Optimizing Healthcare Policies through Healthcare Delivery and Insurance Design

Ting-Yu Ho

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Zelda B. Zabinsky, Chair

Shan Liu, Chair

Paul A. Fishman

Program Authorized to Offer Degree:
Industrial and Systems Engineering

University of Washington

Abstract

Optimizing Healthcare Policies through Healthcare Delivery and Insurance Design

Ting-Yu Ho

Co-Chairs of the Supervisory Committee:
Professor Zelda B. Zabinsky
Industrial and Systems Engineering

Professor Shan Liu
Industrial and Systems Engineering

The problems of efficient prevention, screening, and treatment interventions for chronic diseases and acute infections have received much attention in recent years because of the rise in American healthcare costs. Chronic diseases generally cannot be prevented by vaccines or cured by one-time medication; however, most of them are controllable through lifestyle changes, early detection, and treatment. Improving the healthcare delivery system and encouraging access and adherence to medical care provide key opportunities for better prevention and control of multi-stage chronic diseases. On the other hand, an acute outbreak of a disease can be prevented with vaccination, an effective prevention intervention; however, low vaccination rate and free-riding behavior in vaccination decisions exist. In this case, healthcare policies need to be optimally and efficiently designed to improve sequential decision making to implement interventions for chronic disease and to prevent the outbreak of acute infectious diseases. This research has two contributions: building sequential decision support models for three healthcare delivery and insurance design problems, and developing corresponding methodologies that can be used to effectively investigate these healthcare problems.

The three healthcare policy optimization problems in this dissertation are:

(i) the implementation of optimal sequential screening and treatment policy with a budget constraint for a chronic hepatitis C virus (HCV) birth-cohort to improve population health outcomes;

(ii) a long-term healthcare insurance cost-sharing policy design for patients with hypertensive heart disease to minimize public/private insurer's and insurants' costs;

(iii) the design of a vaccination reimbursement and cost-sharing policy to reduce medical treatment cost and prevent the spread of seasonal influenza.

The first contribution of this dissertation is to build decision support models that represent disease transmission and progression for these three health healthcare applications. These models capture the uncertainties of disease transmission/progression and interactions between decision makers and target population. The three problems are modeled with stochastic processes, game theoretic approaches, and agent-based simulations to forecast population health, intervention effects, and costs over time; however, these problems are complicated and determining an optimal healthcare policy is challenging. Therefore, efficient methodologies need to be developed so that optimal/near-optimal healthcare policies can be easily implemented.

The second contribution is to develop three efficient methodologies that can be applied to three corresponding healthcare policy optimization problems. The methodologies include several common important characteristics. The first characteristic is to approximate models to alleviate the curse of dimensionality. These methodologies focus on building low-fidelity optimization problems or models so that a healthcare policy with good quality can be easily calculated. The low-fidelity approximation allows decision makers to efficiently identify a near-optimal healthcare intervention policy. The second characteristic is to use the structure of the problems to bound or guarantee an optimal solution so that these methodologies can be extended to large-scale healthcare management problems and further improve solution quality. The second objective focuses on incorporating global optimization techniques or optimal control theorems into methodologies to provide an optimal or near-optimal solution with improved computation time. The improvement allows decision makers to trade-off computation effort with policy quality in large-scale healthcare management problems.

These decision support models:

(i) provide useful insights regarding HCV elimination to maximize overall health outcomes measured by quality-adjusted life-years (QALYs) for different annual budgets and birth-cohorts

in baby boomers;

(ii) design flexible and differential cost-sharing provisions for specific patient groups to incent the patients with multi-stage chronic diseases for promoting healthcare service choices;

(iii) enable a public/private insurer to set optimal insurance policies, including reimbursement and cost-sharing rate, to maximize its utility which is in terms of medication and vaccination cost and social benefit during the flu season.

The illustrations of intervention implementation for these three diseases pave the way for new, and more efficient methods for preventing and controlling other disease problems. Also, the models and methodologies applied to these three healthcare policies provide contributions to understanding and future implementations to healthcare delivery and insurance design.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to deeply express my sincere appreciation to Profesor Zelda B. Zabinsky and Professor Shan Liu for invaluable academic and life guidance. They provided warm encouragement and patient advising throughout my research. I would like to extend my thanks to Professor Paul A. Fishman for providing constructive insight. I would also like to thank my committee member Professor Abraham D. Flaxman for his helpful comments and willingness in the final defense. I would also wish to thank Dr. Patty Buchanan for sharing her useful and helpful life experience. My research presented in this dissertation was supported by College of Engineering Dean Fellowship, the Department of Industrial and Systems Engineering at the University of Washington, and the National Science Foundation. My gratitude also goes to my colleagues Aven Samareh, Hasan Manzour, Jue Gong, and Larissa Petroianu for their support and friendship throughout my doctoral program.

Moving to Seattle for my Ph.D. education not only brought me a new life but also great friends. I would like to thank all my Taiwanese friends for being with me in my wonderful times in Seattle. My special thanks go to my dear friends and roommates Stephen Hu and Francis Ray Lin for sharing the most remarkable moments of my time. I also would like to thank Hao-Hsiang Wu, An-Tsu Chen, Professor Hao Huang, and Professor Shi-Chung Chang for their supports and helps in my professional career.

My mother Su-Chen Tsai, my father Cheng-Tsung Ho, and my sister Yihsin Ho have always been passionate supporters of my dreams. I would like to express my deepest gratitude for their encouragement and unconditional love. I also would like to thank my mother-in-law, Kuiying Chen and Su-Hwa Yang Teng for their supports when I was short-handed.

Neither my education nor my life in Seattle and St. Louis would have been so wonderful without Frances Kuo, my perfect wife. I thank her for being with me and bringing me two sons, Kayson Ho and Edmund Ho, in my journey. You make my life more completed than I ever thought it could be.

DEDICATION

to Frances, my perfect wife

Chapter 1

INTRODUCTION

Prevention, screening, and treatment interventions for chronic diseases and acute infections account for primary drivers of U.S. health expenditures [5, 67]. Chronic diseases generally cannot be prevented by vaccines or cured by one-time medication; however, most of them are controllable through lifestyle changes, early detection, and treatment. Improving the healthcare delivery system and encouraging access and adherence to medical care provide key opportunities for better prevention and control of multi-stage chronic diseases [108, 155]. In contrast to chronic disease, an outbreak of acute infection can often be prevented by vaccination before the disease starts to spread in a population. However, despite the known success of vaccines in reducing morbidity and complications associated with the disease, immunization rates remain relatively low [30]. Given these considerations, in order to maximize health population outcomes, sequential budget allocation for multiple interventions need to be improved from the healthcare systems' perspective, and vaccination and medical care seeking behavior need to be properly incentivized from a self-interest individuals' perspective.

Identifying optimal healthcare intervention policies that enable maximum population/individual health outcomes may lead to a significant societal impact on prevention and control of chronic and infectious diseases (e.g., chronic hepatitis C and HIV) [47, 96]. It is, however, computationally challenging due to complicated disease transmission/evolution models and the non-linear and dynamic property of the optimization problems. Research on how to efficiently identify near-optimal/optimal healthcare interventions for disease management is therefore necessary.

Three real-world healthcare problems studied in this dissertation are multi-stage chronic diseases, i.e., HCV and hypertension, and acute infectious disease, i.e., seasonal influenza. We first study the prevention and management of chronic disease from a system perspective. For chronic disease that involves multiple interventions from a government/public payer, (e.g., chronic hepatitis C (HCV), depression, and opioid epidemic), dynamic resource allocation problems have been

investigated in recent years. Operations research (OR) tools such as mathematical programming, simulation optimization, Markov decision process (MDP) and system dynamics models are becoming more common in healthcare policy modeling research. One of the desired goals for healthcare policymakers is to improve population health outcomes through reductions in disease prevalence, incidence, morbidity, and mortality. However, a dynamic resource allocation problem is challenging because it involves making a sequence of intervention decisions ahead of budgetary planning cycle over multiple time-periods. Factors that impede optimal resource allocation include complex population-level disease dynamics, resource constraints, intervention priority, and uncertain intervention effects. Therefore, it is necessary to develop an efficient solution algorithm to solve sequential decision problems for a large population.

Second, this research focuses on chronic diseases that can be prevented and managed with an individual's treatment-seeking behavior (e.g., hypertension and diabetes). The cost-sharing policies in a health insurance scheme have been employed as an effective mechanism to incent the use of medications toward lower medical expenditures; however, the level of cost-sharing may influence patients' willingness to manage chronic diseases [104, 133]. While various healthcare mechanisms are designed and analyzed to reduce moral hazard or eliminate insurance overhead [53, 54, 167], it is necessary to consider long-term interaction between insurer and patients, and the effect of health insurance design on dynamic health outcomes and patient's behavior.

The third healthcare problem is a vaccine-preventable outbreak of acute infectious disease, e.g., seasonal influenza. The best way to prevent and control outbreaks is to encourage vaccination uptake before infection and treatment uptake once getting infected [71, 123]; however, critical vaccination coverage for acute infectious disease usually cannot be achieved by voluntary vaccination. Moreover, increased out-of-pocket expenses for patients are associated with lower rates of treatment uptake; on the other hand, lower cost-sharing may result in a lower vaccination rate because of the relatively low medical cost for treatment. To improve the vaccination coverage rate and further prevent the spread of acute infectious disease, an incentive-based mechanism between the insurer/public payer and patients is necessary.

This dissertation has two primary objectives for each above mentioned real-world application:

(i) to design sequential decision support models;

(ii) to develop corresponding optimization methodologies to efficiently discover policies/interventions for these healthcare problems.

For each of the three healthcare policy-making problems, the research challenges and research contributions are presented with respect to the two primary research objectives as follows.

- *Design a sequential healthcare decision-making methodology for efficiently implementing optimal sequential screening and treatment policies with the budget constraint for chronic HCV birth-cohort to improve population health outcomes.* Two challenges are considered: (i) since the sequential budget allocation problem is non-linear and involves dynamic disease progression and evolutionary population, how to capture the property of the dynamic model with reasonable computational complexity of the optimization problem is challenging; (ii) with the complicated dynamic resource allocation problem, how to efficiently identify an optimal (or near-optimal) policy is a challenge.

  Contributions

  - A low-fidelity model to approximate the overall health outcomes is formulated by utilizing the structural property of Markovian disease progression.
  - Integrating the low-fidelity model into an original optimization model, a multi-fidelity rollout algorithm is designed to efficiently identify a near-optimal sequential intervention policy. The rollout algorithm with a branch-and-bound framework was then designed to provide an optimality guarantee and bounds on the optimality gap of healthcare interventions.
  - A numerical study provides useful insights that allocating $4 or $8 billion annual budgets are able to effectively eliminate HCV infection in 19 and 11 years, respectively.

- *Design a long-term optimal healthcare insurance cost-sharing policy for patients with hypertensive heart disease to minimize public/private insurer's and insurants' respective costs.* Two challenges are considered: (i) with the hierarchical strategies/moves between the insurer and the insurant, insurant's health outcome is determined by both insurer's announced cost-sharing policy and insurant's treatment action. However, the policy and the action depend on

not only the expenditure and effectiveness of healthcare services but also the uncertainty of disease progression. Considering corresponding health results associated with both insurer and insurant's current and future decision-making behaviors, how to formulate the long-term interaction between the insurer and insurants is a challenge; (ii) the insurer needs prudent decision-making tools for not only reducing long-term medical costs but improving patients' health outcomes. Both insurer and insurant make decisions with consideration of natural disease progression and each other's current and future response, which result in a complicated decision-making process. Thus, how to determine optimal cost-sharing strategies for the insurer and treatment-receiving decisions for the insurant is a challenge.

Contributions

- A Stackelberg stochastic game model was applied to derive optimal cost-sharing policies for a public/private insurer and an insurant's optimal decision on multiple treatments. This game-theoretic approach captures the insurer and insurants' long-term hierarchical interactions subject to underlying uncertainties of disease progression.
- A set of conditions was proposed to ensure the existence of monotonic optimal policies for both insurer and an insurant by examining the structural properties of the stochastic Stackelberg game model.
- The analysis results in the hypertension case show that insurants' behavior regarding medical utilization and treatment adherence complements related case studies, and the analysis sheds light on an important connection among dynamic health status, cost-sharing policy, and medical treatment adherence.

- *Design an optimal vaccination reimbursement and cost-sharing policy for reducing medical treatment cost and preventing the spread of seasonal influenza.* Two challenges are considered: (i) there exists a multi-time interaction among insurer and insurants during flu season. The untrackable population-level health status is dynamically influenced by each insurant's individual-level decisions and Influenza virus transmission. Therefore, how to model the complex interconnections between the spread of infectious disease and the dynamic interaction between the insurer and insurants is a challenge; (ii) it is computationally expensive to derive

an optimal policy due to stochastic disease transmission in a large population size network. How to design a methodology to efficiently derive the optimal vaccination reimbursement and cost-sharing policy is a challenge.

Contributions

- The dynamic interaction between a single insurer and multiple insurants was modeled as a Stackelberg vaccination game, which allows characterizes the individual-level model of the insurant by his/her own attributes (i.e., age and risk types). Coupling the game-theoretic decision-making model, an agent-based simulation model is developed to simulate population-level spreading of flu.
- A low-fidelity model was built based on a random forest regression to reduce simulation computation time, and a probabilistic branch and bound algorithm was then applied to analyze the intervention policies.
- Experimental results indicate that reimbursement and cost-sharing policies are effective approaches to encourage vaccination behavior and minimize the overall medication cost with respect to different vaccine efficacy and attack rates.

In summary, this dissertation provides insights into healthcare delivery and insurance design problems with stochastic and simulation models and three multi-fidelity optimization methodologies.

Chapter 2

BACKGROUND AND LITERATURE REVIEW

2.1 Optimized Healthcare Resource Allocation for Chronic Disease

Chronic disease management has received special attention from a limited resource allocation perspective. Deo et al. [46] studied capacity allocation problem in community-based healthcare delivery for a chronic disease and Rauner er al. [120] developed a risk-group oriented chronic disease progression model embedded within a multi-objective optimization approach. Alistar el al. [4] and Deo et al. [47] studied HIV prevention and control with a limited resource and capacity, Khademi et al. [87] examined treatment allocation under different classes of admissible policies for HIV control using an approximate dynamic programming approach, and Deo et al. [48] studied the optimal allocation of rapid diagnosis devices and location of labs and diagnostic equipment in a resource-limited setting to improve the efficiency of early infant diagnosis (EID) of HIV. Martin et al. [105] studied the optimal HCV treatment program strategy for injecting drug users. Li et al. [96] studied the optimal implementation of HCV birth-cohort screening and treatment strategies with limited budget. Ayer et al. [11] studied the optimal treatment strategies to prioritize patients in prisons. This stream of research mainly focuses on the optimal strategies for a population.

Research also focuses on organ allocation among patients with end-stage chronic diseases, e.g., liver cancer or kidney failure, [1, 19, 88, 140, 141, 166]. Moreover, research on optimal time to initiating therapy or acceptance of organs is also growing. Lee et al. [94] studied the cost-effectiveness of initiation of dialysis for patients affiliated with chronic kidney failure, Shechter et al. [130] studied the optimal time to initiate HIV therapy under ordered health states, Mason et al. [106] studied the influence of adherence on the optimal timing of statin treatment for diabetes patients, Helm et al. [74] investigated the timing of periodic monitor for glaucoma patients, Liu et al. [97] developed a framework to guide optimal patient treatment decisions for chronic diseases considering that treatment technologies improve over time, and Bertsimas et al. [20] developed models for the analysis and design of clinical trials testing to analyze chemotherapy combinations in

order to improve clinical outcomes. This stream of research mainly focuses on the optimal treatment strategies for individual patients.

2.2 Optimized Health Insurance Policy Design

A substantial literature on optimal health insurance design focuses on how to effectively control rising healthcare expenditures, welfare analysis, and spread risk. Arrow [8] studied the welfare changes associated with a change in the coinsurance rate considering the supply and demand of medical care. Rothschild and Stiglitz [122] studied the existence of equilibrium in a competitive insurance market which describes the supply and demand functions of the two participants who shop for insurance and companies that sell insurance. Designing optimal health insurance policies to make tradeoffs between risk sharing and the agency problem, e.g., moral hazard, is also discussed [7, 44, 115]. The effect of excess health insurance on the welfare loss has been estimated [58]. Many studies have collected data on the association among health insurance coverage, out-of-pocket payments and health care utilization [60, 104, 112]. Goldman and Joyce [71] showed that each 10% increase in cost-sharing results in 2% to 6% decrease in the use of prescription drugs. An important result found by Fishman et al. [60] is that unmet deductibles lower the likelihood for an individual to make an initial psychotherapy visit for treatment of depression. Mann et al. [104] showed that for patients with chronic diseases, lowering cost-sharing may improve adherence but the impact on clinical and economic outcomes is less certain.

Many researchers have attempted to address healthcare issues using game theory (e.g., global budget payment [37], contraceptive choice [36], and the physician-patient relationship [163]). Mechanism design is a mathematical tool in game theory that has been widely used to study insurance policies. In these games, a leader chooses the game's payoff structure, the other players who act as the followers receive information relevant to the payoff structure and act accordingly. The main characteristic of the mechanism design is that the leader is able to choose the mechanism that minimizes his/her expected cost [113]. In other words, the mechanism is designed to provide incentives to the players, and then can potentially improve their objectives, e.g., improving health states or reducing healthcare costs. Considering two types of cost structure, over-utilization cost, and insurance overhead, Zhu [167] studied possible reduction of structure cost via the implementation of

healthcare financing mechanisms with the participation of the government and the private insurer and establishing a Stackelberg game theoretical framework. Ellis and Manning [54] examined optimal health insurance by modeling the case of one treatment option and one prevention option and considered the tradeoff between individuals' risk reduction moral hazard; however, they include only one health treatment option and do not model correlations over time. Further Ellis et al. [53] further studied optimal insurance for the healthcare market with multiple health treatment alternatives. They examined how an optimal cost-sharing policy is affected by the correlation structure of random shocks affecting demand for health care both across options and over time.

2.3 Optimized Vaccination Incentive Policy Design

Seasonal influenza is an infectious disease caused by the influenza virus and easily spread from person to person. With age-related increased risk, the elderly and young on average have a greater risk of being infected and developing severe complications.

Seasonal influenza is, in fact, preventable by the seasonal flu vaccine, which protects against the most common flu viruses. Vaccination incentives, either rewards or punishments, have been shown in a vast amount of research for significantly improving vaccination rate. For example, [26] showed that college students were more willing to get a flu vaccine when offered a $20 reward (19% vs. 9%). [63] explored the conditions under which the free-rider problem can be overcome through the use of taxes and subsidies. Furthermore, convincing evidence has suggested that reducing out-of-pocket costs and adding incentive are effective interventions for improving vaccination coverage and overcoming vaccination hesitancy ([25, 23]). [152] even found that severe epidemics are unable to be prevented unless vaccination incentives are offered.

Flu treatments for influenza symptoms can be a second-line of defense against the spread of seasonal flu, e.g., antiviral flu drugs are taken to decrease the severity and duration of flu symptoms. Medical treatments are able to prevent serious flu complications and shorten recovery time. In some cases, they may be used to prevent flu. However, a high cost-sharing in healthcare insurance may reduce the chance that individuals will make a doctor visit. Cost-sharing is used to change the utilization of services or prescription drugs for the enrollees of public or private health insurance. The introduction of cost-sharing charge will decrease the utilization of most types of medical services

([6, 162]). [69] showed that increased cost-sharing in healthcare insurance is associated with lower rates of drug treatment, poorer adherence among existing patients, and more interruption of continuation of therapy. For example, [61] showed that unmet deductibles lower the likelihood for an individual to make an initial therapy visit for treatment of depression.

In the past decades, there is extensive literature studying how to simulate the spread of influenza and promote vaccination policy using game theory. However, most of the extant literature considers only a population's perspective on decision making. Model-based analysis of vaccination and a few theories suggested that eradicating a vaccine-preventable infectious disease is difficult or impossible [13, 65, 12, 152]). [139] analyzed the rationale of a mandatory vaccination strategy in the Susceptible-Infectious-Recovered (SIR) model using measles infection as an example. [13] developed a game theoretical interpretation based on an epidemic model of the rational exemption phenomenon, and showed that under a purely voluntary policy, rational exemption makes eradication impossible. [127] studied the effect of an immunization program promoted by the government against the propagation of a contagious infection. [14] focused on smallpox and studied the conflict between self-interest and group interest. [65] parameterized an epidemiological game-theoretic model of influenza vaccination with questionnaire data on actual perceptions of influenza and its vaccine to compare Nash equilibrium vaccination strategies driven by self-interest with utilitarian strategies. [121] applied the proposed general approach for population game to various simple vaccination games. [64] studied the roles of individual imitation behavior and population structure in vaccination. [118] simulated transmission of a vaccine-preventable Susceptible-Exposed-Infectious-Recovered infection through a random, static contact network and examined vaccinating behavior based on neighborhood.

Although a wide variety of research has been studied for controlling and preventing infectious diseases using vaccination promotion, they either adopt *mandatory vaccination* or model a *population-level response* toward interventions instead of considering individual responses. [157] showed that there is a strong public health and scientific rationale for studying the interaction of disease dynamics and individual behavior. For example, fast disease spread will increase the risk perception so that people feel stronger incentives to take preventive measures. One of the focuses of this thesis is to integrate individual determinants of medical service utilization into flu epidemic modeling.

Several approaches have been applied to model a flu epidemic. One approach is the deterministic compartmental model commonly used in epidemiology, such as the SIR model based on ordinary differential equations ([86, 164, 39, 142, 40]). Although SIR model and corresponding extended models, e.g., with or without vaccination, have become important tools in analyzing the spread and control of infectious diseases by mandatory vaccinations, they are not able to incorporate spatial and temporal factors, e.g., distinct population structure, and dynamic interaction of individuals, which drive more realistic modeling results and reflect the heterogeneous environment. Another type of model for the spreading of a disease is the Cellular Automata simulation model, which takes into account spatial parameters ([15, 158]). This method involves two-dimensional cellular automaton to model location-specific characteristics of the susceptible population. In addition, the probabilistic nature of infectious disease transmission can be modeled by adopting stochastic parameters in the simulation. However, movement behavior and dynamic interactions among individuals cannot be captured in a Cellular Automaton model.

Agent-based simulation modeling has been widely employed for highly infectious disease studies due to its advanced capability of tracking the interaction among infectious and susceptible individuals in communities located in a network, and addressing the naturally stochastic nature of the transmission process ([116]). An agent-based simulation consists of a population of individual actors called agents, an environment, and a set of learning rules or adaptive processes. Each individual agent in agent-based simulation collects information from its surroundings or neighbors and uses it to determine how to act. Agent-based simulation has been successfully demonstrated in many epidemiology studies ([117, 49, 72]) and many simulation programs are available in an open-source format, including FluTE, EpiFire, GEM and GSAM, ([73]). Several studies also incorporated vaccination dynamic behavior into agent-based simulation ([93, 64, 57]).

Chapter 3

OPTIMAL HEPATITIS C SCREENING AND TREATMENT POLICIES

This chapter describes the first healthcare application: *Design a sequential healthcare decision-making methodology for efficiently implementing optimal sequential screening and treatment policies with the budget constraint for chronic HCV birth-cohort to improve population health outcomes*. A discrete-time finite-horizon budget allocation problem with HCV progression is modeled within a closed birth-cohort population. A low-fidelity approximation that preserves the population dynamics under a stationary policy is proposed to address the computational challenges associated with large-state and multiple-period dynamic decision-making problems. The low-fidelity approximation is embedded into the high-fidelity optimization model and branch-and-bound scheme to efficiently identify an optimal/near-optimal non-stationary sequential intervention policy.

Most of this chapter appeared in [80] and [81].

3.1 Model Formulation

Section 3.1.1 introduces a model for population disease management and its connection to intervention policies. To maximize overall QALYs through resource allocation, Section 3.1.2 proposes a high-fidelity optimization problem regarding intervention policies under budget constraints. Section 3.1.3 uses the structural property of the Markov model and proposes a low-fidelity model.

3.1.1 System dynamics for population disease management

Consider a disease with multiple health conditions, individuals may receive different interventions due to their awareness of the disease. The closed population in a healthcare system is divided into different groups based on the intervention they may receive. A group is indexed by $g \in \mathscr{G}$, where $\mathscr{G}$ is the set of all groups. Let $\widetilde{\mathscr{G}} \subset \mathscr{G}$ denote the set of groups waiting for interventions. The individuals in the same group $g \in \widetilde{\mathscr{G}}$ may receive the same intervention. Define the group-based health category $s \in \mathscr{S}_{HS} = \cup_{g \in \mathscr{G}} \mathscr{S}_g$, where $\mathscr{S}_{HS}$ is the set of all group-based health categories and $\mathscr{S}_g$ is the set of

group-based health categories in group g. The total number of health categories is denoted $|\mathscr{S}_{HS}|$.

The decision variable $I^s_{g,t}$ is the fraction of population in health category $s \in \mathscr{S}_g$ and $g \in \widetilde{\mathscr{G}}$ that is offered an intervention at the beginning of time period t, with

$$0 \leq I^s_{g,t} \leq 1, \forall s \in \mathscr{S}_g, g \in \widetilde{\mathscr{G}}, t \in \left\{1, \ldots, n_q\right\}. \tag{3.1}$$

For convenience, denote a sequential policy as $I = \left\{I^s_{g,t}\right\}_{s \in \mathscr{S}_g, g \in \widetilde{\mathscr{G}}, t=1,\ldots,n_q}$, where n_q is the number of time periods in the entire lifespan under a lifetime horizon, and a policy at time t as $I_t = \left\{I^s_{g,t}\right\}_{s \in \mathscr{S}_g, g \in \widetilde{\mathscr{G}}}$.

Let N^s_t be the population size in each group-based health category $s \in \mathscr{S}_g$ in time period t. The population dynamics are driven by the natural history of disease and intervention policies. As a result, given population size in each group-based health category and health intervention policy, the dynamics of the disease progression and population evolution is given by

$$\mathbf{N}_{t+1} = \mathbf{N}_t \mathbf{P}(I_t), \tag{3.2}$$

where $\mathbf{N}_t$ is the 1-by-$|\mathscr{S}_{HS}|$ row vector of population size with N^s_t in each group-based health category $s \in \mathscr{S}_{HS}$ in the beginning of time period t, $\mathbf{P}(I_t)$ is the $|\mathscr{S}_{HS}|$-by-$|\mathscr{S}_{HS}|$ matrix of transition probabilities in time t based on intervention policies I_t and natural history of the disease. An example of the matrix of transition probabilities in (3.2) for HCV is given in Section 3.4.1 and Appendix B.

