</td>
<td>0.1433</td>
<td>0.1877</td>
<td>0.2350</td>
<td>0</td>
<td>7</td>
<td>6.76</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>0</td>
<td>0.0025</td>
<td>0</td>
<td>0.0025</td>
<td>0.1330</td>
<td>0.1339</td>
<td>0.0854</td>
<td>0.0985</td>
<td>0.5433</td>
<td>0</td>
<td>7</td>
<td>6.25</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0366</td>
<td>0.0366</td>
<td>0</td>
<td>0.9268</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0097</td>
<td>0.0097</td>
<td>0.0991</td>
<td>0.2063</td>
<td>0.6752</td>
<td>5</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0020</td>
<td>0.0713</td>
<td>0.0020</td>
<td>0.0339</td>
<td>0</td>
<td>0.1118</td>
<td>0.7790</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 8: Test TPM for IFR property of the NIWC TPM for women.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>1</td>
<td>0.4672</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>1</td>
<td>0.8736</td>
<td>0.4232</td>
<td>0.2609</td>
<td>0.1278</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>1</td>
<td>0.9861</td>
<td>0.8481</td>
<td>0.4425</td>
<td>0.2635</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>1</td>
<td>0.9996</td>
<td>0.9738</td>
<td>0.7830</td>
<td>0.4648</td>
<td>0.3264</td>
<td>0.1229</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>1</td>
<td>0.9990</td>
<td>0.9990</td>
<td>0.8754</td>
<td>0.6655</td>
<td>0.5137</td>
<td>0.3901</td>
<td>0.2051</td>
<td>0.2051</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>1</td>
<td>1</td>
<td>0.9988</td>
<td>0.8873</td>
<td>0.7758</td>
<td>0.5661</td>
<td>0.4227</td>
<td>0.2350</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>1</td>
<td>1</td>
<td>0.9975</td>
<td>0.9975</td>
<td>0.9950</td>
<td>0.8611</td>
<td>0.7272</td>
<td>0.6418</td>
<td>0.5433</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9634</td>
<td>0.9268</td>
<td>0.9268</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9903</td>
<td>0.9806</td>
<td>0.8815</td>
<td>0.6752</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9980</td>
<td>0.9267</td>
<td>0.9247</td>
<td>0.8908</td>
<td>0.8908</td>
<td>0.7790</td>
</tr>
</tbody>
</table>
Table 9: Glycosylated hemoglobin (HbA1c) used in the RMDP-TM for men. HbA1c range definition at diagnosis, the mean natural HbA1c values for each HbA1c state at diagnosis (prior to initiating medication), the initial HbA1c distributions at diagnosis, and 3-month HbA1c transition probability matrices for men.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c Range</td>
<td>&lt;6</td>
<td>[6.6,6.7)</td>
<td>[6.7,7)</td>
<td>[7,7.5)</td>
<td>[7.5,8)</td>
<td>[8,8.5)</td>
<td>[8.5,9)</td>
<td>[9,9.5)</td>
<td>[9.5,10)</td>
<td>10</td>
</tr>
<tr>
<td>Mean HbA1c value (%)</td>
<td>5.69</td>
<td>6.25</td>
<td>6.73</td>
<td>7.24</td>
<td>7.74</td>
<td>8.24</td>
<td>8.74</td>
<td>9.21</td>
<td>9.73</td>
<td>11.59</td>
</tr>
<tr>
<td>Initial HbA1c Distribution</td>
<td>0.0694</td>
<td>0.1388</td>
<td>0.1968</td>
<td>0.1626</td>
<td>0.1138</td>
<td>0.0919</td>
<td>0.0619</td>
<td>0.0424</td>
<td>0.0328</td>
<td>0.0896</td>
</tr>
</tbody>
</table>