3.1.2 High-fidelity optimization formulation

The goal is to maximize the discounted cumulative quality-adjusted life-years (QALYs) of the closed population. The *QALY* objective function which sums discounted QALYs over the entire lifetime horizon of the population given population size and interventions,

$$\begin{aligned} QALY(\mathbf{N}_1, I) &= \sum_{t=1}^{n_q} (1+\delta)^{-(t-1)} \mathbf{N}_t \mathbf{u}' \\ &= \sum_{t=1}^{n_q} \sum_{s \in \mathscr{S}_{HS}} (1+\delta)^{-(t-1)} u^s N^s_t, \end{aligned} \tag{3.3}$$

where $\mathbf{N}_1$ is the initial population size vector in the beginning of the first time period, n_q is the number of time periods in the entire lifetime horizon, $\mathbf{u}$ is the 1-by-$|\mathscr{S}_{HS}|$ row vector of health utility multipliers (and $\mathbf{u}'$ its transpose), the utility multiplier u^s in $\mathbf{u}$ is the health utility multiplier per time period corresponding to each group-based health category $s \in \mathscr{S}_{HS}$, and δ is the discount rate.

The disease has a chance to progress every time period, as t ranges from 1 to n_q. The disease progression probabilities in the transition probability matrix $\mathbf{P}(I_t)$ may be age-dependent, and thus $\mathbf{P}(I_t)$ may be updated as the population ages.

The budget constraint in each time period t is expressed as

$$\sum_{g \in \widetilde{\mathscr{G}}} \sum_{s \in \mathscr{S}_g} C_g^s I_{g,t}^s N_t^s \leq B_t, \ \forall t \left\{1, \ldots, n_q\right\}, \tag{3.4}$$

where C_g^s is the cost per person for receiving intervention in group $g \in \widetilde{\mathscr{G}}$ in health category $s \in \mathscr{S}_g$, and B_t is the available budget during time period t. Also, every group $g \in \widetilde{\mathscr{G}}$ is able to make its budget allocation rule for different populations within the group due to its characteristic, e.g., equitable allocation of resources, or give priority to patients with poor health category. An example of budget allocation within each group for HCV is given in Section 3.4.1.

The goal of the optimization model is to maximize the overall discounted QALYs for the population under the population's lifetime horizon. The high-fidelity QALY maximization problem HQMP is described as follows,

$$\text{HQMP:} \qquad \max_{I} QALY\left(\mathbf{N}_1, I\right) \tag{3.5}$$

$$\begin{aligned} \text{s.t.} \quad & \text{Fraction of population receiving interventions } (3.1) \\ & \text{Population dynamics } (3.2) \\ & \text{Budget constraint } (3.4) \\ & \text{Budget allocation rule within each group } g, \forall g \in \widetilde{\mathscr{G}} \end{aligned}$$

3.1.3 Low-fidelity model

Unfortunately, solving HQMP is computationally impractical due to state space explosion. To reduce computation, a value function approximation that enables us to optimize the intervention policy iteratively is proposed. A low-fidelity model is proposed to approximate the QALYs as the population evolves, using a property of an absorbing Markov chain with stationary policies. In the low-fidelity model, the transition matrix is assumed to be stationary, that is, the policies I_t are the same from time period t to the time period that all of the starting population have been absorbed (i.e., all persons are dead). The property of absorbing Markov chains allows us to quickly calculate the matrix of expected number of time periods spent in each state until absorption, beginning at time period t,

$$\mathbf{E}\left(I_t\right) = \left(\mathbf{I} - \mathbf{Q}\left(I_t\right)\right)^{-1}, \tag{3.6}$$

where $\mathbf{I}$ is the appropriately sized identity matrix and $\mathbf{Q}\left(I_t\right)$ is the submatrix of the transition probability matrix $\mathbf{P}\left(I_t\right)$ eliminating the row and column associated with the mortality state [161].

The approximated QALYs from time t with starting population $\mathbf{N}_t$ to the time that the entire population has reached the mortality state under stationary policy I_t is expressed as the Low-Fidelity

Model (LFM),

$$\begin{aligned} LFM(\mathbf{N}_t, I_t) &= \mathbf{N}_t \mathbf{E}(I_t)\,\mathbf{u}' \\ &= \sum_{s \in \mathscr{S}_{HS}} \sum_{v \in \mathscr{S}_{HS}} N_t^s u^v E^{s,v}(I_t), \end{aligned} \tag{3.7}$$

where $E^{s,v}(I_t)$ is the s,vth element of the matrix $\mathbf{E}(I_t)$, representing the expected number of time periods spent in state v, starting in state s until absorption.

The low-fidelity model differs from HQMP in several ways. First, the transition matrix is stationary, which implies that the intervention percentages are kept the same in all time periods from t onward. In the high-fidelity model, the intervention policy is allowed to change every time period. Second, since the matrix of expected number of time periods spent in each state, $\mathbf{E}(I_t)$, is estimated by fixed progression parameters, transition probabilities are age-independent and remain the same as the age of the starting population at starting time period t; in contrast, the disease progression is allowed to be age-dependent in the high-fidelity model. Third, the overall QALYs in LFM is not discounted, as done in the high-fidelity model. The number of decision variables is much reduced (to one time period only) in the low-fidelity model, and thus the optimal stationary policy that maximizes LFM is easily calculated.

3.2 Multi-Fidelity Rollout Algorithm (MF-RA)

The high-fidelity optimization model HQMP has the advantage of more accurately representing the disease progression as the population ages, and allowing time-varying intervention policies; however, it is computationally impractical to solve. The low-fidelity model, on the other hand, has the advantage of quick calculations; however, it is only an approximation to the high-fidelity model. The low-fidelity model is now embedded into HQMP and Multi-Fidelity Rollout Algorithm (MF-RA) is designed for scalable computation.

An overview of the process flow of MF-RA with the high-fidelity disease progression model and the low-fidelity approximate model is illustrated in Fig. 3.1. The initial population size vector $\mathbf{N}_1$ is updated by the high-fidelity Markovian model with natural disease progression and intervention policies (as in Equation (3.2)). The low-fidelity model is used to quickly approximate the impact of a policy decision (as in Equations (3.6) and (3.7)) and then the QALY optimization uses Equation

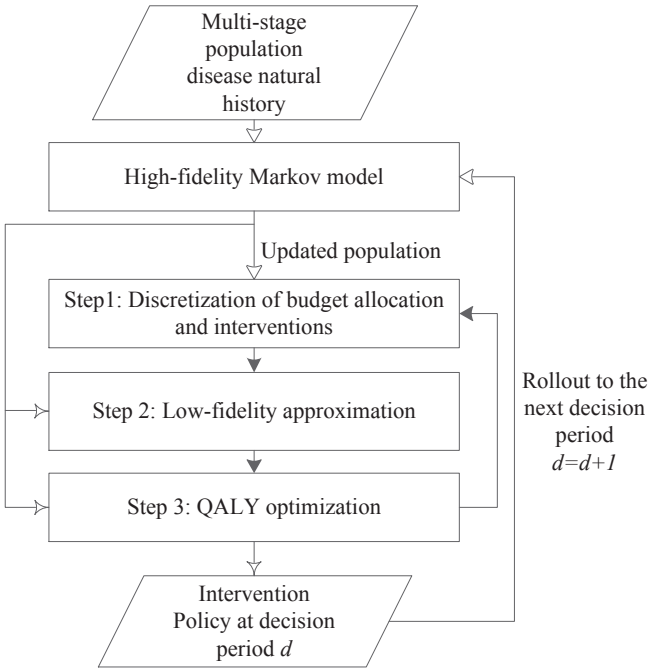


Figure 3.1: Overview of the process flow of MF-RA

(3.3) to provide a more accurate estimate of the result of the policy decision. The output of MF-RA includes the policy decisions for all decision periods, the detailed evolution of population size, the budget consumed, and the corresponding overall QALYs.

A reformulation of the budget constraints is in Section 3.2.1, and the MF-RA is presented in Section 3.2.2. The notation used in MF-RA is summarized in Table 3.1. The computational complexity and properties of MF-RA are discussed in Appendix A.

3.2.1 Budget constraint reformulation

Budget allocation policies generally do not need to change frequently. Thus, a decision period index d is introduced to distinguish the policy-changing period from time period t used for disease progression. For example, in the HCV model in Section 3.4, the time period t is quarterly (3 months) and the decision period is every two years. Let n_q and n_d be the numbers of time periods in a lifetime horizon and decision periods, respectively, and $n_{q,dec}$ be the number of time periods associated with each budget allocation policy decision. In addition, healthcare systems often have a limited planning

Table 3.1: Notation in MF-RA

Indices	
t	index of time period, $t = 1, \ldots, n_q$
d	index of decision period, $d = 1, \ldots, n_d$
k	index of discretized policy, $k = 1, \ldots, n_p$
Ak	index of policy obtained by low-fidelity approximation, $Ak = A1, \ldots, An_p$
$\mathscr{T}_d$	set of time period indices in decision period d
Input Variables	
$\mathbf{N}_1$	vector of population sizes at time period 1
$\mathbf{P}$	matrix of transition probabilities
$\mathbf{u}$	vector of health utility multiplier
B_t	Budget at time period t
C_g^s	intervention cost for health category $s \in \mathscr{S}_{HS}$ and group $g \in \widetilde{\mathscr{G}}$.
n_q	number of time periods in a lifetime horizon
n_d	number of decision periods
$n_{q,dec}$	number of time periods between policy changes
$n_{q,budget}$	number of time periods in budget-planning cycle, $n_{q,budget} = n_d \cdot n_{q,dec}$
n_{dis}	number of discretized budgets in B_t
n_p	number of budget allocation policies for different combinations of discretized interventions, $n_p = C_{|\widetilde{\mathscr{G}}|-1}^{|\widetilde{\mathscr{G}}|+n_{dis}-1}$
Decision Variables	
$x_g^{(d)}$	fraction of budget in decision period d allocated to interventions in group $g \in \widetilde{\mathscr{G}}$, and $\mathbf{x}^{(d)} = \left\{x_1^{(d)}, \ldots, x_{\lvert\widetilde{\mathscr{G}}\rvert}^{(d)}\right\}$ is the set of budget fractions, $\forall d \in \{1, \ldots, n_d\}$
$I_{g,t}^s$	fraction of population in health category $s \in \mathscr{S}_{HS}$ and group $g \in \widetilde{\mathscr{G}}$ who are offered interventions in time period $t \in \left\{1, \ldots, n_q\right\}$
Intermediate Variables	
$x_{g,k}^{(d)}$	kth combination of discretized budget fractions in decision period d allocated to interventions in group $g \in \widetilde{\mathscr{G}}$, and $\mathbf{x}_k^{(d)} = \left\{x_{1,k}^{(d)}, \ldots, x_{\lvert\widetilde{\mathscr{G}}\rvert,k}^{(d)}\right\}$ is the set of the kth combination of discretized budget fractions, $\forall k \in \left\{1, \ldots, n_p\right\}$
$I_{s,g,k}^{(t)}$	kth combination of discretized population fractions in health category $s \in \mathscr{S}_g$ who are offered interventions in group $g \in \mathscr{G}$ in time period $t \in \mathscr{T}_d$, and $I_k^{(t)} = \{I_{s,g,k}^{(t)}\}_{s \in \mathscr{S}_g, g \in \widetilde{\mathscr{G}}}$ is the set of the kth combination of discretized population fractions, $\forall k \in \left\{1, \ldots, n_p\right\}$
$x_{g,Ak}^{(d)}$	kth combination of budget fractions in decision period d allocated to interventions in group $g \in \mathscr{G}$ obtained by low-fidelity approximation following $\mathbf{x}_k^{(d')}$ in decision period $d' < d$, and $\mathbf{x}_{Ak}^{(d)} = \left\{x_{1,Ak}^{(d)}, \ldots, x_{\lvert\widetilde{\mathscr{G}}\rvert,Ak}^{(d)}\right\}$ is the set of the kth combination of budget fractions, $Ak \in \left\{A1, \ldots, An_p\right\}$
$I_{s,g,Ak}^{(t)}$	kth combination of population fractions in health category $s \in \mathscr{S}_g$ who are offered interventions in group $g \in \widetilde{\mathscr{G}}$ obtained by low-fidelity approximation in time period $t \in \mathscr{T}_d$, and $I_{Ak}^{(t)} = \left\{I_{s,g,Ak}^{(t)}\right\}_{s \in \mathscr{S}_{HS}, g \in \widetilde{\mathscr{G}}}$ is the set of the kth combination of population fractions who are offered interventions, $\forall Ak \in \left\{A1, \ldots, An_p\right\}$
Outputs	
$x_g^{(d)*}$	optimal fraction of budget in decision period d allocated to interventions in group $g \in \widetilde{\mathscr{G}}$ obtained by MF-RA, and $x^{(d)*} = \left\{x_1^{(d)*}, \ldots, x_{\lvert\widetilde{\mathscr{G}}\rvert}^{(d)*}\right\}$ is the set of the optimal fractions of budget, $\forall d \in \{1, \ldots, n_d\}$
$I_{s,g}^{(t)*}$	optimal fraction of population in health category $s \in \mathscr{S}_g$ and group $g \in \widetilde{\mathscr{G}}$ who are offered interventions in time period $t \in \mathscr{T}_d$, and $I^{(t)*} = \left\{I_{s,g}^{(t)*}\right\}_{s \in \mathscr{S}_g, g \in \widetilde{\mathscr{G}}}$, $\forall t \in \{1, \ldots, n_q\}$

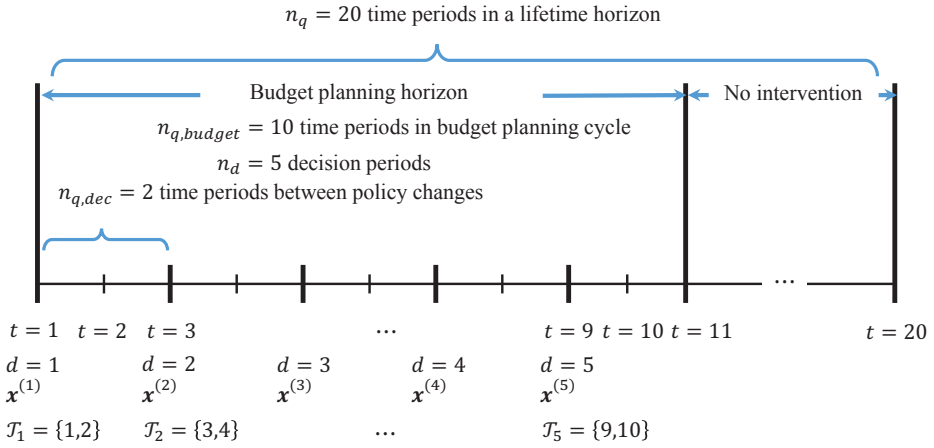


Figure 3.2: Illustration of time and decision periods and decision variables

horizon. Set the planning horizon to be $n_{q,budget} = n_d \cdot n_{q,dec}$ time periods. Let $\mathcal{T}_d$ denote the set of time period indices in decision period d, i.e., $\mathcal{T}_d = \{1 + n_{q,dec}(d-1), \ldots, n_{q,dec}d\}$.

It is assumed that there is no intervention after the pre-determined budget planning time periods $n_{q,budget}$, that is, set the budget and interventions to zero for the remaining time, i.e., $B_t = 0$, $I^s_{g,t} = 0$, $\forall s \in \mathcal{S}_g$, $g \in \widetilde{\mathcal{G}}$, for $t \in \{n_{q,budget} + 1, \ldots, n_q\}$.

An example to illustrate the notation is shown in Fig. 3.2. In this example, the lifetime horizon is twenty time periods and the number of decision periods is five, i.e., $n_q = 20$ and $n_d = 5$. Also, the policy is allowed to change every other time period, i.e., $n_{q,dec} = 2$. Thus, the budget planning horizon is $n_{q,budget} = 5 \cdot 2 = 10$ time periods and the sets of time period indices are $\mathcal{T}_1 = \{1,2\}$, $\mathcal{T}_2 = \{3,4\}$, …, and $\mathcal{T}_5 = \{9,10\}$. No intervention is implemented after 10 time periods, that is, $I^s_{g,t} = 0$, $\forall s \in \mathcal{S}_g$, $g \in \widetilde{\mathcal{G}}$, for $t \in \{11, \ldots, 20\}$.

In order to convert the budget constraint in (3.4) from time period t to decision period d, the decision variables $x^{(d)}_g \in [0,1]$ are introduced as the fraction of budget used for intervention in group $g \in \widetilde{\mathcal{G}}$ in decision period d and $\mathbf{x}^{(d)} = \left\{x^{(d)}_1, \ldots, x^{(d)}_{|\widetilde{\mathcal{G}}|}\right\}$ (see Fig. 3.2). The budget constraints are

therefore reformulated as

$$\sum_{s\in\mathscr{S}_g} C_g^s I_{g,t}^s N_t^s \quad \leq \quad x_g^{(d)} B_t, \forall g \in \widetilde{\mathscr{G}}, t \in \mathscr{T}_d, d \in \{1,\ldots,n_d\} \tag{3.8}$$

and

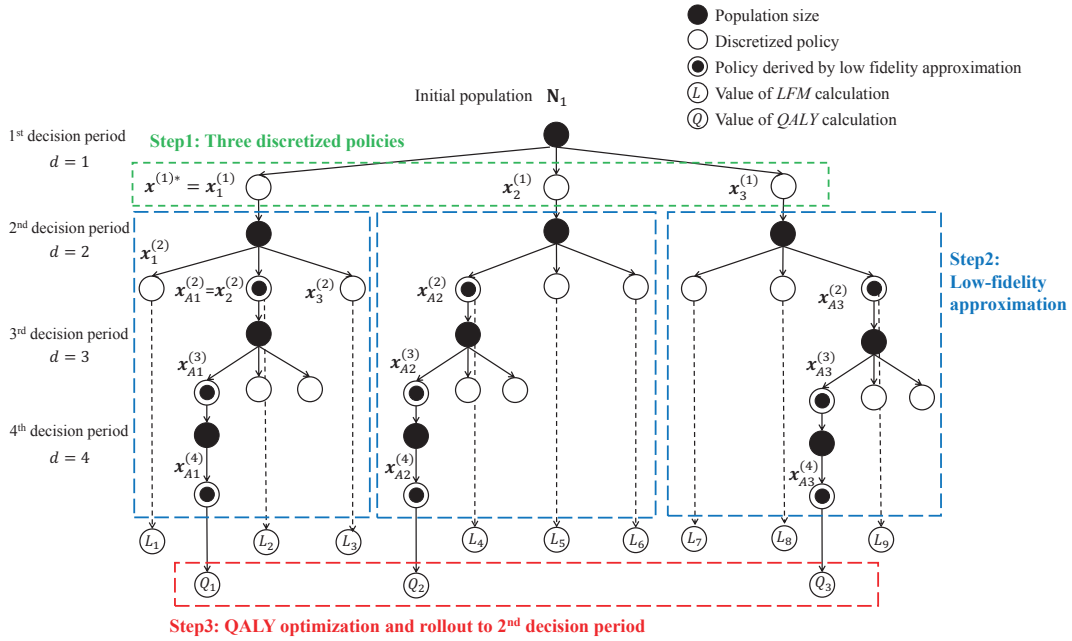
$$\sum_{g\in\widetilde{\mathscr{G}}} x_g^{(d)} \quad = \quad 1,\ \forall d \in \{1,\ldots,n_d\}, \tag{3.9}$$

which indicate that for each decision period d, the fractions of budget allocated to each intervention remain the same, but $I_{g,t}^s$, $t \in \mathscr{T}_d$, may change every time period based on the budget allocation rule of group g and population size.

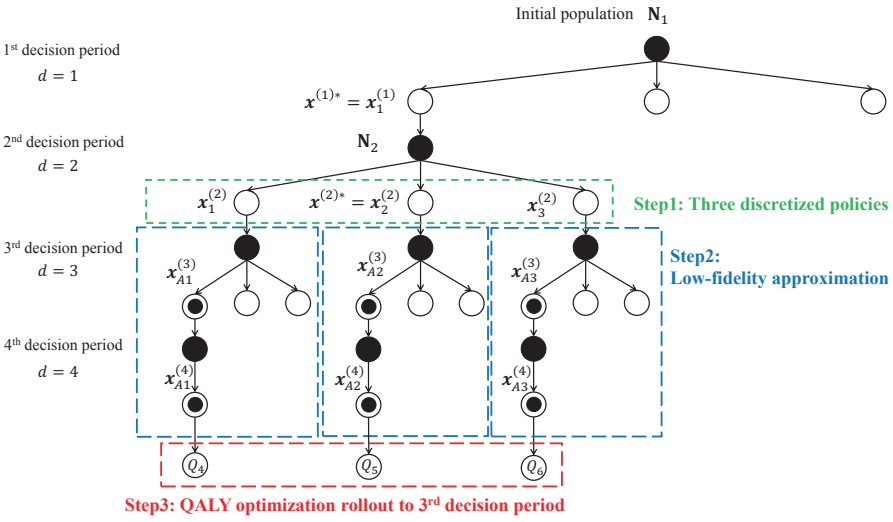
3.2.2 Detailed description of MF-RA

The idea of MF-RA is to embed the *LFM* into the detailed optimization model HQMP using a roll-out approach. In decision period d, MF-RA discretizes the policy space and for each discretized policy, advances to the next decision period. Given the current population, MF-RA calculates the best stationary policy using the simplified optimization model *LFM* and proceeds to the next decision period. A promising policy path is constructed by repeating these steps until the end of the planning horizon. Then for each discretized policy in decision period d, MF-RA calculates the total discounted QALYs for the entire promising policy path, selects the best discretized policy, and applies it to roll out to the next decision period.

An example of MF-RA is illustrated in Fig. 3.3 with four time periods and four decision periods ($t = d$). At $d = 1$, MF-RA first considers three policies, $\mathbf{x}_1^{(1)}$, $\mathbf{x}_2^{(1)}$, and $\mathbf{x}_3^{(1)}$ in Fig. 3.3(a). MF-RA proceeds to the next decision period to discretize three more policies, for each discretized intervention policy in decision period 1. *LFM* is then calculated as in (3.7), denoted $L_1, L_2, \ldots, L_9$ in Fig. 3.3(a). The largest *LFM* associated with each first decision policy is used to select the most promising decision in the second decision period, denoted $\mathbf{x}_{A1}^{(2)}$, $\mathbf{x}_{A1}^{(3)}$, and $\mathbf{x}_{A1}^{(4)}$ in Fig. 3.3(a). MF-RA continues in this manner and constructs three sequential intervention policies $\left(\mathbf{x}_1^{(1)}, \mathbf{x}_{A1}^{(2)}, \mathbf{x}_{A1}^{(3)}, \mathbf{x}_{A1}^{(4)}\right)$, $\left(\mathbf{x}_2^{(1)}, \mathbf{x}_{A2}^{(2)}, \mathbf{x}_{A2}^{(3)}, \mathbf{x}_{A2}^{(4)}\right)$, and $\left(\mathbf{x}_3^{(1)},\ \mathbf{x}_{A3}^{(2)}, \mathbf{x}_{A3}^{(3)}, \mathbf{x}_{A3}^{(4)}\right)$. The high-fidelity model computes QALY (using (3.2) and (3.3)) for each of these policies, yielding Q_1, Q_2, and Q_3 in Fig. 3.3(a).



(a) Illustration of MF-RA in the first decision period



(b) Illustration of MF-RA in the second decision period

Figure 3.3: Illustration of first and second decision periods of MF-RA

The largest QALY (e.g., Q_1 in Fig. 3.3(a)) is used to indicate the best budget allocation policy in the first decision period, denoted $\mathbf{x}^{(1)*}$. the associated intervention policy $\left\{I_{s,g}^{(t)*}\right\}_{s\in\mathscr{S}_{HS},g\in\widetilde{\mathscr{G}},t\in\mathscr{T}_1}$ is calculated by using $\mathbf{x}^{(1)*}$ and budget allocation rule within each group.

Fig. 3.3(b) illustrates MF-RA in the second decision period. The population size $\mathbf{N}_2$ at the second decision period is evolved using (3.2) and the budget allocation policy $\mathbf{x}^{(1)*}$. The algorithm continues, as shown in Fig. 3.3(b) to construct promising policies using the low-fidelity approximation and then evaluate QALYs at the most promising policies. In this example, Q_5 is the largest among Q_4, Q_5, and Q_6 in Fig. 3.3(b), indicating the best policy in the second decision period is $\mathbf{x}_2^{(2)}$, labeled $\mathbf{x}^{(2)*}$.

MF-RA finds the sequential budget allocation policies for all decision periods, $\mathbf{x}^{(1)*},\mathbf{x}^{(2)*},\ldots,\mathbf{x}^{(n_d)*}$, and the corresponding fractions of population to receive intervention $I^{(1)*},I^{(2)*},\ldots,I^{(n_{q,budget})*}$. A formal description follows of three steps to compute the best policy $\mathbf{x}^{(d)*}$ at each decision period d and associated intervention policy $I^{(t)*},\forall t\in\mathscr{T}_d$.

Step 1: Discretization of budget allocation and interventions

At the start of decision period d (and associated time period $t=1+(d-1)n_{q,dec}$), the population size $\mathbf{N}_t$ and the budget B_t are given. Step 1 construct as set of discretized budget allocation policies $\mathbf{x}^{(d)}=\left\{x_1^{(d)},\ldots,x_{|\widetilde{\mathscr{G}}|}^{(d)}\right\}$ and associated discretized intervention policies $I_{g,t}^{s}$ for $s\in\mathscr{S}_g$, $g\in\widetilde{\mathscr{G}}$ and $t\in\mathscr{T}_d$ that satisfy constraints (3.8) and (3.9).

First, discretize $x_g^{(d)}$ uniformly over the unit interval, so that $x_g^{(d)}\in\left\{0,\frac{1}{n_{dis}},\ldots,\frac{n_{dis}-1}{n_{dis}},1\right\}$, yielding $n_{dis}+1$ values for each $g\in\widetilde{\mathscr{G}}$. There are

$$n_p=C_{|\widetilde{\mathscr{G}}|-1}^{|\widetilde{\mathscr{G}}|+n_{dis}-1}=\frac{(|\widetilde{\mathscr{G}}|+n_{dis}-1)!}{(|\widetilde{\mathscr{G}}|-1)!|\widetilde{\mathscr{G}}|!} \tag{3.10}$$

different combinations of budget allocation policies that satisfy (3.9). To denote these discrete budget allocation policies, let $\mathbf{x}_k^{(d)}=\left\{x_{1,k}^{(d)},\ldots,x_{|\widetilde{\mathscr{G}}|,k}^{(d)}\right\}$ be the set of kth budget allocation policy, $k\in\{1,\ldots,n_p\}$ for $g\in\widetilde{\mathscr{G}}$ at decision period d. Since the budget allocation policy changes every decision period, keep the same discretized fraction of budget $\mathbf{x}_k^{(d)}$ for every $n_{q,dec}$ time periods, i.e., $t\in\mathscr{T}_d$.

Given the budget allocation policies $\mathbf{x}_k^{(d)}$ for time period t and all $g\in\widetilde{\mathscr{G}}$, obtain the associated

discretized interventions that satisfy the budget constraint (3.8), using the evolved population size $\mathbf{N}_t$. Let $I^t_{s,g,k}$ be the fraction of population in health category $s \in \mathscr{S}_g$ and $g \in \widetilde{\mathscr{G}}$ that receive intervention within budget constraint. This construction of $I^t_{s,g,k}$ takes a budget allocation rule in each group into account, such as, treat the sickest patient first. An example of how to obtain $I^t_{s,g,k}$ from $\mathbf{x}^{(d)}_k$ is given in Section 3.4.1. Once $I^t_{s,g,k}$ is obtained, the population size is evolved according to (3.2) to the next time period $t+1$, i.e., $\mathbf{N}_{1+dn_{q,dec}} = \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_d} \mathbf{P}\left(I^{(t)}_k\right)$, and thus is repeated for $t \in \mathscr{T}_d$ until the next decision period.

Step 2: Low-fidelity approximation

After deriving the discretized policy $\mathbf{x}^{(d)}_k, \forall k = \{1,\ldots,n_p\}$ in decision period d (time period $1+(d-1)n_{q,dec}$) and associated evolved population size in decision period $d+1$ (time period $1+dn_{q,dec}$), MF-RA then iteratively uses LFM to develop promising sequential policies (from decision period $d+1$ to decision period n_d) following each candidate policy $\mathbf{x}^{(d)}_k$, $\forall k = \{1,\ldots,n_p\}$. Similar to the discretization process in Step 1, fractions of population receiving intervention $I^{(t)}_k$, $\forall t \in \mathscr{T}_{d+1}$, $\forall k \in \{1,\ldots,n_p\}$ are identified by mapping from discretized budget allocation policies $\mathbf{x}^{(d+1)}_k$, $k \in \{1,\ldots,n_p\}$. Then, identify the promising policy $I^{(1+dn_{q,dec})}_{Ak}$ (that is, identify $\mathbf{x}^{(d+1)}_{Ak}$ following $\mathbf{x}^{(d)}_k$) by maximizing LFM,

$$I^{(1+dn_{q,dec})}_{Ak} = \underset{I^{(1+dn_{q,dec})}_k, k\in\{1,\ldots,n_p\}}{\arg\max} LFM\left(\mathbf{N}_{1+dn_{q,dec}}, I^{(1+dn_{q,dec})}_k\right). \tag{3.11}$$

MF-RA updates population size by $\mathbf{N}_{1+(d+1)n_{q,dec}} = \mathbf{N}_{1+dn_{q,dec}} \prod_{t\in\mathscr{T}_{d+1}} \mathbf{P}\left(I^{(t)}_{Ak}\right)$, and then proceeds to the next decision period $d+2$. By iteratively maximizing LFM from decision period $d+1$ to n_d, the promising sequential intervention policies are developed from decision period d to n_d, i.e., $\mathbf{x}^{(d)}_k$, $\mathbf{x}^{(d+1)}_{Ak},\ldots,\mathbf{x}^{(n_d)}_{Ak}$. Therefore, for all discretized policies $\mathbf{x}^{(d)}_k, \forall k \in \{1,\ldots,n_p\}$, compute n_p promising sequential interventions from the current time period to the last time period in $\mathscr{T}_{n_d}$, i.e., $I^{(t)}_k, \forall t \in \mathscr{T}_d$ and $I^{(t)}_{Ak}, \forall t \in \mathscr{T}_{d+1} \cup \cdots \cup \mathscr{T}_{n_d}$, for all $k \in \{1,\ldots,n_p\}$.

Step 3: QALY optimization and rollout

In Step 3, MF-RA selects the best policy $\mathbf{x}^{(d)*}$ in decision period d with maximum overall discounted QALYs among all n_p promising sequential interventions derived in Step 2, and then proceeds to the next decision period $d+1$. After calculating overall discounted QALYs using HQMP for all n_p sequential interventions, MF-RA then selects the best sequential intervention with the maximum QALYs and extracts the best fraction of budget policy during decision period d as the best budget allocation policy, $\mathbf{x}^{(d)*}$. By implementing associated fractions of population $I^{(t)*}, \forall t \in \mathscr{T}_d$ in decision period d, MF-RA advances to the next decision period $d+1$ and the population size is updated by (3.2).

3.3 Rollout Algorithm with Branch-and-Bound (RA-BnB)

Healthcare decision models in population disease management problems are commonly formulated using Markov decision processes (MDPs), which are challenging to solve due to large state-space and action- and time-dependent transition probabilities [136]. In the previous research [80], Multi-Fidelity Rollout Algorithm (MF-RA) is proposed to derive sequential policy for screening and treatment with budget constraints; while MF-RA is capable of efficiently identifying good sequential policies (population computation is quadratic in the number of decision periods), the optimality of the derived policy is not guaranteed. This paper designs an algorithm that provides an optimality gap on intermediate solutions and achieves an optimality guarantee on the final healthcare intervention solution when the gap is zero. Accordingly, decision makers are able to examine the trade-off between policy quality and necessitated computational effort. The approach is illustrated by a numerical example of screening and treatment policy implementation for chronic hepatitis C virus (HCV) infection over a budget planning period.

The proposed Rollout Algorithm with Branch-and-Bound (RA-BnB) is set in a branch-and-bound framework and incorporates the idea of approximate dynamic programing in MF-RA. MF-RA constructs promising policy paths by exploiting a low fidelity model and incorporating it into the Markov disease progression model. In RA-BnB, the non-promising policy paths may be pruned in early iterations by finding a good incumbent solution quickly (via the idea in MF-RA). That is, if those paths that are known to be sub-optimal are eliminated, then the resulting promising paths will

consist of sequential policies that are likely to lead to the optimal solution. Also, an optimality gap can be easily retrieved with a good incumbent solution and an easily calculated upper bound.

The success of implementing branch-and-bound to solve the QALYs maximization problem depends on efficiently determining a good feasible solution and providing upper bounds in the search space. The following subsections introduce how to efficiently compute upper bounds and an incumbent, and then present the details of the RA-BnB algorithm.

3.3.1 Upper bound computation

At time period t, given previous allocation policies before time period t, $\{x_g^{(t')}\}_{g\in\widetilde{\mathscr{G}},t'\in\{1,\ldots,t-1\}}$, and an initial population size $\mathbf{N}_1$, an upper bound on the objective function in QALYs maximization problem can be derived by solving the following upper bound optimization problem,

Upper bound optimization problem

$$\max_{\left\{x_g^{(t'')}\right\}_{g\in\widetilde{\mathscr{G}},t''\in\{t,\ldots,n_T\}}} \quad QALY\left(\mathbf{N}_1,\left\{x_g^{(t')}\right\}_{g\in\widetilde{\mathscr{G}},t'\in\{1,\ldots,t-1\}},\left\{x_g^{(t'')}\right\}_{g\in\widetilde{\mathscr{G}},t''\in\{t,\ldots,n_T\}}\right) \tag{3.12}$$

$$\text{s.t.} \quad 0\le x_g^{(t'')}\le 1, \forall g\in\widetilde{\mathscr{G}},\ \forall t''\in\{t,\ldots,n_T\} \tag{3.13}$$
$$\text{and } (3.2).$$

In the upper bound optimization problem, Constraint (3.4) in the QALYs maximization problem is eliminated, thereby allocating the same amount of budget B_t to every group $g\in\widetilde{\mathscr{G}}$. The complexity of evaluating different budget allocation policies among different groups is therefore eliminated. In addition, the *QALY* function (3.3) is monotonically non-decreasing in $x_g^{(t)}$ due to the model structure, where interventions have a non-negative impact on QALYs. Thus, the upper bound optimization problem is easily solved by setting $x_g^{(t'')}$ to its maximum value of one, i.e., the optimal budget allocation in the upper bound optimization is $x_g^{(t'')ub}=1$, $\forall g\in\widetilde{\mathscr{G}}$ and $\forall t''\in\{t,\ldots,n_T\}$. Therefore, an upper bound of overall QALYs can be obtained by running only one lifelong-time population disease model.

Theorem 1. *Given the same initial population size* $\mathbf{N}_1$ *and determined allocation policies from time*

period 1 *to* $t-1$, $\left\{x_g^{(t')}\right\}_{g\in\widetilde{\mathscr{G}},t'\in\{1,\ldots,t-1\}}$, *the overall QALYs derived in the upper bound optimization problem is no less than its counterpart derived in QALY optimization problem. That is,*

$$\begin{aligned} &QALY\left(\mathbf{N}_1,\left\{x_g^{(t')}\right\}_{g\in\widetilde{\mathscr{G}},t'\in\{1,\ldots,t-1\}},\left\{x_g^{(t'')ub}\right\}_{g\in\widetilde{\mathscr{G}},t''\in\{t,\ldots,n_T\}}\right) \\ &\geq\ QALY\left(\mathbf{N}_1,\left\{x_g^{(t')}\right\}_{g\in\widetilde{\mathscr{G}},t'\in\{1,\ldots,t-1\}},\left\{x_g^{(t'')*}\right\}_{g\in\widetilde{\mathscr{G}},t''\in\{t,\ldots,n_T\}}\right) \end{aligned} \tag{3.14}$$

where $x_g^{(t'')ub}$ *and* $x_g^{(t'')*}$ *are the optimal fractions of budget for group* $g\in\widetilde{\mathscr{G}}$ *in time period* t'' *derived in the upper bound optimization problem and QALYs maximization problem, respectively.*

Proof. Since the QALYs maximization problem is a more restricted version of the upper bound optimization problem, its optimal value is less than or equal to the optimal value of the upper bound problem. Thus, (3.14) always holds. □

3.3.2 Low fidelity approximation

The Low Fidelity Approximation algorithm is similar to the rollout algorithm MF-RA in [80]. It constructs a promising policy path by iteratively applying low fidelity models until the end of the planning horizon. The low fidelity model is based on a common feature in disease progression models; that death is an absorbing state. Treating each individual dynamics as an absorbing Markov chain under a stationary intervention policy, the matrix of expected number of time periods spent in each state until absorption, beginning at time period t, can be quickly calculated.

$$\mathbf{E}_t\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathscr{G}}}\right)=\left(\mathbf{I}-\mathbf{Q}_t\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathscr{G}}}\right)\right)^{-1} \tag{3.15}$$

where $\mathbf{I}$ is the appropriately sized identity matrix and $\mathbf{Q}_t$ is the submatrix of the transition probability matrix $\mathbf{P}_t$ obtained by eliminating the row and column associated with the mortality state.

The approximated QALYs from time t with starting population $\mathbf{N}_t$ to the time that the entire population has reached the mortality state under stationary policy $\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathscr{G}}}$ is expressed as the

Low Fidelity Model (LFM),

$$\begin{aligned} LFM\left(\mathbf{N}_t, \left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathscr{G}}}\right) &= \mathbf{N}_t\mathbf{E}_t\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathscr{G}}}\right)\mathbf{u}' \\ &= \sum_{s\in\mathscr{S}_{HS}}\sum_{v\in\mathscr{S}_{HS}} N_t^s u^v E_t^{s,v}\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathscr{G}}}\right) \end{aligned} \tag{3.16}$$

where $E_t^{s,v}$ is the s,vth element of the matrix $\mathbf{E}_t$, representing the expected number of time periods spent in state v, starting in state s until absorption. For details of the Low Fidelity Approximation algorithm which iteratively uses LFM to develop promising sequential policies, the reader is referred to the work in [80].

The Low Fidelity Approximation provides an incumbent solution which is easily computed. The incumbent solution and the QALYs upper bound calculated in Section 3.3.1 are incorporated into a branch-and-bound framework which is detailed in the following section.

3.3.3 RA-BnB algorithm

A decision period index d is introduced to distinguish the policy-changing periods from time period t used for disease progression. Suppose each decision period d contains n_B time periods and let n_D denote the number of decision periods. The number of time periods is $n_T = n_D n_B$. Decision variables $x_g^{(d)} \in [0,1]$ denote the fraction of budget used for intervention in group g during decision period d, $d \in \{1,2,\ldots,n_D\}$. The reformulation indicates that during decision period d, the percentages of budget allocated to each intervention policy, $x_g^{(t)}$, $\forall t \in \mathscr{T}_d$ are the same, where $\mathscr{T}_d = \{n_B(d-1)+1,\ldots,n_B d\}$ is the set of time period indices in decision period d. That is, the same percentages of budget for each intervention policy are kept in every n_B time periods.

Also, discretize the continuous solution space into finite policies by discretizing the budget into m_B pieces, $x_g^{(d)}$ is therefore uniformly discretized and has m_B+1 possible values, i.e., $x_g^{(d)} \in \{0, \frac{1}{m_B}, \ldots, \frac{m_B-1}{m_B}, 1\}$. Therefore, there are $m_P = C(|\widetilde{\mathscr{G}}|+m_B-1, |\widetilde{\mathscr{G}}|-1)$ different combinations of budget allocation policies since the sum of the fractions of budget for each group equals one in each time period. Let $x_{g,k_1,\ldots,k_d}^{(d)}$ denote the k_dth, $k_d \in \{1,\ldots,m_P\}$, budget allocation policy in group g given prior policies $x_{g,k_1}^{(1)}$, $\cdots$, $x_{g,k_1,\ldots,k_{d-1}}^{(d-1)}$. For the sake of clarity, let $\mathbf{x}_{k_1,\ldots,k_d}^{(d)} = \{x_{1,k_1,\ldots,k_d}^{(d)},\ldots,x_{|\widetilde{\mathscr{G}}|,k_1,\ldots,k_d}^{(d)}\}$ denote the set of the k_dth combination of discretized budget in decision period d for all groups in $\widetilde{\mathscr{G}}$.

To illustrate, the example in Figure 3.4(a) has four decision periods, $n_D = 4$, and each contains one time period, $n_B = 1$, so there are four total time periods, $n_T = 4$. The budget in the first decision period is discretized into two pieces, $m_B = 2$, and therefore there are three possible values of $x^{(1)}_{g,k_1}$ for each group $g \in \{A,B\}$, i.e., $x^{(1)}_{A,k_1} \in \{0, 0.5, 1\}$, and $x^{(1)}_{B,k_1} \in \{0, 0.5, 1\}$. Since the total budget allocation must equal one, there are three candidate budget allocation policies, i.e., $m_P = C(2+2-1, 2-1) = 3$ with $\mathbf{x}^{(1)}_1 = \{0,1\}$, $\mathbf{x}^{(1)}_2 = \{0.5, 0.5\}$, and $\mathbf{x}^{(1)}_3 = \{1,0\}$.

RA-BnB uses a breadth first search strategy for the rollout approximation which explores all discretized policies in the dth decision period before proceeding to the discretized policies in the $(d+1)$th decision period. In the decision period d, m_p combinations of budget allocation policies following policies in the $(d-1)$th decision period are generated. For each discretized policy, the Low Fidelity Approximation attempts to determine good feasible solutions, and the upper bound of overall QALYs over the lifetime horizon is also calculated. For convenience, let $Q^U(\mathbf{N}_1, \{\mathbf{x}^{(1)}_{k_1}, \ldots, \mathbf{x}^{(d)}_{k_1,\ldots,k_d}\})$ denote the function for solving the upper bound optimization problem in Section 3.3.1, and $Q^U_{k_1 \cdots k_d}$ be the corresponding returned value, i.e., the solution of upper bound optimization problem; let $Q^L(\mathbf{N}_1, \{\mathbf{x}^{(1)}_{k_1}, \ldots, \mathbf{x}^{(d)}_{k_1,\ldots,k_d}\})$ denote the function for running Low Fidelity Approximation in Section 3.3.2, and $Q^L_{k_1 \cdots k_d}$ the corresponding returned overall QALYs. Note that both returned overall QALYs $Q^L_{k_1 \cdots k_d}$ and $Q^U_{k_1 \cdots k_d}$ are calculated by the *QALY* function (3.3).

RA-BnB

Given: the initial population size $\mathbf{N}_1$, transition probability matrix $\mathbf{P}_t$, budget B_t, health utility multipliers $\mathbf{u}$, number of pieces in discretized budget m_B, number of time periods between policy changing n_B, number of decision periods n_D, number of time periods in lifetime horizon $n_T = n_D n_B$, and $d = 1$.

Step 1: **Branching** If $d = 1$, $\mathbb{X}^{(1)} = \{\mathbf{x}^{(1)}_{k_1}, \forall k_1 \in \{1, \ldots, m_P\}\}$. If $d \geq 2$, discretize each policy $\mathbf{x}^{(d-1)}_{k_1,\ldots,k_{d-1}} \in \mathbb{X}^{(d-1)}$ into m_P different combinations and store them in $\mathbb{X}^{(d)}$, i.e.,

$$\mathbb{X}^{(d)} = \left\{ \mathbf{x}^{(d)}_{k_1,\ldots,k_d} : \mathbf{x}^{(d-1)}_{k_1,\ldots,k_{d-1}} \in \mathbb{X}^{(d-1)} \text{ and } k_d \in \{1, \ldots, m_P\} \right\}.$$

Step 2: **Upper Bound Update** Evaluate $Q^U\left(\mathbf{N}_1, \left\{\mathbf{x}^{(1)}_{k_1}, \ldots, \mathbf{x}^{(d)}_{k_1,\ldots,k_d}\right\}\right)$ and compute upper bounds of

overall QALYs, $Q^U_{k_1\cdots k_d}$ for all the newly generated policies. Set $Q^{(d)}_U$ to the maximum value among all $Q^U_{k_1\cdots k_d}$.

Step 3: **Elite Set** If $d=1$, $\mathbb{E}^{(d)}=\mathbb{X}^{(d)}$. If $d\geq 2$, sort all $Q^U_{k_1\cdots k_d}$ and select the top $(d-1)m_P$ policies and store them in the elite set $\mathbb{E}^{(d)}$. The number of policies in the elite set increases with d and is proportional to combinations of budget allocation policies m_p.

Step 4: **Low Fidelity Approximation** Compute $Q^L_{k_1\cdots k_d}$ by evaluating $Q^L\left(\mathbf{N}_1,\left\{\mathbf{x}^{(1)}_{k_1},\ldots,\mathbf{x}^{(d)}_{k_1,\ldots,k_d}\right\}\right)$, $\forall \mathbf{x}^{(d)}_{k_1,\ldots,k_d}\in\mathbb{E}^{(d)}$ for each $\mathbf{x}^{(d)}_{k_1,\ldots,k_d}$ as follows:

Step 4-1: Set $\widetilde{d}\leftarrow d+1$.

Step 4-2: **Policies Discretization** Discretize budget and obtain m_P budget allocation policies in decision period $\widetilde{d}$, $\mathbf{x}^{(\widetilde{d})}_{k_1,\ldots,k_{\widetilde{d}}}$, $\forall k_{\widetilde{d}}\in\{1,\ldots,m_P\}$.

Step 4-3: **Low Fidelity Model** Identify the promising policy $\mathbf{x}^{(\widetilde{d})}_{Ak_d}$ in decision period $\widetilde{d}$ by selecting the maximum approximated QALYs among m_P LFM functions calculated by Equation (3.16). Ak_d is the index of the derived budget allocation policy for k_dth policy, i.e., $\mathbf{x}^{(d)}_{k_1,\ldots,k_d}$, evaluation.

Step 4-4: **Population Update** Update population size $\mathbf{N}_{\widetilde{d}}$ by Equation (3.2). If $\widetilde{d}<n_D$, proceed to the next decision period $\widetilde{d}\leftarrow\widetilde{d}+1$ and go to Step 4-2. If $\widetilde{d}=n_D$, continue to Step 4-5.

Step 4-5: **Promising Policy** Identify complete sequential promising policy $\mathbf{x}^{(d+1)}_{Ak_d}$, $\mathbf{x}^{(d+2)}_{Ak_d}$, $\cdots$, $\mathbf{x}^{(n_D)}_{Ak_d}$ and corresponding QALYs $Q^L_{k_1\cdots k_d}$.

Step 5: **Incumbent Update** Set $Q^{(d)}_L$ to the maximum value among all $Q^L_{k_1\cdots k_d}$ and label the corresponding policy as $\mathbf{X}^{(d)}_L$. If $Q^{(d)}_L>Q^{(d-1)}_L$, set $Q^*=Q^{(d)}_L$ and $\mathbf{X}^*=\mathbf{X}^{(d)}_L$. Calculate optimality gap,

$$Q^{(d)}_U-Q^*. \tag{3.17}$$

Step 6: **Fathoming** Update $\mathbb{X}^{(d)}$ by discarding all policies whose upper bounds are less than Q^*

$$\mathbb{X}^{(d)}=\mathbb{X}^{(d)}\setminus\left\{\mathbf{x}^{(d)}_{k_1,\ldots,k_d}:Q^U\left(\mathbf{N}_1,\left\{\mathbf{x}^{(1)}_{k_1},\ldots,\mathbf{x}^{(d)}_{k_1,\ldots,k_d}\right\}\right)<Q^*\right\}.$$

Step 7: **Stopping rule** If

$$\frac{Q_U^{(d)} - Q^*}{Q^* - Q_{no\,budget}} \leq \delta,$$

where δ is stopping threshold for a scaled performance measurement, then terminate the algorithm and output the complete sequential policy $\mathbf{X}^*$; otherwise, update population size by Equation (3.2), and proceed to Step 1 for the next decision period, $d = d + 1$.

Theorem 2. *RA-BnB guarantees optimality (w.r.t. the level of discretization) when the gap is zero.*

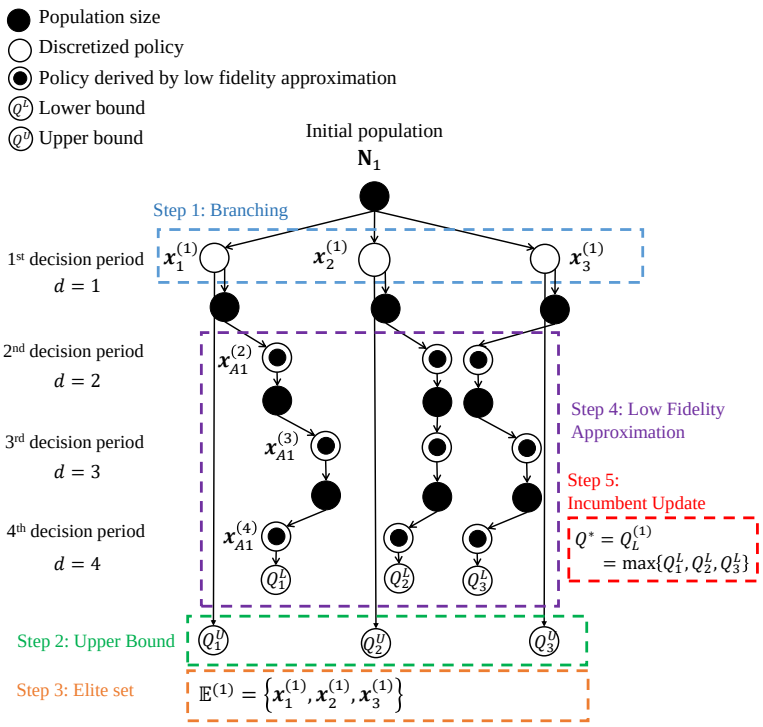
Proof. For each generated policy in the dth decision period, $Q^U_{k_1 \cdots k_d}$ is determined by evaluating (3.3) with the given policy decisions $k_1, \cdots, k_d$ and setting future policy decisions equal to one, for $k_{d+1}, \cdots, k_{n_D}$. Since the $d+1$ decision period has a policy less than or equal to one, $Q^U_{k_1 \cdots k_d k_{d+1}} \leq Q^U_{k_1 \cdots k_d}$, and thus $Q_U^{(d)}$ is monotonically non-increasing over iterations. In addition, Q^* is monotonically non-decreasing since the incumbent solution improves over iterations. Using the property of the bounds,

$$Q^* \leq Q_{opt} \leq Q_U^{(d)},$$

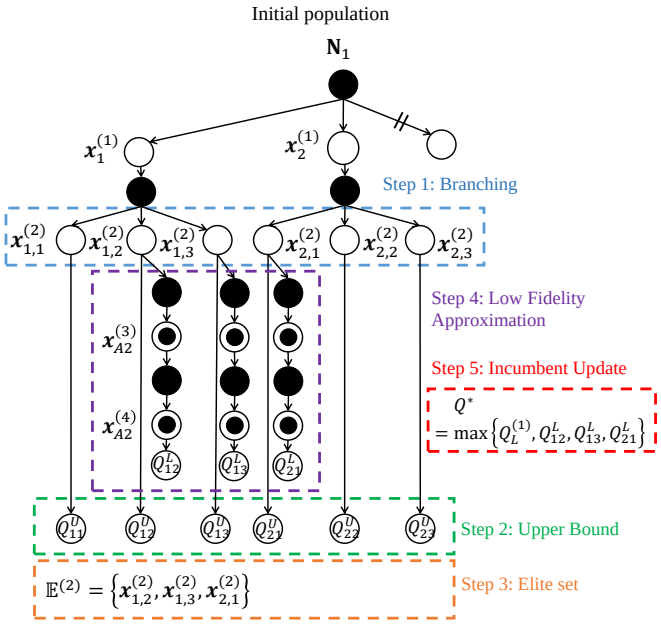
when $Q^* = Q_U^{(d)}$, the derived incumbent solution is therefore an optimal solution, $Q^* = Q_{opt}$. □

An example problem

To illustrate the working of the RA-BnB, refer again to the example in Figure 3.4. In the first decision period (Figure 3.4(a)), the set of discretized policies in Step 1 is $\mathbb{X}^{(1)} = \{\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}, \mathbf{x}_3^{(1)}\}$. In Step 2, RA-BnB computes the upper bound for these discretized policies, i.e., Q_1^U, Q_2^U, Q_3^U, and selects the largest as the upper bound, i.e., $Q_U^{(1)} = \max\{Q_1^U, Q_2^U, Q_3^U\}$. In this example, Q_2^U is the largest. In Step 3, the elite set of potential policies is selected, set $\mathbb{E}^{(1)} = \mathbb{X}^{(1)} = \{\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}, \mathbf{x}_3^{(1)}\}$. In Step 4, Low Fidelity Approximation is implemented for the policies in $\mathbb{E}^{(1)}$, obtaining Q_1^L, Q_2^L, and Q_3^L. In Step 5, the maximum value among the results is $Q_L^{(1)} = \max\{Q_1^L, Q_2^L, Q_3^L\}$. In this example, Q_1^L is the maximum, and the corresponding policy is $\mathbf{X}_L^{(1)} = \left[\mathbf{x}_1^{(1)}, \mathbf{x}_{A1}^{(2)}, \mathbf{x}_{A1}^{(3)}, \mathbf{x}_{A1}^{(4)}\right]$. Then, update the incumbent solution, i.e., $Q^* = Q_L^{(1)}$. In the example in Step 6, $Q_1^U > Q^*$, $Q_2^U > Q^*$, and $Q_3^U < Q^*$, so the right branch is pruned and $\mathbb{X}^{(1)} = \mathbb{X}^{(1)} \setminus \{\mathbf{x}_3^{(1)}\} = \{\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}\}$. In Step 7, if $\frac{Q_U^{(1)} - Q^*}{Q^* - Q_{no\,budget}} > \delta$, go to Step 1 and proceed to the second decision period, $d = 2$ (Figure 3.4(b)).