TPM

| HbA1c state 1 | 0.6667 | 0.2444 | 0.0889 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HbA1c state 2 | 0.1346 | 0.5000 | 0.3654 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HbA1c state 3 | 0.0794 | 0.2222 | 0.3810 | 0.2857 | 0 | 0.0317 | 0 | 0 | 0 | 0 |
| HbA1c state 4 | 0.0175 | 0.0877 | 0.2807 | 0.3860 | 0.1579 | 0.0702 | 0 | 0 | 0 | 0 |
| HbA1c state 5 | 0 | 0.0465 | 0.2093 | 0.2326 | 0.1628 | 0.0930 | 0.0233 | 0 | 0 | 0 |
| HbA1c state 6 | 0 | 0 | 0.0800 | 0.2000 | 0.3600 | 0.2000 | 0.0400 | 0 | 0 | 0.0400 |
| HbA1c state 7 | 0.1071 | 0 | 0.0357 | 0.1071 | 0.0714 | 0.1429 | 0.2143 | 0.2500 | 0.0357 | 0.0357 |
| HbA1c state 8 | 0 | 0.0833 | 0 | 0.0833 | 0.2500 | 0.1667 | 0 | 0.2500 | 0.0833 | 0.0833 |
| HbA1c state 9 | 0.0556 | 0.0556 | 0 | 0.0556 | 0.1667 | 0.1111 | 0.1111 | 0.2222 | 0.0556 | 0.1667 |
| HbA1c state 10 | 0 | 0 | 0.0588 | 0.1176 | 0.0588 | 0.1765 | 0.1176 | 0.0588 | 0.0882 | 0.3236 |

Table 10: Test TPM for IFR property of the MLE of the TPM for men.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>1</td>
<td>0.3333</td>
<td>0.0889</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>1</td>
<td>0.8654</td>
<td>0.3654</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>1</td>
<td>0.9206</td>
<td>0.6984</td>
<td>0.3174</td>
<td>0.0317</td>
<td>0.0317</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>1</td>
<td>0.9825</td>
<td>0.8948</td>
<td>0.6141</td>
<td>0.2281</td>
<td>0.0702</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>1</td>
<td>1</td>
<td>0.9535</td>
<td>0.7442</td>
<td>0.5116</td>
<td>0.2791</td>
<td>0.1163</td>
<td>0.0233</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9200</td>
<td>0.8400</td>
<td>0.6400</td>
<td>0.2800</td>
<td>0.0800</td>
<td>0.0400</td>
<td>0.0400</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>1</td>
<td>0.8929</td>
<td>0.8929</td>
<td>0.8572</td>
<td>0.7501</td>
<td>0.6787</td>
<td>0.5358</td>
<td>0.3214</td>
<td>0.0714</td>
<td>0.0357</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>1</td>
<td>1</td>
<td>0.9167</td>
<td>0.9167</td>
<td>0.8334</td>
<td>0.5834</td>
<td>0.4166</td>
<td>0.4166</td>
<td>0.1666</td>
<td>0.0833</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>1</td>
<td>0.9444</td>
<td>0.8888</td>
<td>0.8888</td>
<td>0.8332</td>
<td>0.6666</td>
<td>0.5555</td>
<td>0.4444</td>
<td>0.2223</td>
<td>0.1667</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9412</td>
<td>0.8236</td>
<td>0.7648</td>
<td>0.5882</td>
<td>0.4706</td>
<td>0.4118</td>
<td>0.3236</td>
</tr>
</tbody>
</table>
Table 11: Left-hand-side maximum deviation of the TPM in RMDP-TM for men.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>0.0693</td>
<td>0.0632</td>
<td>0.0418</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>0.0467</td>
<td>0.0683</td>
<td>0.0658</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>0.0336</td>
<td>0.0516</td>
<td>0.0603</td>
<td>0.0561</td>
<td>0</td>
<td>0.0218</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>0.0171</td>
<td>0.0369</td>
<td>0.0587</td>
<td>0.0636</td>
<td>0.0476</td>
<td>0.0334</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>0</td>
<td>0.0317</td>
<td>0.0612</td>
<td>0.0635</td>
<td>0.0635</td>
<td>0.0555</td>
<td>0.0437</td>
<td>0.0227</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>0</td>
<td>0</td>
<td>0.0535</td>
<td>0.0535</td>
<td>0.0789</td>
<td>0.0946</td>
<td>0.0789</td>
<td>0.0386</td>
<td>0</td>
<td>0.0386</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>0.0576</td>
<td>0</td>
<td>0.0346</td>
<td>0.0576</td>
<td>0.0480</td>
<td>0.0652</td>
<td>0.0764</td>
<td>0.0807</td>
<td>0.0346</td>
<td>0.0346</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>0</td>
<td>0.0786</td>
<td>0</td>
<td>0.0786</td>
<td>0.1232</td>
<td>0.1060</td>
<td>0</td>
<td>0.1232</td>
<td>0.0786</td>
<td>0.0786</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>0.0532</td>
<td>0.0532</td>
<td>0</td>
<td>0.0532</td>
<td>0.0866</td>
<td>0.0730</td>
<td>0.0966</td>
<td>0.0532</td>
<td>0.0866</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>0</td>
<td>0</td>
<td>0.0398</td>
<td>0.0545</td>
<td>0.0398</td>
<td>0.0644</td>
<td>0.0545</td>
<td>0.0398</td>
<td>0.0480</td>
<td>0.0791</td>
</tr>
</tbody>
</table>