(a) Illustration of RA-BnB in the first decision period



(b) Illustration of RA-BnB in the second decision period

Figure 3.4: Illustration of first and second decision periods of RA-BnB

Figure 3.5: Illustration example of Low Fidelity Approximation in the first decision period in RA-BnB

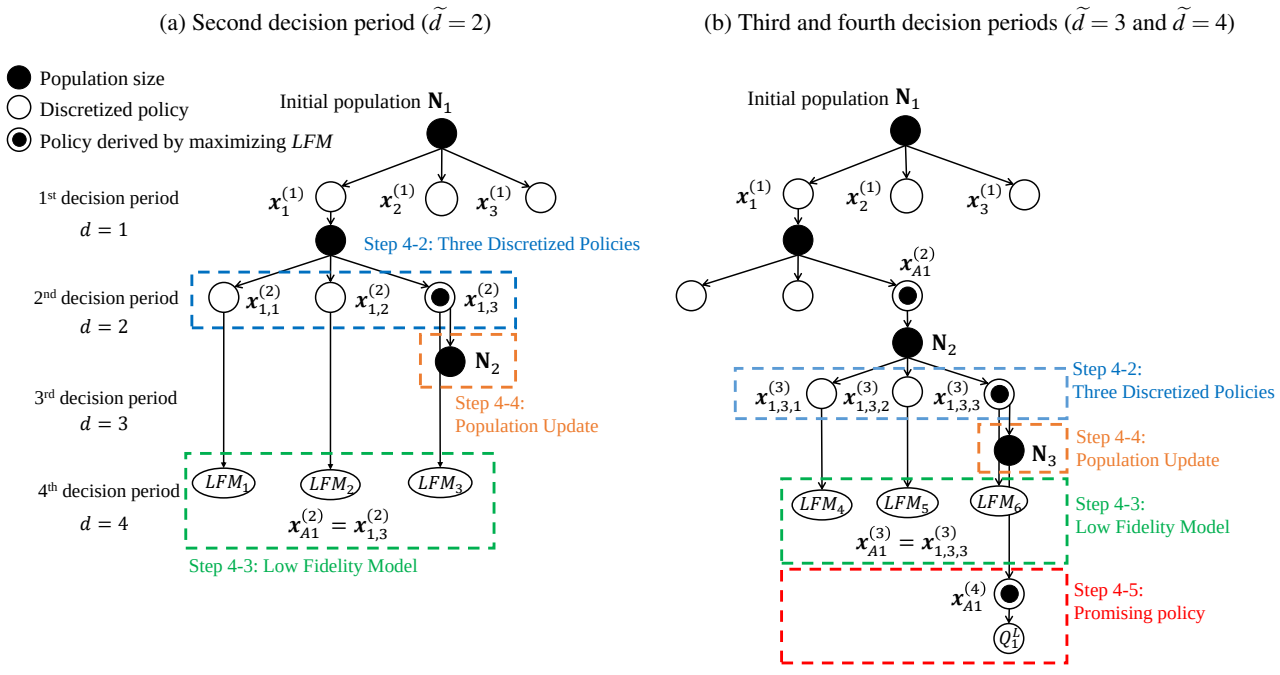


Figure 3.5 presents an illustration of the Low Fidelity Approximation in Steps 4-1 through 4-5, using $\mathbf{x}_1^{(1)}$ as an example. The goal is to construct the incumbent solution $\mathbf{x}_1^{(1)}$, $\mathbf{x}_{A1}^{(2)}$, $\mathbf{x}_{A1}^{(3)}$, $\mathbf{x}_{A1}^{(4)}$, and value Q_1^L from promising policies during decision periods 2 to 4 given $\mathbf{x}_1^{(1)}$. As shown in Figure 3.5(a), set $\widetilde{d} = 1+1 = 2$ in Step 4-1 and Low Fidelity Approximation enumerates 3 policies, i.e., $\mathbf{x}_{1,1}^{(2)}$, $\mathbf{x}_{1,2}^{(2)}$, and $\mathbf{x}_{1,3}^{(2)}$, as in Step 4-2. Step 4-3 calculates the corresponding approximate QALYs of low fidelity models, i.e., LFM_1, LFM_2, and LFM_3. In this example, LFM_3 has the maximum value among LFM_1, LFM_2, and LFM_3; $\mathbf{x}_{1,3}^{(2)}$ is selected as the allocation policy at the second decision period, i.e., $\mathbf{x}_{A1}^{(2)} = \mathbf{x}_{1,3}^{(2)}$. In Step 4-4, the population size is evolved to $\mathbf{N}_2$ using selected policy $\mathbf{x}_{1,3}^{(2)}$. As shown in Figure 3.5(b), Low Fidelity Approximation then proceeds to the 3rd decision period. In this example, the most promising policy in Step 4-3 for the 3rd decision period is $\mathbf{x}_{1,3,3}^{(3)}$, so $\mathbf{x}_{A1}^{(3)} = \mathbf{x}_{1,3,3}^{(3)}$. Repeating the same Steps 4-2 through 4-4 derives the policy $\mathbf{x}_{A1}^{(4)}$ in the 4th decision period and Step 4-5 identifies the promising policy $\mathbf{x}_{A1}^{(2)}$, $\mathbf{x}_{A1}^{(3)}$, $\mathbf{x}_{A1}^{(4)}$.

In the second decision period (back to Figure 3.4(b)), RA-BnB discretizes the policies which follows the policies in $\mathbb{X}^{(1)}$, i.e., $\mathbf{x}_{1,1}^{(2)}$, $\mathbf{x}_{1,2}^{(2)}$ and $\mathbf{x}_{1,3}^{(2)}$ following $\mathbf{x}_1^{(1)}$, and $\mathbf{x}_{2,1}^{(2)}$, $\mathbf{x}_{2,2}^{(2)}$ and $\mathbf{x}_{2,3}^{(2)}$ following $\mathbf{x}_2^{(1)}$. The set of policies in the second decision period is $\mathbb{X}^{(2)} = \{\mathbf{x}_{1,1}^{(2)}, \mathbf{x}_{1,2}^{(2)}, \mathbf{x}_{1,3}^{(2)}, \mathbf{x}_{2,1}^{(2)}, \mathbf{x}_{2,2}^{(2)}, \mathbf{x}_{2,3}^{(2)}\}$. In Step 2, upper bounds are computed for all the newly generated policies in $\mathbb{X}^{(2)}$, i.e., Q_{11}^U, Q_{12}^U, Q_{13}^U, Q_{21}^U, Q_{22}^U, and Q_{23}^U. Suppose Q_{13}^U is selected as the largest upper bound, i.e., $Q_U^{(2)} = \max\{Q_{11}^U, \ldots, Q_{23}^U\} = Q_{13}^U$. In Step 3, Q_{11}^U,...,Q_{23}^U are sorted and the top 3 policies are put into the elite set $\mathbb{E}^{(2)}$. Say $Q_{13}^U > Q_{12}^U > Q_{21}^U > Q_{11}^U > Q_{22}^U > Q_{23}^U$, then $\mathbb{E}^{(2)} = \{\mathbf{x}_{1,2}^{(2)}, \mathbf{x}_{1,3}^{(2)}, \mathbf{x}_{2,1}^{(2)}\}$. In Step 4, the Low Fidelity Approximation is implemented, yielding Q_{12}^L, Q_{13}^L, and Q_{21}^L. In Step 5, the policy with maximum overall QALYs is $Q_L^{(2)} = \max\{Q_{12}^L, Q_{13}^L, Q_{21}^L\} = Q_{12}^L$, thus $\mathbf{X}_L^{(2)} = \left[\mathbf{x}_1^{(1)}, \mathbf{x}_{1,2}^{(2)}, \mathbf{x}_{A2}^{(3)}, \mathbf{x}_{A2}^{(4)}\right]$. If $Q_L^{(2)} > Q_L^{(1)}$, $Q^* = Q_L^{(2)} = Q_{12}^L$ and $\mathbf{X}^* = \mathbf{X}_L^{(2)}$. In Step 6, if $Q_{22}^U < Q^*$ and $Q_{23}^U < Q^*$, two branches $\mathbf{x}_{2,2}^{(2)}$ and $\mathbf{x}_{2,3}^{(2)}$ are pruned. Update $\mathbb{X}^{(2)}$ by discarding all policies whose upper bounds are less than Q^*, i.e., $\mathbb{X}^{(2)} = \mathbb{X}^{(2)} \setminus \{\mathbf{x}_{2,2}^{(2)}, \mathbf{x}_{2,3}^{(2)}\}$ In Step 7, suppose $\frac{Q_U^{(2)} - Q^*}{Q^* - Q_{no\,budget}} < \delta$, RA-BnB is thus terminated. The output sequential policy is $\mathbf{X}^* = \left[\mathbf{x}_1^{(1)}, \mathbf{x}_{1,2}^{(2)}, \mathbf{x}_{A2}^{(3)}, \mathbf{x}_{A2}^{(4)}\right]$ with corresponding QALYs Q^* and the optimality gap is $Q_U^{(d)} - Q^*$.

3.4 Numerical Results for Hepatitis C Screening and Treatment Policies Implementation

This section applies MF-RA and RA-BnB to obtain screening and treatment policy decisions under budget constraints for chronic hepatitis C birth-cohorts. Section 3.4.1 extends the HCV compart-

mental simulation model in prior works ([96, 98]) by grouping health categories into five major groups including the degree of liver damage, and describes the HCV high-fidelity optimization problem to maximize total QALYs with a budget constraint and the HCV low-fidelity approximation. Sections 3.4.2-3.4.6 conduct numerical experiments to optimize screening and treatment policies using MF-RA.

3.4.1 HCV Problem formulation

Fig. 3.6 illustrates a hepatitis C progression model representing American birth-cohorts going through possible screening and treatment interventions. Fibrosis staging is measured by the Metavir score (F0-F4), with possible transitions occurring every quarter, thus the HCV progression time period t is set to correspond to three months (a quarter). The population in the HCV healthcare management system is divided into five groups. Individuals with different health status exist in different groups. The subscripts A, B, C, D, and M are used to differentiate the different groups, and combine with health categories to obtain 40 group-based health categories, denoted in sets as follows: target screening $\mathscr{S}_A = \{H_A, F0_A, F1_A, F2_A, F3_A, F4_A\}$, target treatment $\mathscr{S}_B = \{F0_B, F1_B, F2_B, F3_B, F4_B\}$, disease-free $\mathscr{S}_C = \{H_C, R1_C, R2_C, R3_C\}$, end-stage liver disease $\mathscr{S}_D = \{DC_D, HCC_D, LT_D, ALT_D\}$, and dead $\mathscr{S}_M = \{M\}$. Using the notation in Section 3, $\mathscr{G} = \{A, B, C, D, M\}$ and the total set of group-based health categories is denoted $\mathscr{S}_{HS} = \mathscr{S}_A \cup \mathscr{S}_B \cup \mathscr{S}_C \cup \mathscr{S}_D \cup \mathscr{S}_M$. The HCV progression models for different cohorts, gender, and treatment status are constructed using corresponding progression parameters for five population groups.

There are two intervention policies, screening for population in group A and treatment for population in group B (i.e., $\widetilde{\mathscr{G}} = \{A, B\}$). The decision variable $I^s_{A,t}$ is the fraction of candidates in group A with health category $s \in \mathscr{S}_A$ during time period t who are offered HCV screening tests, and $I^s_{B,t}$ is the fraction of patients in group B with health category $s \in \mathscr{S}_B$ who are offered treatment during time period t. The set of screening fractions is denoted as

$$I_{A,t} = \left\{ I^{H_A}_{A,t}, I^{F0_A}_{A,t} I^{F1_A}_{A,t} I^{F2_A}_{A,t} I^{F3_A}_{A,t} I^{F4_A}_{A,t} \right\},$$

and the set of treatment fractions is denoted as

$$I_{B,t} = \left\{ I_{B,t}^{F0_B} I_{B,t}^{F1_B} I_{B,t}^{F2_B} I_{B,t}^{F3_B} I_{B,t}^{F4_B} \right\}.$$

The evolution of the number of individuals in each group-based health category is expressed as a Markov model in Equation (3.2) where $\mathbf{N}_t$ is the 1-by-40 row vector of population size in each group-based health category $s \in \mathscr{S}_{HS}$ in time t, and $\mathbf{P}(I_{A,t}, I_{B,t})$ is the 40-by-40 transition probability matrix in time t based on screening and treatment allocation policy $I_{A,t}$ and $I_{B,t}$. The transition probabilities in $\mathbf{P}$ follow HCV natural history. Patients in all groups are eventually absorbed into mortality (M). The model dynamics combining disease progression and patient flow show how individuals move between these group-based health categories in a general population model (details are provided in Appendix B).

As illustrated in Fig. 3.6, individuals transition through five groups (A, B, C, D, and M) as determined by the Markov model and transition probability matrix.

Through screening intervention, individuals in group A may transition to group B (HCV+) or group C (HCV-). Those individuals not screened remain in group A and follow the disease progression. The screened individuals without the infection transition from $H_A \rightarrow H_C$. The HCV+ individuals become treatment candidates and maintain their health states ($F0_A \rightarrow F0_B$, $F1_A \rightarrow F1_B$, $F2_A \rightarrow F2_B$, $F3_A \rightarrow F3_B$, $F4_A \rightarrow F4_B$). Individuals in group B may transition to group C or group D. Those with no fibrosis may transition to the healthy state ($F0_B \rightarrow H_C$, $F0_A \rightarrow H_A$) through spontaneous virus clearance. Other patients can progress to corresponding recovered states ($F0_B \rightarrow R1_C$, $F1_B \rightarrow R1_C$, $F2_B \rightarrow R2_C$, $F3_B \rightarrow R2_C$, and $F4_B \rightarrow R3_C$) with effective treatment. With no treatment, individuals in group B may transition to group D according to disease progression. While individuals in group C were determined to be HCV-, there is a chance that they become reinfected and again rejoin group A ($H_C \rightarrow F0_A$, $R1_C \rightarrow F0_A$, $R2_C \rightarrow F2_A$, $R3_C \rightarrow F4_A$). There is a chance that individuals in any state die, i.e., transition to group M.

The high-fidelity optimization formulation for HCV management follows HQMP in Section 3.1.2 with decision variables $I_t = \{I_{A,t}, I_{B,t}\}$, $\forall t \in \{1, \ldots, n_q\}$, and the quarterly budget constraints

are described as follows,

$$\sum_{s\in\mathscr{S}_A} C_A^s I_{A,t}^s N_t^s + \sum_{s\in\mathscr{S}_B} C_B^s I_{B,t}^s N_t^s \leq B_t,\ \forall t \left\{1,\ldots,n_q\right\}, \tag{3.18}$$

where C_A^s is the cost per person in state $s \in \mathscr{S}_A$ for screening, and C_B^s is the cost per person in state $s \in \mathscr{S}_B$ for receiving treatment. The low-fidelity model for HCV management in time t follows LFM in Equation (3.7) in Section 3.1.3 with stationary decision variables $I_t = \{I_{A,t}, I_{B,t}\}$.

As in Equation (3.8), the decision variables are redefined to be the fraction of budget used for treatment during the decision period d, denoted as $x_B^{(d)} \in [0,1], \forall d \in \{1,\ldots,n_d\}$, and thus the fraction of budget used for screening is expressed as $x_A^{(d)} = 1 - x_B^{(d)}$. The quarterly budget constraints are therefore reformulated as

$$\sum_{s\in\mathscr{S}_B} C_B^s I_{B,t}^s N_t^s \quad \leq \quad x_B^{(d)} B_t, \forall t \in \mathscr{T}_d, d \in \{1,\ldots,n_d\} \tag{3.19}$$

and

$$\sum_{s\in\mathscr{S}_A} C_A^s I_{A,t}^s N_t^s \quad \leq \quad (1 - x_B^{(d)}) B_t, \forall t \in \mathscr{T}_d, d \in \{1,\ldots,n_d\} \tag{3.20}$$

which indicate that the fraction of budget allocated to screening and treatment is the same during the decision period d; however, the population may change during $t \in \mathscr{T}_d$ and consequently $I_{A,t}^s$ and $I_{B,t}^s$, $t \in \mathscr{T}_d$, may change every time period based on the pre-determined budget allocation rule within groups A and B.

The budget allocation rule in groups A and B are different. Since the health status is unknown for candidates in group A, the screening fractions in all states in group A are uniform. Therefore,

$$\begin{aligned} I_{A,t}^{H_A} = I_{A,t}^{F0_A} = I_{A,t}^{F1_A} = I_{A,t}^{F2_A} = I_{A,t}^{F3_A} = I_{A,t}^{F4_A} \\ = \quad \min\left\{ \frac{\left(1 - x_B^{(d)}\right) B_t}{\sum_{s\in\mathscr{S}_A} N_t^s C_A^s}, 1 \right\}. \end{aligned} \tag{3.21}$$

The treatment allocation rule for group B prioritizes patients with more severe liver damage first (i.e., $F4_B$ to $F0_B$). The maximum value of $I_{B,t}^s$, $s \in \mathscr{S}_B$ cannot exceed 1, and the minimum value

is 0. Thus, given the population size in group B and a value of $x_B^{(d)}$, prioritizing treatment strategy, $I_{B,t}^{F4_B}$, $I_{B,t}^{F3_B}$, $I_{B,t}^{F2_B}$, $I_{B,t}^{F1_B}$, and $I_{B,t}^{F0_B}$ is determined by Equations (3.22)-(3.26),

$$I_{B,t}^{F4_B} = \min\left\{\frac{x_B^{(d)} B_t}{C_B^{F4_B} N_t^{F4_B}}, 1\right\} \tag{3.22}$$

$$I_{B,t}^{F3_B} = \max\left\{\min\left\{\frac{x_B^{(d)} B_t - I_{B,t}^{F4_B} N_t^{F4_B} C_B^{F4_B}}{C_B^{F3_B} N_t^{F3_B}}, 1\right\}, 0\right\} \tag{3.23}$$

$$I_{B,t}^{F2_B} = \max\left\{\min\left\{\frac{x_B^{(d)} B_t - \sum_{s\in\{F3_B,F4_B\}} I_{B,t}^{s} N_t^{s} C_B^{s}}{C_B^{F2_B} N_t^{F2_B}}, 1\right\}, 0\right\} \tag{3.24}$$

$$I_{B,t}^{F1_B} = \max\left\{\min\left\{\frac{x_B^{(d)} B_t - \sum_{s\in\{F2_B,F3_B,F4_B\}} I_{B,t}^{s} N_t^{s} C_B^{s}}{C_B^{F1_B} N_t^{F1_B}}, 1\right\}, 0\right\} \tag{3.25}$$

$$I_{B,t}^{F0_B} = \max\left\{\min\left\{\frac{x_B^{(d)} B_t - \sum_{s\in\{F1_B,F2_B,F3_B,F4_B\}} I_{B,t}^{s} N_t^{s} C_B^{s}}{C_B^{F0_B} N_t^{F0_B}}, 1\right\}, 0\right\}. \tag{3.26}$$

MF-RA is applied to optimize HCV screening and treatment policies over 10 years under different annual budget constraints for all Americans aged $40-69$ years in year 2015 (born between $1945-1975$). Three birth-cohorts, i.e., 40 to 49, 50 to 59, and 60 to 69, consisting of both genders are considered, and a gender-blind policy is adopted. The total budget is divided according to the population proportions. Three annual budgets, \$1 billion/year, \$4 billion/year, and \$8 billion/year, are explored for HCV management in 10 years. The annual budget is divided equally into four quarterly budgets (i.e., \$0.25 billion/quarter, \$1 billion/quarter, \$2 billion/quarter). The lifetime horizon for the HCV progression model is 60 years, i.e., $n_q = 240$ quarters, until everyone in the cohort is dead. The budget is discretized to ten budget splits ($n_{dis} = 10$, $\left|\widetilde{\mathscr{G}}\right| = 2$ and thus $n_p = C_{2-1}^{2+10-1} = 11$) in MF-RA.

Since MF-RA only gives an approximation to an optimal solution, a grid search is evaluated to compare the quality of solutions. The grid search approach is discretized to the same level as MF-RA, that is, $n_{dis} = 10$. Grid search systematically evaluates every combination of decision variables for all decision periods. The QALYs is evaluated using (3.2) and (3.3) for $(n_p)^{n_d}$ policy combinations in the experiments.

Input data (cohort characteristics, HCV progression parameters, costs, and utilities) are summa-

rized in Appendix C. In the model, the rates of HCV progression are age- and gender-dependent. It is assumed that healthy individuals in target screening population (i.e., group A) may avoid risky behaviors so that new infection is negligible. It is also assumed that a widespread screening and treatment program may reduce individual risky behaviors so that HCV reinfection probability is linearly depreciated every two years in an aging cohort. This initial annual reinfection probability is estimated to be 0.0032. Note that the initial population distribution across different groups (groups A and B in Fig. 3.6) is calculated from US Census data and HCV infection awareness is estimated at 50%.

MF-RA is implemented to observe the trend of budget allocation under different amounts of budget. The following outputs from the numerical experiments include: 1) incremental discounted QALYs, 2) total numbers of individuals screened and treated, and 3) computation time. All numerical computations were carried out on a Intel coreI7 processor at 3.6GHz with 16GB RAM and Matlab R2015b.

The first experiment described in Section 3.4.2 determines screening and treatment policies in five decision periods ($n_{q,dec} = 8$ and $n_d = 5$), i.e., year 1, 3, 5, 7, 9, over a 10-year budget planning horizon ($n_{q,budget} = 40$ quarters). The performance and computation time for MF-RA and grid search approach is compared in Section 3.4.2. The experimental result shows that MF-RA derives similar overall QALYs as grid search with significantly improved computation time.

The second experiment described in Section 3.4.3 determines screening and treatment implementation policies in ten decision periods ($n_{q,dec} = 4$ and $n_d = 10$), i.e., year 1, 2, 3, ..., 10, over 10-year budget planning horizon ($n_{q,budget} = 40$ quarters) to demonstrate that MF-RA is applicable to a higher-dimension sequential allocation problem. The numerical results from the grid search approach were not obtained after 48 hours computation time; thus, the results from MF-RA with ten decision periods are compared to the results from Experiment 1. Results show that MF-RA is able to provide good sequential policies quickly.

The third experiment described in Section 3.4.4 conducts an analysis of HCV elimination using different annual budgets over 30 years to show that MF-RA can provide insights on controlling HCV in the long term. The planning horizon is now extended to 30 years ($n_{q,budget} = 120$ quarters) with fifteen decision periods ($n_{q,dec} = 8$ and $n_d = 15$), i.e., year 1, 3, 5, ..., 29. The experiment captures the trade-off between budget and HCV near-elimination time over three different annual budgets.

Results show that MF-RA is able to evaluate the performance of long-term disease control under different budgets.

The fourth experiment described in Section 3.4.5 conducts a sensitivity analysis of HCV elimination on different reinfection probabilities over 30 years to show that MF-RA can provide insights on managing long-term allocation when the reinfection parameter is uncertain. As in Experiment 3, the planning horizon is extended to 30 years ($n_{q,budget} = 120$ quarters) with fifteen decision periods ($n_{q,dec} = 8$ and $n_d = 15$), i.e., year 1, 3, 5, ..., 29. Results show that MF-RA is able to adjust screening and treatment budget allocation policies with reinfection probabilities. Even with higher reinfection probability, HCV is nearly eliminated within budget constraints.

3.4.2 Experiment 1: General screening and treatment setting

The results from the first experiment are summarized in Table 3.2, including the optimal percentages of budget allocated to treatment over ten years, computation time, the incremental QALYs, (i.e., the difference between the total QALYs from derived screening and treatment policies and the "do nothing" case), and the total number of individuals screened and treated under three budget scenarios for three separate birth-cohorts. Results show that, when the budget is limited to \$1 billion per year, the grid search approach allocates the entire budget to treatment during the ten budget planning years for the $50-59$ and $60-69$ birth-cohorts, while 10% of the budget is allocated to screening in the $9-10$ years for the $40-49$ birth-cohort. This is due to the fact that the initial infected population size of group B in the $40-49$ birth-cohort is smaller than its counterparts in other cohorts (initial infected population sizes in groups A and B: $547,882$ for $40-49$ birth-cohort, $1,224,436$ for the $50-59$ birth-cohort, and $981,916$ for the $60-69$ birth-cohort). In contrast, some percentage of budget starts to be allocated to screening after the 5th year when the budget is \$4 billion/year and the 3rd year when the budget is \$8 billion/year.

The MF-RA approximate solution is generally very close to the grid search optimal solution. In the \$1 billion/year case, the derived sequential polices are nearly identical (except the $60-69$ birth-cohort, where 20% of the budget is allocated to screening during $5-8$ years). In the \$4 and \$8 billion/year cases, the sequential policies are the same for the $60-69$ birth-cohort; however, for the $50-59$ birth-cohort, the policies differ slightly. It is also observed that whenever MF-RA allocates

Table 3.2: Experiment 1: Optimal percentage of budget used for treatment over first 10 years for three birth-cohorts (Overall discounted QALYs in the "do nothing" case: 741,332,012 for 40-49 birth-cohort, 551,132,227 for 50-59 birth-cohort, and 273,348,618 for 60-69 birth-cohort) (MF-RA: multi-fidelity rollout algorithm, GS: grid search)

		1 billion/year		**4 billion/year**		**8 billion/year**	
Birth-cohort	**Year**	**MF-RA**	**GS**	**MF-RA**	**GS**	**MF-RA**	**GS**
40-49	1-2	100	100	100	100	60	60
	3-4	100	100	40	10	10	10
	5-6	100	100	10	20	20	20
	7-8	100	100	10	10	10	10
	9-10	90	90	30	40	20	20
Incremental QALYs		298,442	298,442	445,565	450,536	598,892	598,892
Total number of individuals screened		233,428	233,428	18,365,254	18,699,781	41,079,735	41,079,735
Total number of individuals treated		234,765	234,765	361,328	349,930	489,659	489,659
Time (sec)		1	1,803	1	1,903	1	1,745
50-59	1-2	100	100	100	100	90	100
	3-4	100	100	90	100	100	40
	5-6	100	100	70	50	20	30
	7-8	100	100	40	30	20	20
	9-10	100	100	30	40	30	40
Incremental QALYs		281,596	281,596	651,112	655,221	907,062	918,597
Total number of individuals screened		0	0	10,714,423	10,833,129	31,629,246	32,119,345
Total number of individuals treated		190,751	190,751	582,788	578,626	849,771	833,432
Time (sec)		1	1,507	1	1,449	1	1,492
60-69	1-2	100	100	100	100	100	100
	3-4	100	100	100	100	90	90
	5-6	80	100	30	30	30	30
	7-8	80	100	20	20	20	20
	9-10	100	100	20	20	0	0
Incremental QALYs		138,173	146,821	357,715	357,715	507,516	507,516
Total number of individuals screened		584,368	135,727	13,845,329	13,845,329	22,480,130	22,480,130
Total number of individuals treated		166,404	182,052	429,866	429,866	610,979	610,979
Time (sec)		1	1,301	1	1,228	1	1,170

Note: For $60-69$ birth-cohort, even though the budget is allocated to treatment over ten years in the \$1 billion/year case, 100% population in group B is treated and thus some remaining budget is allocated to screening in the last quarter.

less budget to treatment (screening) than grid search, more budget will be allocated to treatment (screening) in the following years. Thus, the total number of individuals being screened and treated in MF-RA and grid search are similar. The incremental QALYs are the same or very close over three budget scenarios for all three cohorts because the two approaches have similar total number of individuals being screened and treated.

To demonstrate that MF-RA holds the property of sequential QALYs improvement discussed in Appendix A, Table D.1 in Appendix D presents the incremental QALYs to show that they are better or the same as MF-RA proceeds.

As shown in Table 3.2, the computation time for MF-RA is much faster than for grid search. It only takes about one second for MF-RA to obtain good sequential policies in all nine scenarios, whereas grid search takes about 20 minutes for the fastest scenario. In terms of the overall number of population computations, for the $40-49$ birth-cohort with $n_q = 240$, $n_p = 11$, and $n_d = 5$, the grid search approach needs $240 \cdot 11^5 = 38,652,240$ population computations using Equation (A.1). In contrast, MF-RA uses only $22,000$ population computations using Equation (A.2), a significant improvement. MF-RA is $\frac{38,652,240}{22,000} = 1,757$ times faster than grid search for this scenario. The ratio of numbers of quarterly population computation is close to the ratio of computation time. Thus, the number of population computations serves as a good estimate of the required computation time.

3.4.3 Experiment 2: Increased number of decision periods

In experiment 2, the number of decision periods is increased with the same 10-year budget planning horizon to observe whether the health outcomes are improved. Comparing to the results in Table 3.2, Table 3.3 shows that increasing the number of decision periods has the potential to provide better population health outcomes in some scenarios. (The incremental QALYs in bold in Table 3.3 indicate those sequential policies that are better than, or the same as, the policies in Table 3.2 under the same birth-cohort and budget.)

For example, making ten annual decisions for the $40-49$ birth-cohort under 8 billion/year provides higher incremental QALYs than its counterpart with five decision periods. This may be explained by providing some flexibility in resource allocation. Using the same example, the annual policy provided by MF-RA in years 1 and 2 is 100% and 10% budget allocation to treatment,

Table 3.3: Experiment 2: Optimal percentage of budget used for treatment over first 10 years for three birth-cohorts using MF-RA with ten decision periods (Overall discounted QALYs in the "do nothing" case: 741,332,012 for 40-49 birth-cohort, 551,132,227 for 50-59 birth-cohort, and 273,348,618 for 60-69 birth-cohort)

Birth-cohort	Year	1 billion/year MF-RA	4 billion/year MF-RA	8 billion/year MF-RA
40-49	[1,2]	[100,100]	[100,100]	[100,10]
	[3,4]	[100,100]	[80,100]	[10,20]
	[5,6]	[100,100]	[10,10]	[10,0]
	[7,8]	[100,100]	[10,0]	[10,0]
	[9,10]	[90,10]	[10,50]	[10,100]
Incremental QALYs		**298,638**	442,534	**606,773**
Time (sec)		7	7	6
50-59	[1,2]	[100,100]	[100,100]	[100,100]
	[3,4]	[100,100]	[100,80]	[60,100]
	[5,6]	[100,100]	[100,70]	[20,40]
	[7,8]	[100,100]	[100,100]	[20,0]
	[9,10]	[100,100]	[100,100]	[20,0]
Incremental QALYs		**281,596**	652,272	916,041
Time (sec)		6	7	6
60-69	[1,2]	[90,100]	[100,100]	[100,100]
	[3,4]	[100,100]	[100,100]	[90,100]
	[5,6]	[100,90]	[40,100]	[100,100]
	[7,8]	[70,100]	[20,20]	[20,0]
	[9,10]	[100,100]	[100,100]	[100,100]
Incremental QALYs		141,444	357,695	**507,516**
Time (sec)		7	6	6

Table 3.4: Experiment 3: Dollars (in billions) spent every two years (every decision period) for 40-49 birth-cohort

	Years														
Annual budget	1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20	21-22	23-24	25-26	27-28	29-30
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
4	8	8	8	8	8	8	8	8	6	0.039	0.020	0.012	0.010	0.008	0.005
8	16	16	16	16	12.4	0.077	0.060	0.052	0.045	0.004	0.002	0.001	0.001	0.001	0.001

respectively, whereas the bi-annual policy is to allocate 60% to treatment in both years.

It should be noted that MF-RA is an approximation and cannot guarantee an optimal policy. However, it is impractical to run grid search with ten decision periods. Take the \$1 billion/year budget for the $40-49$ birth-cohort as an example, the average computation time per quarterly population computation is $\frac{1,803}{38,652,240} = 4.7 \cdot 10^{-5}$ seconds, and the total number of quarterly population computation is $240 \cdot (11)^{10} \approx 6.2 \cdot 10^{12}$ by Equation (A.1). Therefore, the estimated computation time for grid search is $\left(4.7 \cdot 10^{-5}\right)\left(6.2 \cdot 10^{12}\right) \approx 2.9 \cdot 10^{8}$ seconds, which is about nine years. Similarly, the computation time for \$1 billion annual budget for the $50-59$ and $60-69$ birth-cohorts is about 7 and 6 years, respectively. Contrasting years of computation time for grid search with seconds for MF-RA illustrates the benefit of MF-RA.

3.4.4 Experiment 3: Hepatitis C control over different annual budgets

In this experiment, MF-RA is used to provide some insights on eradicating HCV infection within a targeted American birth-cohort. An analysis is conducted with different annual budgets for the $40-49$ birth-cohort to estimate the infected population size over a 30-year budget planning. The policies can change every two years and thus there are fifteen decision periods, i.e., year 1, 3, 5, ..., 29, over 30 years.

Fig. 3.7 shows the decrease in size of the infected population over time under four scenarios: no intervention, i.e., zero budget, \$1 billion/year, \$4 billion/year, and \$8 billion/year. It is observed that when there is no intervention, the infected population size slightly decreases every year mainly due to background death. In the \$1 billion/year case, almost all of the budget is allocated to treatment in the first ten years, decreasing the infected population size significantly in the first few years;

however, there are still about $200,000$ HCV infected patients in the 10th year. On the contrary, in the \$4 and \$8 billion/year cases, it takes about 18 and 10 years, respectively, to nearly eliminate the hepatitis C infection.

Table 3.4 shows actual spending every two years (i.e., every decision period). When the budget is limited to \$1billion/year, all of the annual budgets are spent; on the other hand, when the budget is \$4 and \$8 billion/year, there is leftover money after HCV is controlled. The total spending that is needed to control HCV in the cases of \$1, \$4, and \$8 billion/year budget are about \$30, \$70, and \$77 billion, respectively. In addition, Figs. D.1 and D.2 and Tables D.2 and D.3 in Appendix D demonstrate the evolution of infected population size and spending over 20 years for the $50-59$ birth-cohort and over 10 years for the $60-69$ birth-cohort.

To estimate the computation time for grid search approach with fifteen decision periods, Equation (A.1) is used to calculate the average computation time per quarterly population computation. Using \$1 billion budget per year for the $40-49$ birth-cohort as an example, the average computation time per quarterly population computation is $\frac{1,803}{38,652,240} = 4.7 \cdot 10^{-5}$ seconds, and the total number of quarterly population computation is $240 \cdot (11)^{15} \approx 1 \cdot 10^{18}$. Thus, the estimated computation time is $\left(4.7 \cdot 10^{-5}\right)\left(1 \cdot 10^{18}\right) \approx 4.7 \cdot 10^{13}$ seconds, which is about 1.5 million years, implying that to derive the optimal elimination policy for 30 years using grid search is practically impossible.

3.4.5 Experiment 4: Sensitivity analysis of hepatitis C elimination on reinfection probabilities

Uncertain parameters in the population disease model will lead to inaccurate evaluation of QALYs, and calculation of time to eliminate disease. In this experiment, reinfection probability is used as an example of an uncertain input, and MF-RA is applied with a high and low reinfection probabilities to establish a range on time to eliminate HCV. A sensitivity analysis using two different reinfection probabilities for the $40-49$ birth-cohort is conducted to estimate the infected population size over a 30 year budget planning horizon where the policies can change every two years. The low and high annual reinfection probabilities are set to 0.001 and 0.039, respectively, based on values in [101] (the value of 0.0032 is used in Experiments 1, 2, and 3).

Fig. 3.8(a) shows that when there is no intervention, the infected population decreases over time due to mortality, as in Fig. 3.7 (with no intervention). Figs. 3.8(b) and 3.8(c) show the infected

Table 3.5: Experiment 4: Summary of sensitivity analysis of hepatitis C elimination on reinfection probabilities

		Low reinfection prob.	**High reinfection prob.**
1 billion	Time (sec)	13	15
	Total number of individuals screened	13,707,466	13,499,173
	Total number of individuals treated	270,269	276,655
4 billion	Time (sec)	15	16
	Total number of individuals screened	38,050,952	38,044,588
	Total number of individuals treated	428,144	521,723
8 billion	Time (sec)	19	18
	Total number of individuals screened	41,077,269	41,211,215
	Total number of individuals treated	487,714	653,032

population trend over 30 years when the budget is limited to \$1 billion per year. It is observed that similar to the case of no intervention, it takes 30 years to control HCV, no matter whether the reinfection probability is low or high. Similar to the results in Section 3.4.2, under both cases of low and high reinfection probabilities, all of the budget is allocated to treatment in the first eight years; however, there is a significant difference in long-term budget allocation between low and high reinfection probabilities when the budget is low. When there is a low reinfection probability, the size of group B (known HCV+) in the beginning of year 11 is nearly zero; however, with a high reinfection probability, the size of group B remains noticeable until year 15. The policy in years $9-10$ for the low reinfection probability case is allocating 90% of the budget to treatment whereas only 50% of the budget is allocated to treatment in the high reinfection probability case.

Figs. 3.8(d) and 3.8(e) show the infected population trend over 30 years when the budget is limited to \$4 billion per year. Both cases start to allocate some budget to screening in the third year. It is observed that it takes about 12 years in the case of low reinfection probability and 14 years in the case of high reinfection probability to eliminate the known HCV infection in group B, and both cases take about 18 years to control HCV. Figs. 3.8(f) and 3.8(g) show similar results when the budget is limited to \$8 billion per year. In both low and high cases, the policies obtained from MF-RA start to allocate some budget to screening in the first year. It takes about 10 years to eliminate the known HCV infection for both cases and the sequential policies are almost the same. From Figs. 3.8(d)-3.8(g), even though low and high reinfection probabilities produce similar sequential policies, a higher budget percentage is allocated to treatment in the case of high reinfection probability.

A higher reinfection probability results in more individuals transiting from group C to group A. Intuitively, it indicates that higher budget should be allocated to screening in order to identify the newly reinfected patients; however, the derived policies may be counter to the idea. For three different budget scenarios, the total number of treated patients (screened individual) is higher (lower) in the case of high reinfection probability than low reinfection probability (the details are shown in Table 3.5). The reason could be that reinfected individuals are patients in more severe health categories, which increase the needs for immediate treatment.

3.4.6 Experiment 5: Hepatitis C elimination study using RA-BnB

Now, apply RA-BnB to obtain an optimal screening and treatment policy under budget constraints for chronic hepatitis C birth-cohorts. Birth-cohorts consist of all Americans aged $40-69$ years in year 2015 (born between $1945-1975$). Consider three birth-cohorts separately, i.e., 40 to 49, 50 to 59, and 60 to 69. The HCV population evolution model is based on a validated HCV progression model in the previous research [80, 98, 99, 100, 96]. There are two intervention policies, i.e., HCV screening in a target population group with unknown disease status, and HCV treatment in the identified HCV+ population group. To show that RA-BnB is capable of providing insights on HCV care management, conduct separate analyses using \$1 billion, \$4 billion, and \$8 billion annual budgets. The budget planning horizons for $40-49$, $50-59$, and $60-69$ birth-cohorts are 30, 20, and 10 years, respectively. The model cycle period t is three months and policy can be changed every two years, e.g., year 1, 3, 5, ..., 29 for a 30 years budget planning horizon. Since budget is set to zero after the budget planning horizon, the numbers of valid budget allocation decision periods are 15, 10, and 5 for the three different birth-cohorts. The budget is discretized to ten budget pieces ($m_B = 10$, $\left|\widetilde{\mathscr{G}}\right| = 2$ and $m_P = C(2+10-1,\ 2-1) = 11$). A stopping threshold of zero ($\delta = 0$) is used in RA-BnB. All numerical computations were carried out on a Intel coreI7 processor at 3.6GHz with 16GB RAM and Matlab R2015b.

Table 3.6 provides the CPU time required by RA-BnB and the number of policies pruned over iterations for age $40-49$, $50-59$, and $60-69$ birth-cohorts. The performance measure $(Q_U^{(d)}-Q^*)/(Q^*-Q_{no\,budget})$ gradually reaches 0 in all nine cases, indicating that the output incumbent policy is the same as the optimal policy, and thus RA-BnB is able to find the optimal budget allo-

cation policy. In the $40-49$ birth-cohort, even though it takes about 9 hours under the \$1 billion case in 15 iterations to find the optimal policy, the incumbent lower bound and upper bound are very close after the 8th iteration. If the stopping threshold is $\delta = 0.005$, RA-BnB terminates in about 1.5 hours. On the other hand, the paper [80] showed that for each budget scenario, enumerating all policies for a 30 years budget planning horizon with 15 decision periods necessitates about $1.5 \cdot 10^6$ years, which implies that to obtain the optimal policy for 30 years using a grid search approach is practically impossible. Figure 3.9 shows the upper bounds and incumbent lower bounds over fifteen iterations for the 40-49 birth-cohort. The upper bound and incumbent lower bound are reasonably close in the first few iterations.

Due to the optimality guarantee, RA-BnB can provide useful insights for policymakers. Using the $40-49$ birth cohort as an example, Figure 3.10 shows the size of infected population over time under no budget, \$1, \$4, and \$8 billion per year. It is observed that when there is no budget, the infected population size slightly decreases every year mainly due to background death until the 30th year. In the \$1 billion/year case, the infected population size is significantly reduced in the first few years; however, there are still more than $200,000$ HCV infected patients in the 11th year, and HCV could not be eliminated until the 30th year. On the other hand, in the \$4 and \$8 billion/year cases, it takes about 19 and 11 years to eliminate the hepatitis C infection, respectively.

Using the $40-49$ birth-cohort as an example, Figure 3.11 shows the optimal fraction of population for treatment and screening. The result in Figure 3.11(a) indicates that when the budget is limited to \$1 billion/year, the fraction of population in target group A receiving screening is 0, i.e., all the budgets are allocated to treatment group, in the first nine years due to immediate increase in health utility. After the 9th year, both fractions keep increasing, i.e., some budget starts to be allocated toward screening. The reason is that without prior screening, the patients in compensated cirrhosis health category in group A are unaware of their health condition, resulting in entering to untreated health categories in group D such as decompensated cirrhosis, hepatocellular carcinoma, or requiring liver transplant, which incur huge health utility loss. Figures 3.11(b) and (c) show that screening begins earlier when the total budget is larger. All the individuals in group A and group B are able to receive interventions after the 19th and the 11th year as the infected population size in groups A and B become stable (refer to Figure 3.10).

Table 3.6: CPU seconds for allocation problem (30-year with 15 decision-periods budget planning for 40-49 birth-cohort, 20-year with 10 decision-periods budget planning for 50-59 birth-cohort, 10-year with 5 decision-periods budget planning for 60-69 birth-cohort)

		1 billion/year			**4 billion/year**			**8 billion/year**		
Birth cohort	**Iter.**	**Time (s)**	**# prune**	**Performance measure**	**Time (s)**	**# prune**	**Performance measure**	**Time (s)**	**# prune**	**Performance measure**
	1	1	0	0.6	1	1	0.04	1	2	0.55
	2	5	0	0.3	4	53	0.03	4	51	0.01
	3	47	700	0.09	26	417	0.02	25	388	0.004
	4	264	6,660	0.03	128	2,655	0.02	102	1,909	0.0005
	5	592	15,987	0.02	408	10,240	0.01	388	10,950	0
40-49	6	1,070	30,345	0.013	742	20,150	0.006	–	–	–
	7	1,637	45,311	0.008	927	25,692	0.002	–	–	–
	8	3,011	82,664	0.005	1,098	31,051	0.0002	–	–	–
	9	6,213	160,492	0.0027	1,210	34,892	0	–	–	–
	⋮	⋮	⋮	⋮	–	–	–	–	–	–
	15	32,667	826,991	0	–	–	–	–	–	–
	1	0	0	0.3	1	0	0.4	0	2	0.1
	2	3	33	0.3	6	43	0.1	3	56	0.04
	3	27	588	0.24	35	644	0.05	16	397	0.02
	4	133	4,559	0.13	130	2,911	0.03	57	1,577	0.006
	5	278	9,971	0.09	307	7,608	0.014	190	7,228	0.0005
50-59	6	493	19,511	0.03	757	19,098	0.005	191	7,260	0
	7	526	21,044	0.0027	2,114	68,823	0.0012	–	–	–
	8	528	21,111	0.0008	2,370	78,163	0.00002	–	–	–
	9	530	21,131	0.0002	2,546	84,562	0	–	–	–
	10	549	21,131	0	–	–	–	–	–	–
	1	0	6	0.13	0	1	0.3	0	5	0.09
	2	1	46	0.13	2	85	0.06	2	56	0.02
60-69	3	4	173	0.12	7	340	0.02	4	159	0.002
	4	11	538	0.09	13	591	0.005	16	833	0
	5	20	538	0	28	591	0	–	–	–

Note 1: '–' indicates that the optimal policy is obtained due to incumbent lower bound = upper bound.

3.5 Conclusions

The performance of MF-RA and RA-BnB were demonstrated with HCV birth-cohort screening and treatment policy implementation over budgetary planning cycle. The numerical results showed that sequential policies and the overall health benefits obtained from MF-RA and grid search are nearly identical under various annual budget and birth-cohort settings. Most importantly, MF-RA significantly outperformed grid search in computation time, especially for the high-dimensional scenario. The improved computation time enables studies with more decision periods that could not previously be addressed.

The proposed MF-RA is used to estimate the required budget and time to mitigate or even eradicate the disease in the target population. On the other hand, RA-BnB allows healthcare decision makers to examine the trade-off between policy quality and necessitated computation effort. The numerical results show that, $1 billion annual budget to address HCV is likely insufficient to have the desired impact on reducing the new incidence of infection for each age-cohort baby-boomer. To eliminate HCV for baby boomers, the government should consider increasing budget as well as implement more effective screening and treatment interventions.

The proposed sequential decision-making methodology may provide policy insights for two important healthcare systems with high HCV prevalence after model calibration with system-specific parameters; both of which address the healthcare needs of vulnerable sub-populations in the U.S. The first system is the Veterans Health Administration (VHA). Over 189,000 veterans were diagnosed with chronic HCV between years 2000 to 2008 [151]. The second system is healthcare for incarcerated populations. Over $500,000$ estimated individuals living with chronic HCV infections are currently incarcerated in federal, state and local correctional facilities [34, 91, 135].

Two limitations of the study should be noted when interpreting the results. An implicit assumption that the individuals' health categories are fully observable; however, for some diseases, such data is difficult to observe and estimate. Partially observable MDP (POMDP) modeling frameworks could be applied to capture partial observability of the individual disease progression and imperfect test information [10, 136]. The second limitation is that this work used a closed population (i.e., birth-cohort) in the formulation. The setting of the closed population in the study highlights the fundamental characteristic that the disease is more prevalent among specific population cohorts;

some healthcare interventions, however, target specific age groups, where the model must include individuals aging in and out of the target age group at each time period. In these cases, an open population provides more desired modeling of disease dynamics and propagation analysis.

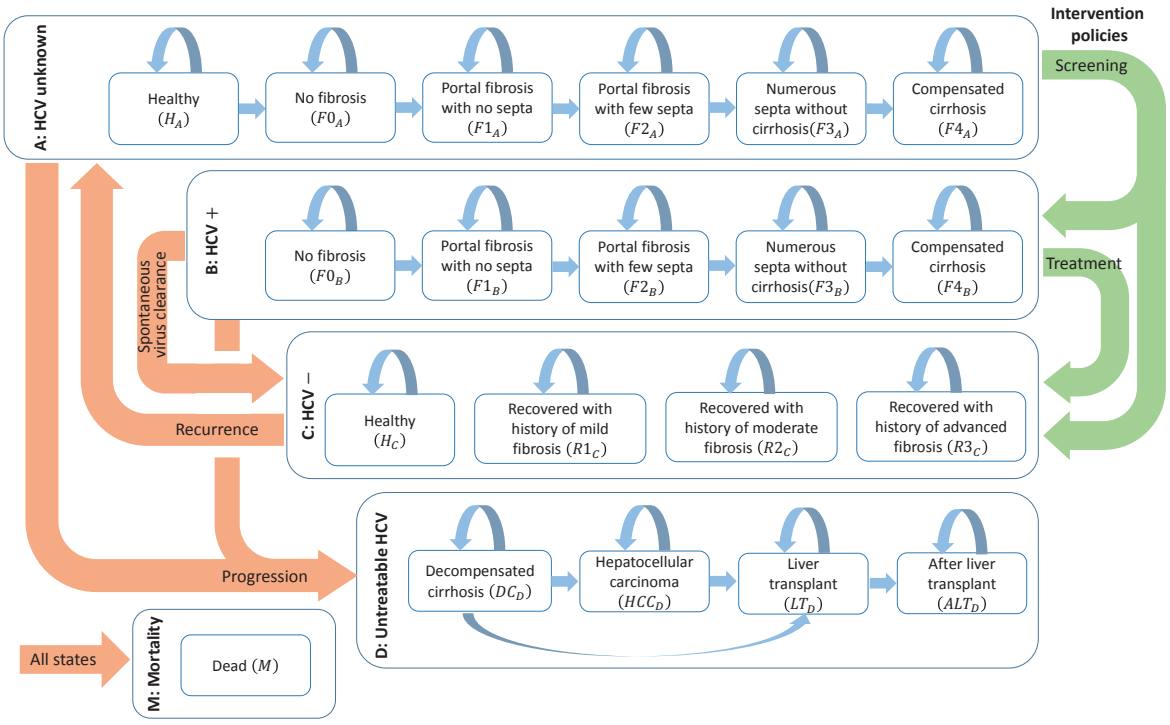


Figure 3.6: Progression model of hepatitis C with screening and treatment intervention policies

Note 1: Similar to earlier work in modeling the progression of chronic HCV infection [124], a patient's health status is evaluated through their liver conditions, including: healthy without HCV (H), no fibrosis $(F0)$, portal fibrosis with no septa $(F1)$, portal fibrosis with few septa $(F2)$, numerous septa without cirrhosis $(F3)$, compensated cirrhosis $(F4)$, recovered with history of mild fibrosis $(R1)$, recovered with history of moderate fibrosis $(R2)$, recovered with history of advanced fibrosis $(R3)$, decompensated cirrhosis (DC), hepatocellular carcinoma (HCC), liver transplant (LT), after liver transplant (ALT), and dead (M).
Note 2: Individuals with different health status exist in different groups. Population is divided into five groups, including group A: target screening population, candidates whose HCV status are unknown, group B: HCV+ population, patients who know they are HCV+ and are potential treatment candidates and patients who are under-going treatment, group C: HCV- population, patients who have been cured of HCV or individuals who know they are not HCV infected through screening, group D: Untreatable population, patients who have HCV but are not able to recover by taking only antiviral treatment, group M: Mortality, those who are dead.