Table 12: Right-hand-side maximum deviation of the TPM in RMDP-TM for men.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>0.3333</td>
<td>0.3489</td>
<td>0.2311</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>0.2578</td>
<td>0.3777</td>
<td>0.3637</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>0.1855</td>
<td>0.2853</td>
<td>0.3332</td>
<td>0.3100</td>
<td>0</td>
<td>0.1203</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>0.0947</td>
<td>0.2041</td>
<td>0.3242</td>
<td>0.3512</td>
<td>0.2631</td>
<td>0.1843</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>0</td>
<td>0</td>
<td>0.1749</td>
<td>0.3379</td>
<td>0.3509</td>
<td>0.3509</td>
<td>0.3066</td>
<td>0.2413</td>
<td>0.1252</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>0</td>
<td>0</td>
<td>0.2955</td>
<td>0.2955</td>
<td>0.4357</td>
<td>0.5229</td>
<td>0.4357</td>
<td>0.2135</td>
<td>0</td>
<td>0.2135</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>0.3184</td>
<td>0</td>
<td>0.1910</td>
<td>0.3184</td>
<td>0.2651</td>
<td>0.3602</td>
<td>0.4224</td>
<td>0.4457</td>
<td>0.1910</td>
<td>0.1910</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>0</td>
<td>0.4346</td>
<td>0</td>
<td>0.4346</td>
<td>0.6808</td>
<td>0.5860</td>
<td>0</td>
<td>0.6808</td>
<td>0.4346</td>
<td>0.4346</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>0.2941</td>
<td>0.2941</td>
<td>0</td>
<td>0.2941</td>
<td>0.4784</td>
<td>0.4035</td>
<td>0.4035</td>
<td>0.5337</td>
<td>0.2941</td>
<td>0.4784</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>0</td>
<td>0</td>
<td>0.2198</td>
<td>0.3010</td>
<td>0.2198</td>
<td>0.3561</td>
<td>0.3010</td>
<td>0.2198</td>
<td>0.2649</td>
<td>0.4370</td>
</tr>
</tbody>
</table>
Table 13: NIWC TPM for men.