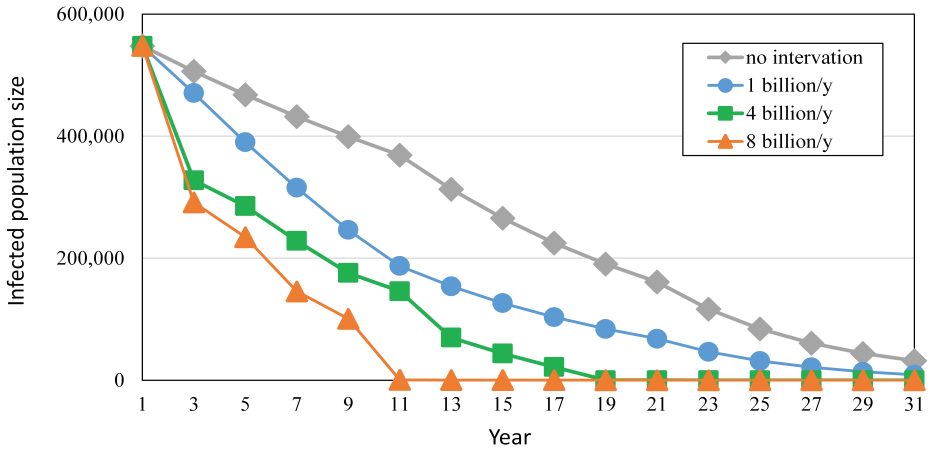
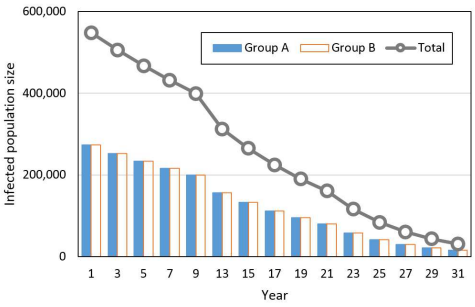
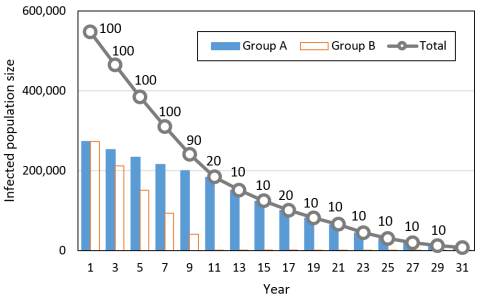


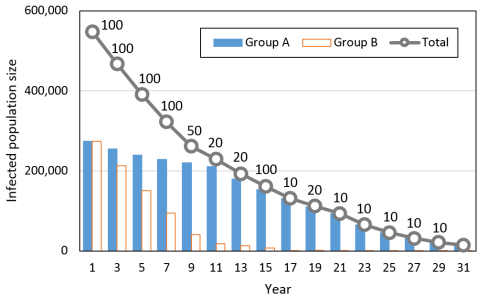
Figure 3.7: Experiment 3: HCV control over 30 years for 40-49 birth-cohort



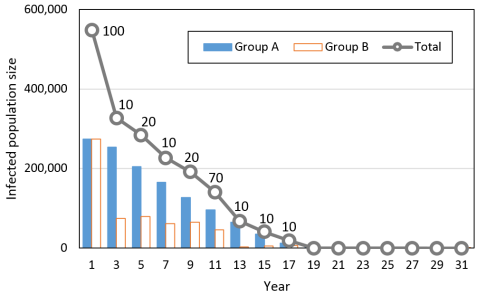
(a) No policy intervention



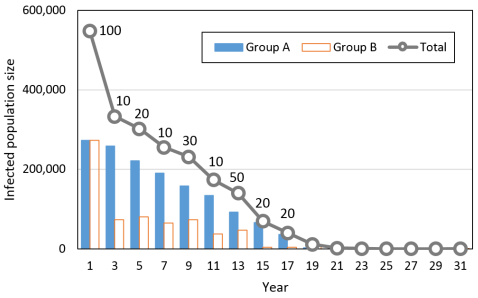
(b) $1 billion/year for low reinfection probability



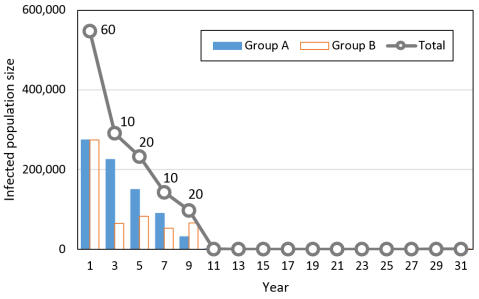
(c) $1 billion/year for high reinfection probability



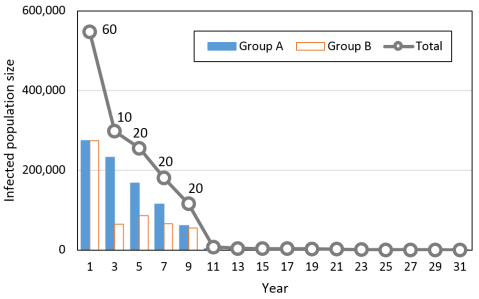
(d) $4 billion/year for low reinfection probability



(e) $4 billion/year for high reinfection probability



(f) $8 billion/year for low reinfection probability



(g) $8 billion/year for high reinfection probability

Figure 3.8: Experiment 4: Sensitivity analysis of HCV elimination on reinfection probability for 40-49 birth-cohort

Note 1: The numbers next to the circles are percentages of budget allocated to treatment. The circles without numbers in later years indicate that there is enough budget to screen and treat everyone in the groups *A* and *B*, respectively.

Figure 3.9: Upper bound and incumbent lower bound over 15 decision periods for 40-49 birth-cohort

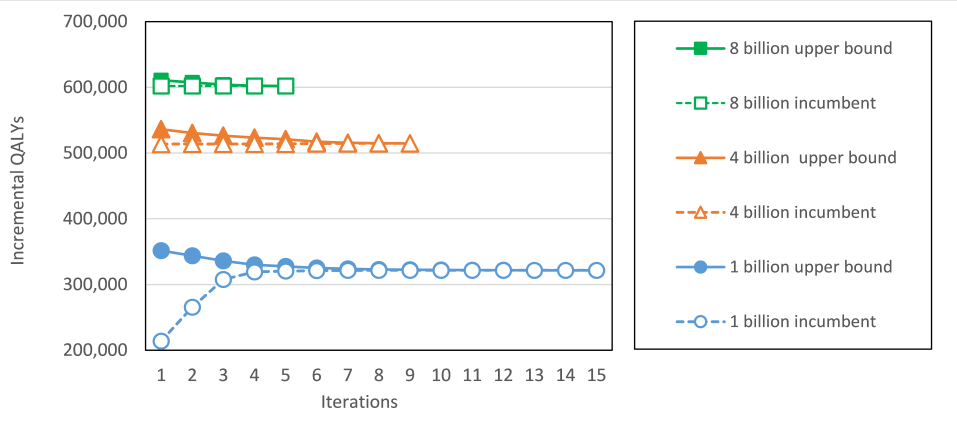


Figure 3.10: HCV control over 30 years with 15 decision periods for 40-49 birth-cohort

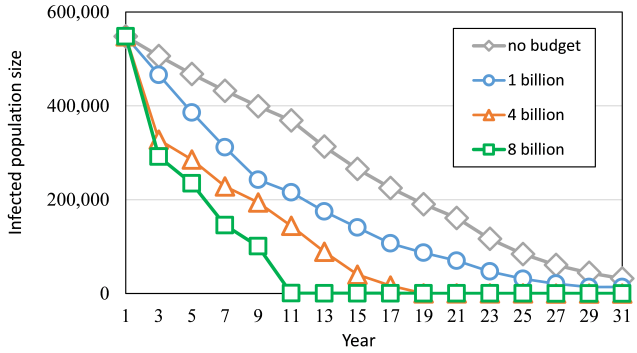
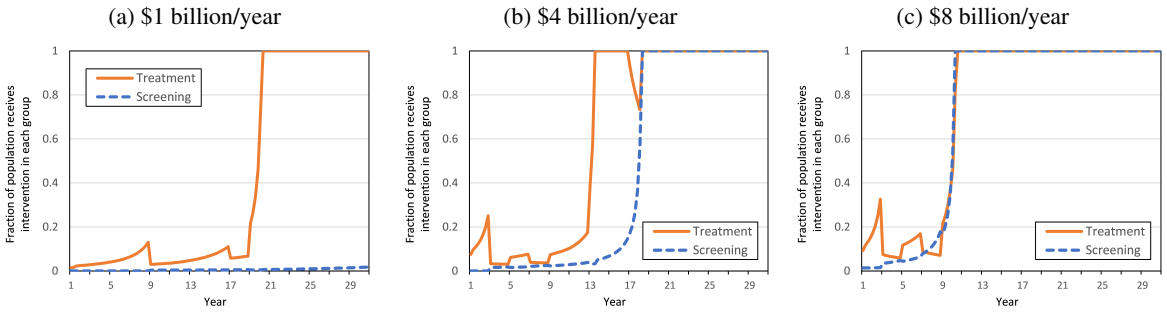


Figure 3.11: Optimal fraction of population for treatment and screening for 40-49 birth-cohort



Chapter 4

OPTIMAL HEALTH INSURANCE DESIGN FOR HYPERTENSION

This chapter describes the second healthcare application: *Design a long-term optimal healthcare insurance cost-sharing policy for insurants with hypertensive heart disease to minimize public/private insurer's and insurants' respective costs*. The problem consists of determining optimal cost-sharing policies for a public/private insurer as well as an insured's optimal decision on multiple treatments based on healthcare expenditures and health status to minimize respective costs. The problem is modeled as a Stochastic Stackelberg Game (SSG), which captures the long-term interactions between an insurer and an insured subject to underlying uncertainties of disease progression (the transition of the insurants across health states). The analysis results show that insurants' behavior regarding medical utilization and treatment adherence complements related case study reviews. To the best of our knowledge, this is the first mathematical model considering the association between cost-sharing and dynamic health outcomes for chronic diseases.

Most of this chapter appeared in [76].

4.1 Stochastic Stackelberg Game

4.1.1 Markov stationary stochastic Stackelberg game model

Consider a medical system with health insurance as a network with an insurer, group of insurants, and money and information flowing from one party to another. For modeling purpose, It is assumed that the insurants' health status/outcomes can be unambiguously observed; thus, the insurer's design of the cost-sharing mechanism is based on not only the cost of the healthcare provided to insurants, but also the observations of their health status. This is a reasonable assumption because, in a hypertension case, the insurant is a group of patients with similar characteristics and those characteristics are known to the insurer. As is known in the literature, insurants with generous insurance spend more on medical care than insurants with less generous insurance; however, to control chronic disease and reduce the cost of expensive treatment in the future, it may be beneficial to the insurer to

provide incentives for the insurant to maintain a chronic disease. Thus, the insurer needs to trade off the cost-savings from a high level of cost-sharing that can reduce medication adherence or even stop treatment with a low level of cost-sharing that reduces the risk of future expense.

It is worth noting that although asymmetric information between insurers and insurants is one of the key factors driving how insurance products are offered and priced, and has been a focus of economic analysis of health care since Arrow's seminal paper from 1963 [7], adverse selection and moral hazard are not the focus in this problem.

Adverse selection happens when insurers attract individuals with high risk, as opposed to a balanced risk group. Adverse selection is primarily affected by premiums, which are important for decisions to enroll in insurance. Premiums and adverse selection are not an issue for the question attempted to address in this chapter because the insurer is assumed to already know the risk pool of individuals they are facing when structuring a cost-sharing arrangement due to the requirement of submitting a rate schedule and cost-sharing arrangements to a regulatory body (usually the Insurance Commissioner). Also, the insurance premium is a fixed cost and does not incent treatment for those insurants (e.g., type of treatment or type of prescription drug). The only factor driving the choice for the treatment on the margin is the level of cost sharing. The focus is on what the insurer can do to incent an insurant to seek appropriate treatment for a chronic disease with cost-sharing. In this chapter, an analysis is presented that demonstrates that insurers may set cost-sharing rates to increase the likelihood that insured individuals select treatments that have mutually beneficial outcomes for the insurer and the insurant.

Moral hazard describes a situation where a party to a contract acts against the interests of the other party but still within the confines of the agreement, e.g., when an insurant uses more services than they were expected to use but to which they are entitled or when the physician's recommendation is to utilize more expensive but less effective treatments. However, there is no moral hazard if those insurants behave consistently with the expectations. Moral hazard may only be an issue when individuals with no chance of positively responding to treatment in fact seek treatment. Considering hypertension, an individual with normal blood pressure would not choose to take antihypertension medication even when cost-sharing is low because it is unnecessary and does not improve their health. Even though the issue of moral hazard from the service providers' side is significant in insurance design literature [55, 167, 168], this research emphasizes the importance of improving an

insurant patient's long-term treatment seeking behavior, which is mainly influenced by factors like income, disease burden, and cost-sharing [24, 160]. These three factors are captured in the game model through the immediate cost functions. Thus, the moral hazard problem is not an issue in the chronic disease scenario discussed in this research.

This model treats the followers in this SSG as a homogeneous group with similar characteristics, e.g., age or risk measures such as smoking status, BMI. Heterogeneity is not an issue for two reasons. First, the issue of heterogeneity relates to selection and this model may be applied to any number of insurant risk distributions. Second, the insurant heterogeneity and the choice to buy insurance is more closely linked to insurance premiums. Premiums may become relevant in an insurant's decision to buy insurance in a subsequent period but will have no impact on the choice to seek treatment for a medical care need. In fact, even though insurance providers consider cost-sharing in conjunction with premium levels as they create insurance products and offer them with the intent of attracting specific risk profiles, evidence shows that only cost-sharing impacts the likelihood that individuals seek treatment independent of premiums [27].

Here, a two-player SSG is applied where the insurer is the leader and the insurant is the follower to investigate the long-term interaction between the insurer and the insurant. The leader takes the first step, i.e., announcing a cost-sharing policy in health insurance, and the follower makes the best response regarding healthcare treatment. The leader interacts with the follower in *repeated* plays that dynamically evolve the health state. The evolution of health status depends on the probabilistic outcomes of the follower's actions in each decision epoch. Thus, the stochastic Stackelberg game presented is a repeated game.

The focus is on Markovian strategies; a common approach in health care modeling. In dynamic decision-making models for chronic disease management, most research assumes Markovian healthcare intervention [94, 130, 140]. An infinite-horizon approach is used for treatment decisions and cost-sharing policy even though the insurant will die in a finite amount of time. This is a reasonable approach since the time between decision epochs is short in comparison with the length of the overall horizon in which decisions are made. Moreover, infinite-horizon models and corresponding actions (interventions and cost-sharing policy) are often assumed to be time-independent (stationary) in medical decision-making and disease management problems [2, 125, 130, 134, 136]. In summary, the repeated game setup is limited to Markov stationary strategies. That is, the insurants

(followers) choose whether to take treatment or not depending only on their current health state and stick to that action when in the same health state in the future. The insurer (leader), on the other hand, chooses a cost-sharing policy based on the determined probability distribution.

The SSG game has a finite state space S, referring to the current health state of the insurant. The players, i.e., insurer and insurant, adopt a pair of actions, causing health state transitions in a Markov manner. The SSG game is applied to the case of hypertension, and the following notation is used throughout the chapter.

- $\mathscr{S}$: Finite set of representation of disease progression over insurant's health status. The health state ($s \in \mathscr{S}$) indicates the insurant's current medical condition. For the hypertension case, the classification of blood pressure for adults is used to determine four health states: "normal" (for those with blood pressures below 120 mmHg systolic and/or below 80 mmHg diastolic); "pre-hypertension" (for those with blood pressures ranging from 120-139 mmHg systolic and/or 80-89 mmHg diastolic); "hypertension stage 1" (for those with blood pressures ranging from 140-159 mmHg systolic and/or 90-99 mmHg diastolic); and "hypertension stage 2" (for those with blood pressures ranging from 160-179 mmHg systolic and/or 100-109 mmHg diastolic) [41].

- $\mathscr{A}_L$ and $\mathscr{A}_F$: Finite action spaces for the insurer (L) and insurant (F), respectively. For the insurer, the action ($a_l \in \mathscr{A}_L$) is the level of cost-sharing in health insurance, e.g., 10, 40, or 70 for copayment strategies. Let the cost-sharing level be different in different health states. For the insurant, the action ($a_f \in \mathscr{A}_F$) is the medical treatment s/he decides to take (e.g., do nothing, adjust lifestyle, and take antihypertension treatment). In general, the insurer could consider different types of cost-sharing rate (both coinsurance and copayment beneficiaries) in different states in the SSG game.

- $T_{ss'}^{a_f}$: State transition function from s to $s^{'}$ based on the action $a_f \in \mathscr{A}_F$ performed by the insurant. The transition function $T_{ss'}^{a_f} : \mathscr{A}_F \times \mathscr{S} \times \mathscr{S} \to [0,1]$ indicates the transition probability from health state s to $s^{'}$ given the actions performed by the insurant, e.g., the change in health status is only associated with the action of the insurant. In general, the probability

can be extended to $T_{ss'}^{a_l a_f}$ if disease progression depends on actions of both the insurer and the insurant.

- $\beta(s)$: Probability that the initial health state is $s \in S$, with $\sum_{s\in S}\beta(s) = 1$.

- $\pi(a_l|s)$ and $\phi(a_f|s)$: $\pi(a_l|s)$ is the probability the insurer takes action a_l when the insurer observes health state s and $\phi(a_f|s)$ is the probability the insurant takes action a_f when the insurant observes state s. It is reasonable to assume that the behavior of the insurant is deterministic (pure strategy) since the action of the insurer is given [42]. That is, in every state $s \in S$, $\phi(a_f|s) = 1$ if the insurant chooses treatment a_f, and 0 otherwise.

- $r_L(s,a_l,a_f)$ and $r_F(s,a_l,a_f)$: Immediate cost functions for the insurer and the insurant, respectively, in state s under the actions $a_l \in \mathscr{A}_L$ and $a_f \in \mathscr{A}_F$. Note that immediate cost functions are formulated based on different types of cost-sharing policies.

- γ_L and γ_F: Discount factors for the insurer and the insurant, respectively, satisfying $0 \leq \gamma_L, \gamma_F < 1$.

- $R_L(s,\pi,a_f)$ and $R_F(s,\pi,a_f)$: Insurer and insurant's expected costs in state s under the actions of the insurant $a_f \in \mathscr{A}_F$, respectively.

- $v_L(s)$ and $v_F(s)$: Insurer and insurant's optimal expected cost functions with initial state $s \in S$, respectively.

With the notation defined above, an illustration of the SSG with progression of hypertensive heart disease and corresponding treatments is shown in Figure 4.1, where the health states and actions are summarized as

$$\mathscr{S} = \{s_1\text{: normal, } s_2\text{: prehypertension, } s_3\text{: hypertension stage 1, } s_4\text{: hypertension stage 2}\},$$

$$\mathscr{A}_L = \{a_{l1}\text{: \$10 copayment, } a_{l2}\text{: \$40 copayment, } a_{l3}\text{: \$70 copayment}\},$$

and $\mathscr{A}_F = \{a_{f1}\text{: do nothing (D), } a_{f2}\text{: lifestyle adjustment (L), } a_{f3}\text{: antihypertensive treatment (T)}\}$.

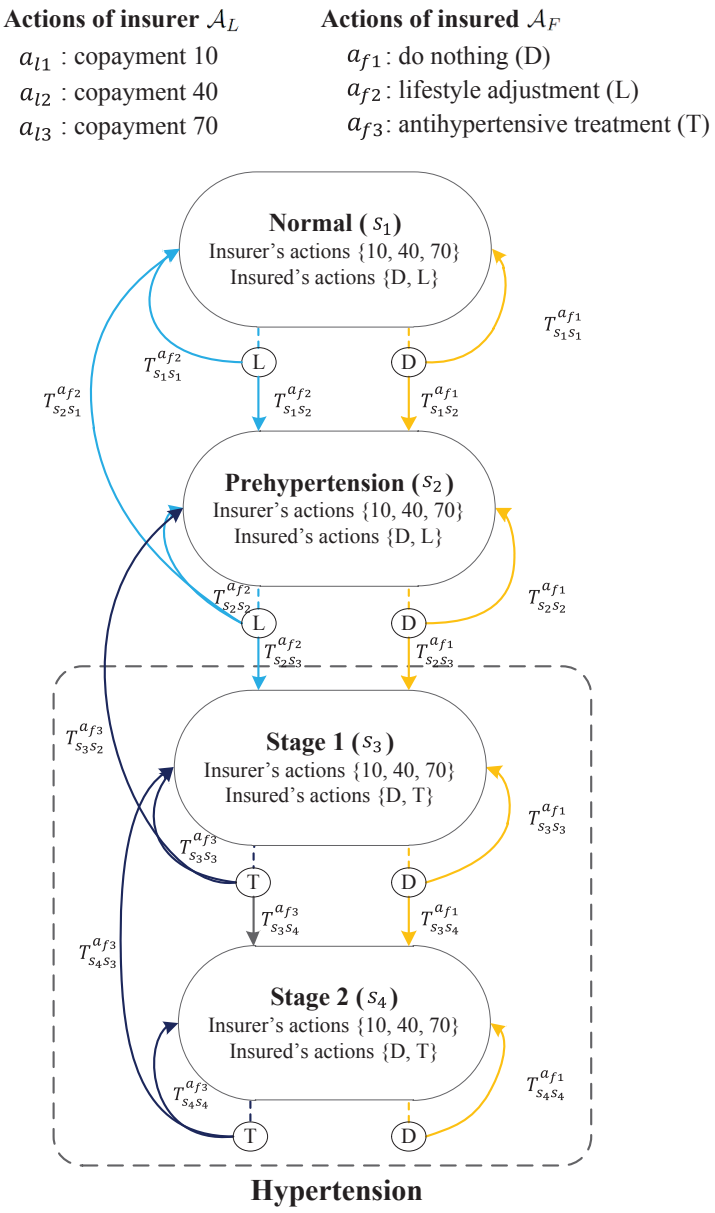


Figure 4.1: Antihypertension treatments, health states, and disease progression using a stochastic Stackelberg game model

The insurant's action set is determined by his/her health states, e.g., the insurant can do nothing (D) or make a lifestyle adjustment (L) in the normal state (s_1) and in the prehypertension state (s_2), while the action sets in hypertension stage 1 (s_3) and stage 2 (s_4) include do nothing (D) and

take antihypertension treatment (T). There are seventeen $T_{ss'}^{a_f}$ state transition probabilities because available actions in each state are different: $T_{s_1s_1}^{a_{f1}}$, $T_{s_1s_2}^{a_{f1}}$, $T_{s_1s_1}^{a_{f2}}$, $T_{s_1s_2}^{a_{f2}}$, $T_{s_2s_2}^{a_{f1}}$, $T_{s_2s_3}^{a_{f1}}$, $T_{s_2s_1}^{a_{f2}}$, $T_{s_2s_2}^{a_{f2}}$, $T_{s_2s_3}^{a_{f2}}$, $T_{s_3s_3}^{a_{f1}}$, $T_{s_3s_4}^{a_{f1}}$, $T_{s_3s_2}^{a_{f3}}$, $T_{s_3s_3}^{a_{f3}}$, $T_{s_3s_4}^{a_{f3}}$, $T_{s_4s_4}^{a_{f1}}$, $T_{s_4s_3}^{a_{f3}}$, and $T_{s_4s_4}^{a_{f3}}$. For example, an insurant in hypertension stage 1 (s_3) may recover to normal state (s_1) by receiving antihypertension treatment and adjusting lifestyle in two consecutive health states ($s_3 \rightarrow s_2 \rightarrow s_1$). Doing nothing will not improve his/her health status, but the health status may worsen, i.e., ($s_3 \rightarrow s_3$ or $s_3 \rightarrow s_4$). Note that actions of the insurer do not affect the health state transition directly. The insurer, i.e., the leader, chooses a cost-sharing policy from 10, 40, and 70 copayments in each health state before the insurant takes an action.

Immediate cost functions

Immediate cost functions for the insurer, $r_L(s,a_l,a_f)$, and the insurant, $r_F(s,a_l,a_f)$, are formulated based on different types of cost-sharing plans. insurants' long-term treatment seeking behaviors are influenced by three main factors: income, disease burden, and cost-sharing [24, 160]. Thus, the immediate cost functions incorporate monthly income, mortality rate, and monthly out-of-pocket payment. In the hypertension case, the insurer's and the insurant's immediate cost function using copayment as an example are as follows,

$$r_L(s,a_l,a_f) = \left(c^{treat}(s,a_f) - s_{CP}(a_l,a_f)\right) + c_L^{effect}(s,a_f), \tag{4.1}$$

and

$$r_F(s,a_l,a_f) = s_{CP}(a_l,a_f) + c_F^{effect}(s,a_f) - b_F^{health}(s,a_f) \tag{4.2}$$

where $c^{treat}(s,a_f)$ is the total cost of medical treatment according to insurant's action a_f in state s, e.g., payment/fee of different healthcare service/treatments, $s_{CP}(a_l,a_f)$ is the copayment level associated with the insurer's copayment setting $a_l \in \mathscr{A}_L$ and the insurant's action $a_f \in \mathscr{A}_F$, $c_L^{effect}(s,a_f)$ and $c_F^{effect}(s,a_f)$ denote the relative costs of side effects from healthcare service or behaviors with action $a_f \in \mathscr{A}_F$ to the insurer and insurant in state s, respectively, e.g., the low quality of different medical service, undesirable effect of a drug, life expectancy, or reduction of the risk of mortality, and $b_F^{health}(s,a_f)$ is the health benefit for the insurant, e.g., income, in state $s \in \mathscr{S}$ with action $a_f \in \mathscr{A}_F$.

Expected cost functions

The expected cost functions for the insurer and the insurant are in terms of immediate costs and future costs. Suppose that the insurant's health state is s, the insurer adopts policy $\pi(a_l|s)$ and the insurant adopts action $a_f \in A_F$. The insurer's and the insurant's expected costs, $R_L(s,\pi,a_f)$ and $R_F(s,\pi,a_f)$, are shown as follows,

$$R_L(s,\pi,a_f) = \sum_{a_l \in A_L} \pi(a_l|s) \left(r_L(s,a_l,a_f) + \gamma_L \sum_{s' \in S} T_{ss'}^{a_f} v_L\left(s'\right) \right) \tag{4.3}$$

and

$$R_F(s,\pi,a_f) = \sum_{a_l \in A_L} \pi(a_l|s) \left(r_F(s,a_l,a_f) + \gamma_F \sum_{s' \in S} T_{ss'}^{a_f} v_F\left(s'\right) \right) \tag{4.4}$$

where γ_L and γ_F are discount factors, and $v_L\left(s'\right) = \min_{\pi,a_f} R_L(s',\pi,a_f)$ and $v_F\left(s'\right) = \min_{\pi,a_f} R_F(s',\pi,a_f)$ are optimal expected costs going forward from a future state s' for the insurer and the insurant. Therefore, a Stackelberg equilibrium computation problem is formulated as a mixed integer nonlinear program (MINLP) as follows:

$$\min_{\pi,\phi,v_L,v_F} \sum_{s \in \mathscr{S}} \beta(s) v_L(s) \tag{4.5}$$

s.t.

$$\pi(a_l|s) \geq 0 \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L \tag{4.6}$$

$$\sum_{a_l} \pi(a_l|s) = 1 \qquad \forall s \in \mathscr{S} \tag{4.7}$$

$$\phi(a_f|s) \in \{0,1\} \qquad \forall s \in \mathscr{S}, a_f \in \mathscr{A}_F \tag{4.8}$$

$$\sum_{a_f} \phi(a_f|s) = 1 \qquad \forall s \in \mathscr{S} \tag{4.9}$$

$$0 \leq R_F(s,\pi,a_f) - v_F(s) \leq (1-\phi(a_f|s))Z \qquad \forall s \in \mathscr{S}, a_f \in \mathscr{A}_F \tag{4.10}$$

$$R_L(s,\pi,a_f) - v_L(s) \leq (1-\phi(a_f|s))Z \qquad \forall s \in \mathscr{S}, a_f \in \mathscr{A}_F \tag{4.11}$$

The Stackelberg equilibrium is defined as (π^*, ϕ_π^*), i.e., the optimal policies derived from the above MINLP for the leader (the insurer) and follower (the insurant). In the above MINLP, the objective (4.5) is to minimize the expected cost of the insurer given the distribution of the insurant's

initial health state s. The insurer's policy is stochastic and has to satisfy constraints (4.6) and (4.7) to ensure a probability distribution. Constraints (4.8) and (4.9) force the insurant's policy to be binary; that is, the insurant adopts exactly one action in each health state $s \in \mathscr{S}$. Constraint (4.10) is used to compute the insurant's best response $\phi(a_f|s)$ to the insurer's policy π in each state s, and Constraint (4.11) ensures the insurer's optimal expected cost given an insurant's deterministic best response. Let Z be a very large constant. Constraint (4.10) is composed of two inequalities where the first indicates that the value of the insurant's optimal expected cost $v_F(s)$ in state s is no larger than the expected cost over all possible actions a_f. The second inequality in Constraint (4.10) ensures that if an action a_f is chosen in state s, i.e., $\phi(a_f|s) = 1$, $v_F(s)$ exactly equals the insurant's expected cost in that state; on the other hand, if $\phi(a_f|s) = 0$, the constraint becomes $0 \leq R_F(s, \pi, a_f) - v_F(s) \leq Z$ and the right-hand side is just a redundant inequality. These two inequalities make sure that the insurant minimizes his/her expected cost over all possible choices a_f in each state s. Similarly, in Constraint (4.11), when the insurant chooses action a_f in state s, the constraint becomes equality, i.e., $v_L(s) = R_L(s, \pi, a_f)$, and thus the insurer's cost must equal the expected cost when the insurant takes action a_f in that state; on the other hand, when $\phi(a_f|s) = 0$, the constraint becomes redundant.

4.1.2 Markov stationary stochastic Stackelberg game solution

A Stackelberg equilibrium in this stochastic Stackelberg game exists if the leader's and follower's strategies are restricted to Markov stationary strategies [95, 154]. The equilibrium is the optimal solution found by solving the MINLP formulation (4.5)−(4.11). However, because of the combination of the binary variables ϕ and the non-convex interaction between continuous variables (i.e., π and v_F in Constraint (4.10), and π and v_L in Constraint (4.11)), the MINLP formulation is computationally challenging to solve using existing MINLP solution methods. The Stackelberg equilibrium computation in [154] is applied to solve the MINLP formulation: discretizing the probabilities π [66] and linearizing these constraints using McCormick inequalities [107] to approximate a Stackelberg equilibrium. The details of the approximation method and corresponding optimization problem are presented in Appendix E.

4.2 Numerical Results for Hypertension Case

The SSG cost-sharing mechanism design is applied to analyze copayment policies for a hypertension case. Nearly 90% of U.S. adults with uncontrolled hypertension have insurance and a source of health care [32], representing a missed opportunity to control hypertension because they were either unaware of their hypertension or they were aware of their hypertension but were not receiving treatment. [114] presented that over 64% of adults who were taking hypertension medication successfully achieved blood pressure less than 140/90 mmHg. However, nonadherence to prescribed medication and insufficiently intensive antihypertension treatment are the two major challenges to controlling high blood pressure [52]. Thus, an expanded effort to incent an increased rate of taking hypertension medication treatment using cost-sharing rates provides an opportunity to improve hypertension control.

Four health states (i.e., normal s_1, prehypertension s_2, hypertension stage 1 s_3, and hypertension stage 2 s_4) are defined in the experiments. The designation is intended to identify those individuals in whom early intervention by an adoption of a healthy lifestyle or adherence to treatment could reduce blood pressure, decrease the rate of deterioration of blood pressure to hypertension levels, or prevent hypertension entirely. Since high blood pressure values are associated with an increase in the risk of mortality as compared to those with normal blood pressure levels, early intervention could improve health outcomes.

Three experiments are conducted. To validate the insurant's response to cost-sharing in the hypertension case, the first experiment is conducted to examine the sensitivity of the insurant's behavior towards variations in the cost-sharing rate and income, and thus compare the mathematical model and quantitative analysis with literature reviews of evidence. The second experiment analyzes the insurer's and the insurant's strategies in a Stackelberg equilibrium to observe their interactions over different insurant's characteristics, i.e., age and income. The insurer's expected cost is also reported. The third experiment analyzes multi-tiered copayments over multiple treatments to show the analysis can be applied to multi-tiered cost-sharing levels. In the third experiment, two types of treatments, one with low cost and the other with higher cost but higher efficacy, are considered.

Table 4.1: Experiment settings

Experiment 1: Fixing insurer's action, find optimal insurant's action	
Varied inputs	insurant's age {20-39,40-59,60-79}
	insurant's income {\$210, \$625, \$2300}
	Copayment level $\{10, 40, 70\}$
insurant's actions	{Do nothing, Lifestyle adjustment, Treatment taking}
Experiment 2: Copayment policy for a single antihypertension treatment	
Varied inputs	insurant's age {20-39,40-59,60-79}
	insurant's income {\$210, \$625, \$2300}
Insurer's actions	Copayment level $\{10, 40, 70\}$ with probabilities 0, 0.5, and 1
insurant's actions	{Do nothing, Lifestyle adjustment, Treatment taking}
Experiment 3: Copayment policy for multiple antihypertension treatments	
Varied inputs	insurant's income {\$210, \$625, \$2300}
	Copayment for treatment A\B {10\10,10\90,45\110}
insurant's actions	{Do nothing, Lifestyle adjustment, Treatment A taking, Treatment B taking}

4.2.1 Parameter settings

The following inputs are varied in the numerical experiments: 1) insurant's age, 2) income in the normal state, and 3) copayment level. The experiment settings are summarized in Table 4.1. The collected outputs from the numerical experiments are as follows: 1) optimal policies for the insurer and insurant, and 2) minimum expected cost to the insurer.

The hypertension-specific parameters and corresponding sources are given in Table 4.2. The insurer's and the insurant's immediate cost functions for copayment plan are presented in Equations (4.1) and (4.2), respectively. The antihypertension treatment cost function $c^{treat}(s, a_f)$ and side-effect functions $c_L^{effect}(s, a_f)$ and $c_F^{effect}(s, a_f)$ are formulated based on the hypertension literature as shown in Table 4.2.

The monthly average antihypertension treatment cost per person $c^{treat}(s, a_f)$ (see Table 4.2) is calculated from [43] and [29]. [29] suggested that an additional monthly investment of \$50 is required for insurants with hypertension stage 2 and is reflected in Table 2. More details about hypertension therapy costs and recommendations can be found in cost-effective analysis studies [153, 111, 103].

The relative cost of the insurer's and insurant's side-effects in their immediate cost functions, $c_L^{effect}(s, a_f)$ and $c_F^{effect}(s, a_f)$, are based on mortality rate of coronary artery disease to factor in

Table 4.2: Hypertension parameter values used in the numerical experiments

Variable	Value	Source
Monthly cost $c^{treat}(s,a_f)$		
Lifestyle adjustments (L) in s_1 or s_2	\$60	Assume
Treatment (T) in s_3	\$80	[43]
Treatment (T) in s_4	\$130	[29]
Mortality $Mortality(s)$		
	(s_1, s_2, s_3, s_4)	
20-39 age group	(0.0001,0.0001,0.0002,0.0003)	Assume
40-59 age group	(0.00012,0.00019,0.00031,0.00038)	[119]
60-79 age group	(0.00047,0.00049,0.00062,0.00065)	[119]
Initial health state distribution $\beta(s)$		
	(s_1, s_2, s_3, s_4)	
20-39 age group	(0.652,0.297,0.044,0.007)	[83]
40-59 age group	(0.465,0.391,0.119,0.026)	[83]
60-79 age group	(0.408,0.413,0.212,0.067)	[83]
Discount rate		
γ_L	0.996	[90]
γ_F	0.996	[90]

increased risk of vascular disease mortality as blood pressure increases [119]. As shown in the study [28], antihypertension treatment reduced the risk of all-cause mortality with relative risk 0.89. Thus, c_L^{effect} and c_F^{effect} are multiplied by 0.89 when lifestyle adjustment or antihypertension treatment are taken. Note that the relative cost of the side-effect when taking antihypertension treatment is less than the side-effect with do-nothing action. Let $c_L^{effect}(s,a_f)$ and $c_F^{effect}(s,a_f)$ be functions of mortality rate $mortality(s)$ as follows,

$$c_L^{effect}(s,a_f) = c_F^{effect}(s,a_f) = \begin{cases} 3(mortality(s)\cdot 10^4)^2 & \text{if } a_f = D \\ 3(0.89\cdot mortality(s)\cdot 10^4)^2 & \text{if } a_f = L \text{ or } T \end{cases} \tag{4.12}$$

where $mortality(s)$ is shown in Table 4.2.

The insurant's benefit $b_F^{health}(s,a_f)$ is represented by person monthly income, which is based on the person monthly income distribution in [149] (\$210: percent below 10, \$625: percent below 20, \$2300: percent below 50); thus, the insurant's benefit is represented as

$$b_F^{health}(s, a_f) = income \cdot \lambda(s), \tag{4.13}$$

where *income* is an input to the experiments as shown in Table 4.1 and $\lambda(s)$ is loss of production rate in state s, $\lambda(s_1) = 1$, $\lambda(s_2) = 0.99$, $\lambda(s_3) = 0.98$, and $\lambda(s_4) = 0.95$.

The probabilities $\beta(s)$ that the initial health state is s are adopted from the distribution of blood pressure measurement in the USA [83] for different age groups. Similar to other hypertension research using Markov models to evaluate the cost-effectiveness of hypertension therapies [90], an annual discount rate of 5% is applied to all future costs for leader and follower (monthly $\gamma_L = \gamma_F = 0.996$).

The discrete-time Markov model, with monthly transition probabilities $T_{ss'}^{a_f}$ from hypertension state s to s', is built based on hypertension literature studies [103, 56, 147]. Also a regression analysis in these studies showed that age is a significant factor for developing hypertension; thus, the transition probabilities are age-dependent and younger age groups are assumed to have better antihypertension treatment efficacy and lower rates to develop high blood pressure. If the health states are extended by incorporating multiple congenital heart disease (CHD), stroke events, and death into health states, the mortality and risk of cardiovascular disease could be applied as in [103]. The following functions is used to convert these probabilities into monthly probabilities, $p_m = 1 - (1 - p_y)^{1/12}$, $p_m = 1 - (1 - p_{5y})^{1/60}$, and $p_m = 1 - (1 - p_{8y})^{1/96}$, where p_m, p_y, p_{5y}, and p_{8y} are monthly, yearly, five-yearly, and eight-yearly probabilities, respectively.

The transition probabilities from normal and prehypertension states to hypertension stage 1 are selected and transformed to monthly data from [147], which studied the rate of development of hypertension with and without treatment for different blood pressures by monitoring blood pressure over eight years. The transition probability from normal to prehypertension state is selected within $15 - 38\%$, which is the probability that blood pressure is raised over five years [103], and transformed to monthly data.

The monthly transition probabilities when the insurant does nothing are presented as follows,

$$
\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array}
\begin{array}{c}
\begin{array}{cccc} s_1 & s_2 & s_3 & s_4 \end{array} \\
\left[\begin{array}{cccc}
0.991-2t & 0.006+t & 0.003+t & 0 \\
0 & 0.992-2t & 0.008+2t & 0 \\
0 & 0 & 0.988-2t & 0.012+2t \\
0 & 0 & 0 & 1
\end{array}\right]
\end{array}
\tag{4.14}
$$

where s_1, s_2, s_3, and s_4 are normal, prehypertension, hypertension stage 1 and stage 2, respectively. Parameter t is used to create age-dependency in the monthly transition probability. Although the literature agrees that the transition probabilities depend on age, there is no agreement in the literature on how to quantify the age dependency, especially since it is often challenged because the data relied on self-measured blood pressure, as discussed in [156]. Let values for t be: -0.001, 0.001, and 0.003 for age groups 20-39, 40-59, and 60-79, respectively.

Similarly, the transition probabilities when the insurant adjusts lifestyle and\or takes antihypertension treatment are presented as follows,

$$
\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array}
\begin{array}{c}
\begin{array}{cccc} s_1 & s_2 & s_3 & s_4 \end{array} \\
\left[\begin{array}{cccc}
0.996-2t & 0.003+t & 0.001+t & 0 \\
0.082-t & 0.915-t & 0.003+2t & 0 \\
0 & 0.082-t & 0.915-t & 0.003+2t \\
0 & 0 & 0.082-t & 0.918+t
\end{array}\right]
\end{array}
\tag{4.15}
$$

.

The transition probabilities from prehypertension to the normal state, from hypertension stage 1 to prehypertension, and from hypertension stage 2 to stage 1 with action treatment or lifestyle adjustment are estimated by yearly 36% failure rate to achieve blood pressure control [56]. More hypertension statistical literature is reported in [32, 33].

4.2.2 Experiment 1: Fixing insurer's action, find optimal insurant's action

The first experiment is performed to determine how sensitive the insurant's behavior is to variations in the cost-sharing rate and income in normal state for three different age groups. This indicates which parameters play the most significant role in the determination of actions as well as the trend

of treatment adherence. Table 4.3 shows the behaviors of insurants in three age groups for different copayments and income levels. It is observed that the insurant always chooses to do nothing in s_1 and adjust lifestyle in s_2, as expected. For treatment taking behavior in the 20-39 age group, the insurant with the lowest income \$210 tends to take antihypertension treatment in s_3 and s_4 with copayment \$10 while choosing to do nothing when copayment level is \$70 or the insurant is in s_4; on the other hand, the insurant with incomes \$625 and \$2300 always takes antihypertension treatment in s_3 and s_4 no matter which copayment level is set. As shown in Table 4.3, the insurant's actions and insurer's expected cost are all the same for the \$625 and \$2300 income levels. The insurant's actions for the \$210 income levels that are different are highlighted in bold in Table 4.3. It indicates that a low-income insurant is more likely to be sensitive to cost-sharing rate than the counterpart with high income. The trend is consistent with the study in [50] which showed insurants with higher copayments are less likely to adhere to recommended treatment regimens than those with lower copayments.

It is also observed that the insurant in the 60-79 age group is more willing to take treatment than the insurant in the 20-39 and 40-59 age groups in the lowest copayment level because of a higher probability of getting worse blood pressure. In most cases, the insurer's expected cost decreases when the the copayment level increases; however, the trend is different for the lowest income level in the 20-39 and 40-59 age groups because of side effects resulting from doing nothing when copayment level is 70, which leads to a higher medication cost in the future.

4.2.3 Experiment 2: Copayment policy for single antihypertension treatment

The second experiment determines the optimal cost-sharing policy for the insurer and optimal action for the insurant in the Stackelberg equilibrium. The two-player SSG for hypertension used in this experiment is shown in Figure 4.1.

See Tables 4.4 and 4.5 for the optimal strategies of the insurant and the insurer, respectively, with three different income levels and three age groups. The optimal copayment policies found in Table 4.5 are always pure strategies even though a mixed strategy is allowed. For example, the insurer's strategy when the insurant is in s_3, in all cases, is a pure strategy of selecting a 70 copayment level, denoted in Table 4.5 as $10:0, 40:0, 70:1$. When solving the MINLP for other

Table 4.3: Experiment 1: Cost-sharing sensitivity over three copayment levels, three income levels, and three age groups. The insurant's actions in the normal state s_1 are always to do nothing (D), and in the prehypertension state s_2 to always adjust lifestyle (L), so they are not shown in the table.

insurant's optimal strategy (D,L,T)									
Income in normal state	210			625			2300		
Copayment	10	40	70	10	40	70	10	40	70
20-39 age group									
Stage 1 s_3	T	T	***D***	T	T	T	T	T	T
Stage 2 s_4	***D***	***D***	***D***	T	T	T	T	T	T
Insurer's expected cost	**2223**	**1527**	**1990**	2292	1553	815	2292	1553	815
40-59 age group									
Stage 1 s_3	T	T	***D***	T	T	T	T	T	T
Stage 2 s_4	T	***D***	***D***	T	T	T	T	T	T
Insurer's expected cost	**5358**	**3541**	**5053**	5358	3630	1903	5358	3630	1903
60-79 age group									
Stage 1 s_3	T	T	T	T	T	T	T	T	T
Stage 2 s_4	T	T	***D***	T	T	T	T	T	T
Insurer's expected cost	21150	18220	16689	21150	18220	15290	21150	18220	15290

D: do nothing L: lifestyle adjustment T: antihypertension treatment

parameter settings, occasionally a mixed strategy was found, such as $10:0$, $40:0.5$, and $70:0.5$, which indicates that the insurer sets the copayment level to \$10, \$40, and \$70 with probabilities 0, 0.5, and 0.5, respectively. In Experiment 2, when the highest copayment level is selected, there is no need for a mixed strategy since it is already at the highest level. Note that the insurer's actions in s_1 and s_2 are not shown in Table 4.5 because the insurer's copayment levels in normal (s_1) and prehypertension (s_2) states will not affect the insurers' expected cost or the actions of the insurant (because the actions of the insurant "do nothing (D)" and "lifestyle adjustment (L)" do not include antihypertension treatment cost).

It is observed that the insurant with medium and high income (income= \$625 and \$2300) desire to take antihypertension treatment even with the highest copayment level setting, where as the insurant with low income (income= \$210) still chooses to not take antihypertension treatment in s_4 even when the insurer decreases the copayment level to 10, i.e., 10 copayment for the 20-39 and 40-59 age groups (highlighted in bold). The reason is because the insurant gains more incentive to

return to the normal state s_1 even though the insurant needs to pay a higher partial healthcare cost for the antihypertension treatment. However, the insurant with low income= \$210 in the 60-79 age group tends to take treatment with copayment level 40 (highlighted in bold) because of a higher probability of raising blood pressure.

The last row of Table 4.5 shows the insurer's minimum expected cost. For each age group, the insurer's minimum cost decreases as the income increases. The result indicates that the insurer should pay more attention to the insurants with low income because of poor adherence treatment-taking behavior.

Furthermore, it is worth noting that the insurer's expected cost improves compared to Experiment 1, which fixes the cost-sharing rate instead of solving for the Stackelberg equilibrium. For example, Table 4.3 shows that for the insurant with income \$210 in the 20-39 and 40-59 age groups, the minimum expected cost for the insurer is 1527 and 3541, respectively, with insurant's optimal response $s_3 : T$, and $s_4 : D$; in contrast, Experiment 2 shows that the insurer's minimum expected cost for the income level 210 is 831 and 2165 with the same insurant's optimal response $s_3 : T$, and $s_4 : D$. This improvement is due to a change in the copayment policy. Moreover, for the insurant with low income 210 in the 60-79 age group in Table 4.3, when the insurer adopts a copayment level \$70, the insurant takes treatment only in hypertension stage 1 (s_3) and does nothing in s_4, and the insurer's expected cost is 16689; in contrast, Table 4.5 shows that for the same insurant, when the copayment levels in the Stackelberg equilibrium are set as \$70 and \$40 in stages 3 and 4, respectively, the insurant takes antihypertension treatment in both hypertension stages and the insurer acquires a reduced expected cost of \$15668.

4.2.4 Experiment 3: Copayment policy for multiple antihypertension treatments

This experiment demonstrates that the analysis can be extended to multiple treatments with different cost-sharing rates. In addition to improving the adherence behavior of the insurant, a customized cost-sharing policy has an impact on the insurant's chronic disease progression, and the insurer can incent an insurant with specific characteristics to receive suitable treatment. Moreover, the different cost-sharing policies over multiple treatments can reduce the insurer's expenses while keeping insurants continuing to take treatments.

Table 4.4: Experiment 2: insurant's strategy ϕ_π^* in the Stackelberg equilibrium with respect to different income levels and age groups

insurant's strategy ϕ_π^*									
Age	20-39			40-59			60-79		
Income in normal state	210	625	2300	210	625	2300	210	625	2300
Normal s_1	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*	*D*
Prehypertension s_2	*L*	*L*	*L*	*L*	*L*	*L*	*L*	*L*	*L*
Stage 1 s_3	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*	*T*
Stage 2 s_4	***D***	*T*	*T*	***D***	*T*	*T*	***T***	*T*	*T*

D: do nothing *L*: lifestyle adjustment *T*: antihypertension treatment

Table 4.5: Experiment 2: Insurer's strategy π^* in the Stackelberg equilibrium with respect to different income levels and age groups

Insurer's strategy π^* (level of copayment: probability that adopts that level)									
Age	20-39			40-59			60-79		
Income in normal state	210	625	2300	210	625	2300	210	625	2300
Stage 1 s_3	10: 0	10: 0	10: 0	10: 0	10: 0	10: 0	10: 0	10: 0	10: 0
	40: 0	40: 0	40: 0	40: 0	40: 0	40: 0	40: 0	40: 0	40: 0
	70: 1	70: 1	70: 1	70: 1	70: 1	70: 1	70: 1	70: 1	70: 1
Stage 2 s_4	10: **1**	10: 0	10: 0	10: **1**	10: 0	10: 0	10: **0**	10: 0	10: 0
	40: **0**	40: 0	40: 0	40: **0**	40: 0	40: 0	40: **1**	40: 0	40: 0
	70: **0**	70: 1	70: 1	70: **0**	70: 1	70: 1	70: **0**	70: 1	70: 1
Insurer's minimum expected cost	831	815	815	2165	1903	1903	15668	15290	15290

Multi-tiered cost-sharing benefits are increasingly being implemented by plan sponsors, especially targeting prescription drug plans. The impact of multi-tiered pharmacy benefits on drug utilization has been the subject of several studies. In common three-tier formulary, the copayment for generic drugs in the first tier is the lowest, a copayment for brand-name drugs preferred by the insurer in the second tier is higher, and the copayment for brand-name drugs not preferred by the insurer in the third tier is the highest. By examining a three-tier measure of insurance coverage on different index depth of drug converge for hypertension, [109] found that lack of drug coverage was associated with cost-related poor adherence, and cost-related poor adherence was associated with income level and out-of-pocket spending for drugs. [144] showed that overall compliance for antihypertensive agents was higher in tier 1 than in tiers 2 and 3. [82] showed that increased copayments or adding a new level of copayments reduced plan expenditures for working aged people enrolled in employer health plans. The reduction mainly benefited health insurance plans because of increased beneficiary out-of-pocket costs. [132] discussed that prescription decision-making depends on communication between the pharmacists and the insurant's preferences due to physicians' insufficient knowledge to assist insurants in managing out-of-pocket costs for prescription drugs. Further, [131] analyzed pharmacy claims from a large three-tier pharmacy benefit plans for chronic medications and found that prescribing generic or preferred medications is associated with improvements in adherence to therapy. Therefore, effective management of insurants' healthcare prescription decisions under multi-tiered benefit plans is required.

Experiment 3 demonstrates that the framework and analysis can be extended to a multi-tier cost-sharing benefit plan. Although there are other analytical works considering two treatments [53, 70], to the best of our knowledge, this work is the first to examine the issues of long-term chronic disease progression. For example, Lisinopril and Enalapril (Vasotec) are both used to treat high blood pressure (hypertension) in adults; however, a 10 mg tablet of Enalapril costs $96, while a 40 mg tablet of generic Lisinopril costs $17 but has more common side effects, e.g., cloudy urine and decrease in urine output.

The third experiment examines a two-tier formulary that contains two types of treatments which have different effectiveness and treatment costs. To reduce medication cost, the insurer tends to increase the copayment for the expensive treatment to drive the insurants to take the cheaper drug even though it may be less effective or require a longer recovery time. Two antihypertension treat-

ments can be chosen by the insurant in hypertension stage 1 (s_3) and stage 2 (s_4). The price of antihypertension treatment A is low ($100) and the price of antihypertension treatment B is high ($160) but more effective. The transition probability with action taking treatment A is the same as Equation (4.15) and the transition probability when taking the more effective treatment B is assumed as follows,

$$\begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{array} \begin{array}{c} \begin{array}{cccc} s_1 & s_2 & s_3 & s_4 \end{array} \\ \begin{bmatrix} 0.996-2t & 0.003+t & 0.001+t & 0 \\ 0.081-t & 0.916-t & 0.003+2t & 0 \\ 0 & 0.182-t & 0.816-t & 0.002+2t \\ 0 & 0 & 0.161-t & 0.839+t \end{bmatrix} \end{array} \tag{4.16}$$

which compared to treatment A, has a higher probability to lower blood pressure.

Three policies of two-tier copayment levels for antihypertension treatments A\B (10\10, 10\90, and 45\110) are examined. Table 4.6 shows that, for three age groups, when the copayment for the treatments A and B are both low (policy 1), all insurants take treatment B due to the better effectiveness of treatment B; however, as the copayment for treatment B increases (policy 2), all insurants change policy from treatment B to treatment A in both s_3 and s_4 states. Policy 3, $45 for treatment A and $110 for treatment B, keeps both insurants continuing to receive treatments, and also has the lowest sum of expected costs over income levels (e.g., 20-30 age group: $6945 for policy 1, $8541 for policy 2, and $4338 for policy 3). Note that, with Policy 3, the insurant with income $210 in the 20-39 age group tends to take treatment A in s_3 while the insurants with income $210 in the 40-59 and 60-79 age groups choose to take treatment B (highlighted in bold). The analysis manifests that the framework allows the insurer to lower medication cost by incenting the insurant to continue taking hypertensive treatments with a multi-tier formulary.

4.3 Structural Properties of the Model

The game-theoretic approach enables an insurer to derive an optimal cost-sharing policy with respect to insurant's corresponding health states and response. However, a deeper insight into the game model is necessary in order to understand both insurer's and insurant's dynamic behaviors.

Table 4.6: Experiment 3: insurant's optimal strategies over three two-tier copayment policies and three age groups. The insurant's actions in the normal state s_1 are always to do nothing (D), and in the prehypertension state s_2 to always adjust lifestyle (L), so they are not shown in the table.

Insurer's copayment policies									
Two-tier policy	Policy 1			Policy 2			Policy 3		
Treatment A\B	10\10			10\90			45\110		
insurant's optimal strategy (D, L, T_A, T_B)									
Income	210	625	2300	210	625	2300	210	625	2300
20-39 age group									
Stage 1 s_3	T_B	T_B	T_B	T_A	T_A	T_A	T_A	T_B	T_B
Stage 2 s_4	T_B	T_B	T_B	T_A	T_A	T_A	T_A	T_A	T_B
Insurer's expected cost	2315	2315	2315	2847	2847	2847	1982	1182	1174
Sum over income levels	6945			8541			4338		
40-59 age group									
Stage 1 s_3	T_B	T_B	T_B	T_A	T_A	T_A	T_B	T_B	T_B
Stage 2 s_4	T_B	T_B	T_B	T_A	T_A	T_A	T_A	T_A	T_B
Insurer's expected cost	4623	4623	4623	5880	5880	5880	2305	2305	2256
Sum over income levels	13869			17640			6866		
60-79 age group									
Stage 1 s_3	T_B	T_B	T_B	T_A	T_A	T_A	T_B	T_B	T_B
Stage 2 s_4	T_B	T_B	T_B	T_A	T_A	T_A	T_A	T_A	T_B
Insurer's expected cost	18903	18903	18903	21080	21080	21080	15300	15300	15177
Sum over income levels	56709			63240			45777		

D: do nothing T_A: antihypertension treatment A T_B: antihypertension treatment B

The next step is to consider structural properties of the proposed optimal health insurance framework-cost-sharing problem. This may help guarantee certain structure on the model output (i.e., the optimal values and policies in Stackelberg equilibrium) with certain structure on the model input. Besides, structural properties can also make policy implementation easier and accelerate computation time for large-scale game models.