| HbA1c State  | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | No. Nonzero | Effective |
|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----| Element | Budget |
| HbA1c state 1| 0.5974 | 0.1812 | 0.2214 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 2.57 |
| HbA1c state 2| 0.0879 | 0.4317 | 0.4804 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 2.32 |
| HbA1c state 3| 0.0458 | 0.1706 | 0.3207 | 0.3109 | 0 | 0.1520 | 0 | 0 | 0 | 0 | 5 | 4.08 |
| HbA1c state 4| 0.0004 | 0.0508 | 0.2220 | 0.3224 | 0.1499 | 0.2545 | 0 | 0 | 0 | 0 | 6 | 5.17 |
| HbA1c state 5| 0 | 0.0148 | 0.1481 | 0.1691 | 0.1691 | 0.1073 | 0.2432 | 0.1485 | 0 | 0 | 7 | 6.62 |
| HbA1c state 6| 0 | 0 | 0.0265 | 0.0265 | 0.1211 | 0.2654 | 0.1211 | 0.1859 | 0 | 0.2535 | 7 | 6.68 |
| HbA1c state 7| 0.0495 | 0 | 0.0011 | 0.0495 | 0.0234 | 0.0777 | 0.1379 | 0.2074 | 0.2267 | 0.2267 | 9 | 8.53 |
| HbA1c state 8| 0 | 0.0047 | 0 | 0.0047 | 0.1268 | 0.0607 | 0 | 0.1268 | 0.1583 | 0.5180 | 7 | 6.17 |
| HbA1c state 9| 0.0024 | 0.0024 | 0 | 0.0024 | 0.0801 | 0.0381 | 0.0381 | 0.1256 | 0.0660 | 0.6450 | 9 | 8.03 |
| HbA1c state 10| 0 | 0 | 0.0190 | 0.0631 | 0.0190 | 0.1121 | 0.0631 | 0.0190 | 0.0402 | 0.6644 | 8 | 7.78 |

Table 14: Test TPM for IFR property of the NIWC TPM for men.

<table>
<thead>
<tr>
<th>HbA1c State</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HbA1c state 1</td>
<td>1</td>
<td>0.4026</td>
<td>0.2214</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 2</td>
<td>1</td>
<td>0.9121</td>
<td>0.4804</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 3</td>
<td>1</td>
<td>0.9542</td>
<td>0.7836</td>
<td>0.4629</td>
<td>0.1520</td>
<td>0.1520</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>HbA1c state 4</td>
<td>1</td>
<td>0.9996</td>
<td>0.9488</td>
<td>0.7268</td>
<td>0.4044</td>
<td>0.2545</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 5</td>
<td>1</td>
<td>1</td>
<td>0.9852</td>
<td>0.8371</td>
<td>0.6680</td>
<td>0.4990</td>
<td>0.3917</td>
<td>0.1485</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HbA1c state 6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9735</td>
<td>0.9470</td>
<td>0.8259</td>
<td>0.5605</td>
<td>0.4394</td>
<td>0.2535</td>
<td>0.2535</td>
</tr>
<tr>
<td>HbA1c state 7</td>
<td>1</td>
<td>0.9505</td>
<td>0.9505</td>
<td>0.9494</td>
<td>0.8999</td>
<td>0.8765</td>
<td>0.7988</td>
<td>0.6609</td>
<td>0.4534</td>
<td>0.2267</td>
</tr>
<tr>
<td>HbA1c state 8</td>
<td>1</td>
<td>1</td>
<td>0.9953</td>
<td>0.9953</td>
<td>0.9906</td>
<td>0.8638</td>
<td>0.8031</td>
<td>0.8031</td>
<td>0.6763</td>
<td>0.5180</td>
</tr>
<tr>
<td>HbA1c state 9</td>
<td>1</td>
<td>0.9976</td>
<td>0.9952</td>
<td>0.9952</td>
<td>0.9928</td>
<td>0.9127</td>
<td>0.8746</td>
<td>0.8365</td>
<td>0.7110</td>
<td>0.6450</td>
</tr>
<tr>
<td>HbA1c state 10</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.9810</td>
<td>0.9179</td>
<td>0.8989</td>
<td>0.7868</td>
<td>0.7236</td>
<td>0.7046</td>
<td>0.6644</td>
</tr>
</tbody>
</table>