Understanding structural properties of the stochastic Stackelberg game may be useful to derive a set of conditions that ensure the existence of monotonic optimal policies or control-limit policy for both players (e.g., should insurance company offer high or low cost-sharing rate? at which health state should patients begin to take treatment?). It may be possible to use structural properties to solve the game problem within a limited computation time.

Research on optimal time to initiating therapy or acceptance of organs is also growing [2, 3, 130]. In this section, structural properties of discounted infinite-horizon Stackelberg stochastic game are studied to derive conditions that ensure the Stackelberg equilibrium of stochastic game involves monotonic cost-sharing policy and monotonic chronic disease-based policy. Given monotonic cost-sharing policy, the conditions that insurant has monotonic chronic disease-based policy is first presented. Second, conditions that monotonic cost-sharing policy for the insurer can guarantee the minimum cost is then presented. For convenience, insurant's actions are indexed $D_T = 1$ (don't take treatment), $T = 2$ (take treatment), $L = 3$ (change lifestyle), and $D_L = 4$ (not change lifestyle). Also, insurant's health states are indexed normal $= 1$, prehypertension $= 2$, stage $1 = 3$, and stage $2 = 4$.

General-sum discounted stochastic games are proved to have at least one Nash equilibrium point in stationary strategies [59]. Approachability is adopted to analyze the multi-objective optimization in finite-time Stackelberg Stochastic Games [84]. In this problem, the Stackelberg equilibrium is presented as a solution to a non-linear integer program.

Definition 1. ***Monotonic chronic disease-based policy***

For a given insurer's cost-sharing policy $\pi(a_l|s)$, if the optimal action the insurant selects, $a_f(s)$, is nonincreasing in insurant's health state s, the optimal policy of the insurant $\phi(a_f|s)$ is called a monotonic chronic disease-based policy.

Definition 2. ***Monotonic cost-sharing policy***

For a given response of an insurant $\phi(a_f|s)$, if the optimal action the insurer selects, $a_l(s)$, is nonincreasing in insurant's health state s, the optimal policy of the insurer $\pi(a_l|s)$ is called a monotonic cost-sharing policy.

Assumption 1. *For a given monotonic cost-sharing policy, $\pi(a_l|s)$, if $s^+ \geq s^-$, then the immediate cost function for the insurant satisfies*

$$r_F(s^+, a_l(s^+), a_f) \geq r_F(s^-, a_l(s^-), a_f) \tag{4.17}$$

for each $a_f \in \mathscr{A}_F$.

For example, assumption 1 indicates that with the same insurant's action, the immediate cost of an insurant with high blood pressure is higher than that of an insurant with low blood pressure.

Assumption 2. *For a given monotonic response of an insurant, $\phi^*(\pi)$, if $s^+ \geq s^-$, then the immediate cost function for the insurer satisfies*

$$r_L(s^+, a_l, a_f^*(s^+)) \geq r_L(s^-, a_l, a_f^*(s^-)) \tag{4.18}$$

for each $a_l \in \mathscr{A}_L$.

An example of assumption 2 is that, given the same cost-sharing policy, the immediate cost for the insurer is higher when the insurant has high blood pressure than when the insurant has low blood pressure.

Assumption 3. $\overline{T}_{sk}^{a_f} = \sum_{s'=k}^{|S|} T_{ss'}^{a_f}$ *is nondecreasing in s for each $k \in \mathscr{S}$, $a_f \in \mathscr{A}_F$.*

Similar to the definition of Increasing Failure Rate (IFR) [130], Assumption 3 implies that given an insurant's action, the insurant with higher blood pressure has a larger probability to go to any particular health state worse than lower blood pressure.

Assumption 4. $\overline{T}_{sk}^{a_f} = \sum_{s'=k}^{|S|} T_{ss'}^{a_f}$ *is submodular in (s, a_f) for each $k \in \mathscr{S}$.*

From the Assumption 4,

$$\overline{T}^{L}_{normal\,k} + \overline{T}^{D_L}_{prehypertension\,k} \geq \overline{T}^{D_L}_{normal\,k} + \overline{T}^{L}_{prehypertension\,k} \tag{4.19}$$

and

$$\overline{T}^{D_T}_{stage1\,k} + \overline{T}^{T}_{stage2\,k} \geq \overline{T}^{T}_{stage1\,k} + \overline{T}^{D_T}_{stage2\,k} \tag{4.20}$$

for each $k \in \mathscr{S}$, implying that the insurant with *prehypertension* should start to adjust lifestyle; similarly, as insurant with hypertension should seek early treatment since it will result in a better health state than late treatment.

Theorem 3. *For a given monotonic cost-sharing policy, $\pi(a_l|s)$, suppose assumptions 1, 3 and 4 hold, and*

$$r_F(s^-, a_l(s^-), a_f^+) + r_F(s^+, a_l(s^+), a_f^-) \geq r_F(s^+, a_l(s^+), a_f^+) + r_F(s^-, a_l(s^-), a_f^-) \tag{4.21}$$

if $s^+ \geq s^-$ and $a_f^+ \geq a_f^-$. Then $a_f(s^-) \geq a_f(s^+)$ if $s^+ \geq s^-$, that is, a monotonic chronic disease-based policy exists.

Proof. The following proof follows the logic in the proof of monotonicity of optimal policies [68].

Claim 1. *$v_F(s)$ is nondecreasing in s.*

Proof. Let $s^+ \geq s^-$ and suppose a_f^* is optimal in s^+.

$$\begin{aligned}
v_F(s^+) &= \min_{a_f \in A_F} \left\{ r_F(s^+, a_l^*(s^+), a_f) + \sum_{s' \in S} T_{s^+s'}^{a_f} v_F(s') \right\} \\
&= r_F(s^+, a_l^*((s^+)), a_f^*) + \sum_{s' \in S} T_{s^+s'}^{a_f^*} v_F(s') \\
&\geq r_F(s^-, a_l^*(s^-), a_f^*) + \sum_{s' \in S} T_{s^+s'}^{a_f^*} v_F(s') \\
&\geq r_F(s^-, a_l^*(s^-), a_f^*) + \sum_{s' \in S} T_{s^-s'}^{a_f^*} v_F(s') \\
&\geq \min_{a_f \in A_F} \left\{ r_F(s^-, a_l^*(s^-), a_f) + \sum_{s' \in S} T_{s^-s'}^{a_f} v_F(s') \right\} \\
&= v_F(s^-)
\end{aligned}$$

This proves claim 1. □

Claim 2. *$\sum_{s' \in S} T_{ss'}^{a_f} v_F(s')$ is submodular in (s, a_f).*

Proof. Submodularity of $\overline{T}_{ss'}^{a_f}$ implies

$$\overline{T}_{s^+k}^{a_f^+} + \overline{T}_{s^-k}^{a_f^-} \leq \overline{T}_{s^+k}^{a_f^-} + \overline{T}_{s^-k}^{a_f^+}, \forall s^+ \geq s^-, a^+ \geq a^-, \text{for each } k \in S.$$

That is, for each $k \in S$,

$$\begin{aligned}
\sum_{s' \geq k} \left[T^{a_f^+}_{s^+ s'} + T^{a_f^-}_{s^- s'} \right] &\leq \sum_{s' \geq k} \left[T^{a_f^-}_{s^+ s'} + T^{a_f^+}_{s^- s'} \right] \\
\sum_{s' \geq k} \left[T^{a_f^+}_{s^+ s'} + T^{a_f^-}_{s^- s'} \right] v_F(s') &\leq \sum_{s' \geq k} \left[T^{a_f^-}_{s^+ s'} + T^{a_f^+}_{s^- s'} \right] v_F(s') \\
\sum_{s' \geq k} T^{a_f^+}_{s^+ s'} v_F(s') - \sum_{s' \geq k} T^{a_f^-}_{s^+ s'} v_F(s') &\leq \sum_{s' \geq k} T^{a_f^+}_{s^- s'} v_F(s') - \sum_{s' \geq k} T^{a_f^-}_{s^- s'} v_F(s').
\end{aligned}$$

This proves claim 2. □

Now, $R_F(s, a_l(s), a_f) = r_F(s, a_l(s), a_f) + \gamma_F \sum_{s \in S} T^{a_f}_{ss'} v_F(s')$ is submodular because $r_F(s, a_f(s), a_l)$ is submodular (Condition (4.21)) and $\sum_{s' \in S} T^{a_f}_{ss'} v_F(s')$ is also submodular (Claim 2). The conclusion of Theorem 3 thus holds because $a_f(s)$ is the largest minimizer of $R_F(s, a_l(s), a_f)$. □

Lemma 4. *Suppose $R_F(s, a_l(s), a_f)$ is submodular given $a_l(s)$. Let $a_f^*(s)$ denote the largest minimizer of $R_F(s, a_l(s), a_f)$, then $a_f^*(s)$ is nonincreasing in s.*

Proof. Suppose not. Then, there exists $s^+ \geq s^-$ such that $a_f^*(s^+) > a_f^*(s^-)$. Thus, $a^*(s^+)$ is not optimal in s^-.

$$0 > R_F(s^-, a_l(s^-), a_f(s^-)) - R_F(s^-, a_l(s^-), a_f(s^+)). \tag{4.22}$$

Since $a_f(s^+)$ is optimal in s^+,

$$R_F(s^+, a_l(s^+), a_f(s^-)) - R_F(s^+, a_l(s^+), a_f(s^+)) \geq 0. \tag{4.23}$$

Combining the two inequalities,

$$R_F(s^+, a_l(s^+), a_f(s^-)) + R_F(s^-, a_l(s^-), a_f(s^+)) > R_F(s^+, a_l(s^+), a_f(s^+)) + R_F(s^-, a_l(s^-), a_f(s^-)). \tag{4.24}$$

This contradicts submodularity. □

Condition (4.21) may be violated for some (s, a_f). The magnitude of the violation of Equation

(4.21) is quantified by defining the following metric,

$$\max_{a_f^+ \geq a_f^-, s^+ \geq s^-} \left\{0, \left[r_F(s^-, a_l(s^-), a_f^+) + r_F(s^+, a_l(s^+), a_f^-)\right] - \left[r_F(s^+, a_l(s^+), a_f^+) + r_F(s^-, a_l(s^-), a_f^-)\right]\right\}. \tag{4.25}$$

Theorem 5. *For a given optimal monotonic response of an insurant $\phi^*(\pi)$, suppose assumptions 2, 3, and 4 hold, and*

$$r_L(s^-, a_l^+, a_f^*(s^-)) + r_L(s^+, a_l^-, a_f^*(s^+)) \geq r_L(s^+, a_l^+, a_f^*(s^+)) + r_L(s^-, a_l^-, a_f^*(s^-)) \tag{4.26}$$

if $s^+ \geq s^-$ and $a_l^+ \geq a_l^-$. Then $a_l(s^-) \geq a_l(s^+)$ if $s^+ \geq s^-$, that is, monotonic cost-sharing policy exists.

The proof is similar to the proof of Theorem 3. Similar to Condition (4.21), Condition (4.26) may also be violated for some (s, a_l). The magnitude of the violation of Condition (4.21) is also quantified by defining the following metric

$$\max_{a_l^+ \geq a_l^-, s^+ \geq s^-} \left\{0, \left[r_L(s^-, a_l^+, a_f^*(s^-)) + r_L(s^+, a_l^-, a_f^*(s^+))\right] - \left[r_L(s^+, a_l^+, a_f^*(s^+)) + r_L(s^-, a_l^-, a_f^*(s^-))\right]\right\}. \tag{4.27}$$

The space of health states can be expanded to have more accurate chronic disease progression modeling, for example, {normal, prehypertension, stage 1 without history of non-fatal CHD or stroke, stage 1 survived a CHD event or stroke, stage 1 with history non-fatal CHD or stroke but no CHD event or stroke in this time period, stage 2 without history of non-fatal CHD or stroke, stage 2 survived a CHD event or stroke, stage 2 with history non-fatal CHD or stroke but no CHD event or stroke in this time period} [90, 103, 126],. In this case, structural properties can be used to accelerate computation time for this large-scale game models and make policy implementation easier.

4.4 Conclusions

In this chapter, a stochastic Stackelberg game is applied to design an optimal cost-sharing mechanism in healthcare insurance. The game-theoretic mechanism connects dynamic health outcome,

cost-sharing policy, and medical treatment for chronic disease by mathematical formulation, coupled with the long-term interaction between the insurer and the insurant with disease progression. The framework is applied to a case of hypertension medication use and blood pressure control. The results complement related evidence in the related study review, and pave a way to an quantitative connection between insurance design and personal medical treatment adherence. In the experiment, the copayment policy is adopted to examine the antihypertensive treatment utilization for insurants with different ages and incomes. The framework could also extend to the multi-tier copayment plan. Moreover, the proposed framework has the potential to provide a personalized health insurance design for targeted insurants and customized health insurance which varies across insurant groups (age, income, and medical history). For future research, the SSG applied in this chapter can also be extended to a scenario where the insurer has incomplete information on the insurant, e.g., race, medical record, and health information. Different types of insurants cause different transition probabilities and immediate cost; this scenario can be solved by transforming the insurer's incomplete information regarding the insurant into an imperfect information game. That is, the problem can be mapped to a Bayesian Stackelberg game by considering the probability of facing different types of homogeneous insurant in the SSG.

It is worth highlighting two limitations of this study. First, as is common in the game-theoretic approach in the literature, the insurer (leader) in the SSG needs to know all the information and knowledge about the insurant (follower) in order to make a decision in the first move. This information includes transition probabilities, i.e., how would health status (e.g., blood pressure) evolve after taking different medical treatment, and the immediate cost or health benefit after taking actions in the specific health state. Second, the SSG developed in this study is a stationary Markov process and does not take the premium and deductible out-of-pocket costs in the health insurance into consideration. Both insurance premiums and deductible out-of-pocket costs are non-stationary time series which are based on historical information, i.e., past insurer and insurants' behaviors. For example, the insurer may increase the premium if claims were made by insurants in previous periods. However, typically, when insurants think of health insurance expenditures, they think about their monthly premium. The truth of the matter is that the majority of health care spending is on health care service utilization and not monthly premiums. It is therefore adequate that the cost-sharing analysis alone can be used to evaluate the effect of healthcare cost.

Affordable care act

The Affordable Care Act (ACA) precludes insurance companies from denying coverage on the basis of any criteria including pre-existing conditions (PECs), what is referred to in the insurance industry as guaranteed issue. In the pre-ACA era, insurance companies could and did exclude individuals based on PECs. Another element of the ACA limits the identifiable sources of risk that could be used to set higher premiums.

In the pre-ACA world, higher premiums or less generous benefits were used to dissuade individuals with predictably higher expenses based on PECs or socio-demographic risk. With that background, insurance companies could offer plans specifically designed for persons with chronic disease, and there are a variety of ways in which this occurs depending on the market segment (e.g., Medicare, Medicaid, commercial, individual and family) and the specific payment model in place. First, within markets in which provider payments are made in a capitated model, insurance premiums are risk-adjusted so that higher payments accrue to plans and providers that enroll individuals that are expected to be more expensive than the average enrollee. If risk-adjusted payments are sufficiently high, health plans will seek individuals with illnesses for which proper management using evidence-based tools are likely to decrease morbidity and therefore expense.

Hypertension applied in this chapter is the perfect example of this because it is a potentially very expensive disease if not well managed. However, evidence-based tools exist and health plans successfully reduce the population prevalence and risk from hypertension. One prominent example of this is the ways in which the Kaiser Permanente health plan in Northern California has demonstrably reduced the burden of hypertension in its 4 million plus member plan and thus earns positive net revenue because they are paid an above average fee to care for a population for whom they can reduce costs to closer to the population average. A successful model like this leads to Centers of Excellence in which purchasers seek out plans that have successfully reduced the burden of illness for a specific illness or population. For example, one hospital in the U.S. market has the national contract for spine surgeries for Walmart, which sends all of its complex spine cases nationally to this particular hospital.

Policy implications

The analysis on hypertension verifies that increasing insurant cost-sharing is associated with declines in treatment adherence, which in turn is associated with poorer health outcomes. Further, the analyses reflect the literature data that those with low income have a decrease in the consumption of medications and thus poorer treatment adherence.

The study suggests that viable health insurance policies, e.g., implement drug reimbursement or dispense more effective and essential medications should be targeted, towards a low-income population, which is similar to the goal of the ACA. Among the ACA's many objectives was an effort to create a robust market for individual and family insurance products, which had been absent for some years prior to the ACA's passage. Through subsidies, the exchanges were intended to incent both lower income and higher risk individuals to buy insurance with penalties used to incent lower risk individuals to also buy insurance. This work complements the efforts of insurers offering plans both motivated by the ACA as well as market reforms to create and offer insurance products that can both attract individuals with different risk profiles as well as use cost sharing to motivate health seeking behavior that increases health outcomes as well as reduces the rate of growth of health care spending.

This study also suggested an analysis for customized health insurance policies. Several group health insurance plans and cost-sharing discussions take a one-size-fits-all approach [22], although the coverage does not fit all insurants needs and is not able to work for certain situations and populations. For example, young adults should be offered a high deductible Health Savings Account (HSA) plan so that they can save thousands of dollars in an account for future medical expenses. Customized health insurance designs have emerged as a potential approach to improve healthcare value. Tiered pharmacy cost shares, for example, different copayments for brand or generic drugs or for drugs on and off formulary, were a first attempt to steer insurants to less costly medications. This experience has been greatly broadened with the advent of so called Value Based Insurance Design (VBID) in which cost shares increasingly vary by insurant and service based on a wide range of insurant factors (risk measures such as smoking status, BMI, blood pressure and whether insurants seek treatment for these needs) and service factors (the degree of evidence that supports a therapy and whether the service has a low value to cost ratio) [38, 143, 145]. Instead of consider-

ing on-going insurance issues such as fee-for-service and moral hazard reduction, health insurance application is able to have an impact on customized health insurance for specific market structure which captures insurant factors and service factors, e.g., low cost-sharing rate for prevention therapy for insurants with high risk for cardiovascular disease because of costly treatments and drugs. Therefore, notwithstanding limitations mentioned above, the paper can pave the way for more research in creating win-win situations across health insurance providers and the insured population with chronic diseases.

Chapter 5

OPTIMAL FLU VACCINATION REIMBURSEMENT POLICY

This chapter describes the third healthcare application: *Design an optimal vaccination reimbursement and cost-sharing policy for reducing medical treatment cost and preventing the spread of seasonal influenza*. An integrated healthcare insurance mechanism, which includes two interventions: vaccination incentive and cost-sharing policies, is developed. To design the mechanism, the dynamic interaction between a single insurer and multiple insurants is modeled as a Stackelberg vaccination game; the game then embedded into an agent-based simulation to model the spread of flu in a population under different insurance policies; finally, Machine learning and simulation optimization approaches are applied to optimize healthcare interventions in the large-scale flu transmission simulation. Simulation results indicate that the proposed methodology efficiently identifies good healthcare policies under different scenarios of flu vaccine efficacy and flu attack rates.

Most of this chapter appeared in [78] and [79].

5.1 Methodology: Incentive Mechanism

This section focuses on incentive mechanism design. The goal of the mechanism is to design an optimal incentive policy in insurance-based setting and find the optimal policies for minimizing the insurer's overall flu-related medical cost and public health burden. Section 5.1.1 builds an agent-based simulation to model the flu transmission. It captures the dynamic population health evolution and the effect of healthcare insurance policies during the spread of seasonal flu. Section 5.1.2 formulates the insurants and the insurer's decision-making problems in the agent-based simulation. Section 5.1.3 demonstrates how to address the computational issue of optimizing a large simulated population network by applying machine learning and simulation optimization techniques.

The proposed mechanism not only estimates the burden of influenza in the heterogeneous population and the impact of influenza vaccination reward and treatment cost-sharing but provides an incentive-optimization solution for effective and efficient healthcare policy management as well.

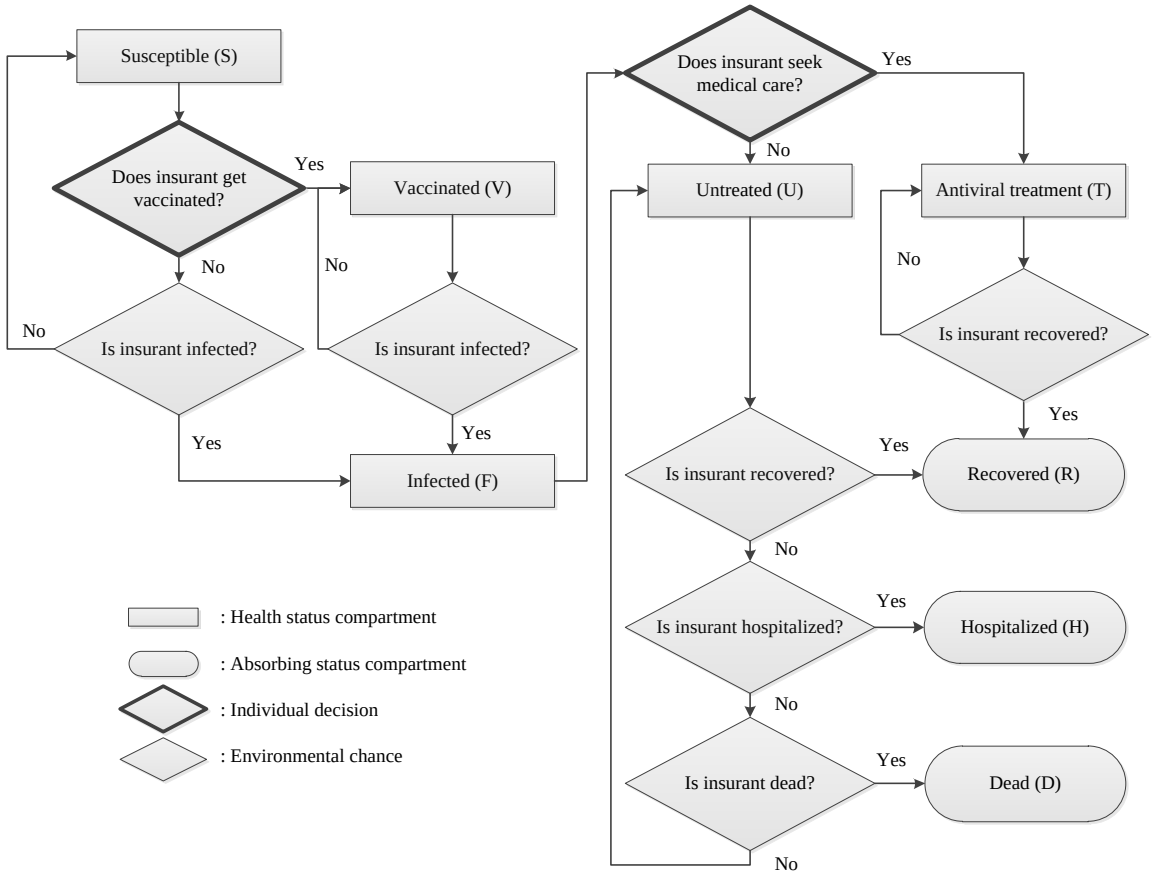


Figure 5.1: Flow diagram for the infection rules and individual decision-making.

5.1.1 Flu Transmission Agent-based Simulation

The agent-based simulation simulates the spread of seasonal influenza with a SIR structure ([86]), incorporating vaccination and corresponding rules that govern the transmission of flu, as depicted in the flow diagram in Figure 5.1. Each individual insurant is represented as an agent in the model. Agents are grouped into eight health state compartments, three of which are absorbing state compartments. The first compartment consists of susceptible (S) agents, who are subject to be infected. The second compartment consists of vaccinated agents (V), who receive flu vaccines. The third compartment consists of infectious agents (F), who are contagious. The fourth compartment consists of untreated agents (U), who are infected with the flu but decide not to receive outpatient healthcare services. The fifth compartment consists of treated agents (T), who receive outpatient healthcare

services after becoming infected. After the infection period, the treated and untreated agents may recover from flu and enter the recovery compartment (R). The seventh compartment consists of hospitalized agents (H), who are being hospitalized for the flu. The eighth compartment consists of agents that die (D) due to flu-related illness. The recovery (R), hospitalized (H) and dead (D) compartments are the three absorbing compartments; the insurants will stay in these compartments until the end of the flu season. .

In the flow diagram in Figure 5.1, agents in the susceptible compartment S may acquire the infection with a given infection probability and move into the infected compartment F. Agents in F may infect their neighbors with a given probability. Agents in the untreated compartment U and the treatment compartment T will recover (move into R) after the infected period. It is assumed that recovered agents are resistant to the seasonal flu and will not return to susceptible compartment S and are resistant to seasonal flu. Untreated agents in U may be hospitalized (move into H) or die from flu-related illness (move into D) based on given hospitalization and death probabilities.

Although adopting a stochastic approach to traverse agents' health states using infection, recovery, hospitalization, and death probabilities, two health state transitions are determined by insurants themselves: the vaccination decision "Does insurant get vaccinated?" (traverse from the compartment S to the compartment V); and the treatment-taking decision "Does insurant seek medical care?" (traverse from the compartment F to the compartment T). The susceptible and infected agents calculate their expected costs associated with the insurance policy and decide whether to get vaccination and/or treatment based on their received expected cost.

5.1.2 Stackelberg Vaccination Game Model

This section derives the expected cost for insurants by modeling the dynamic interaction between the insurer and insurants as a Stackelberg vaccination game played by a single insurer (an insurance company or public payer) and multiple insurants. The insurer and insurants are modeled as leader and followers, respectively. The insurer moves first by announcing the incentive policies before the flu season, i.e., vaccination reward and coinsruance rate to insurants, and each individual insurant responds by calculating two expected costs. Expected cost of vaccination is used to determine whether to get a vaccine in each time step after the announcement. If a non-vaccinated insurant

is infected, the expected cost of treatment is used to determine whether or not to seek treatment. Incorporating the game theoretic approach into the agent-based simulation model allows individual agents to act in their own self-interest in response to the influenza evolution and neighboring environment. Coinsurance rate, which the insurants pay a percentage of a medical charge for the share of the costs of a health care service, is adopted as a form of cost-sharing policy in health insurance.

Vaccination and Antiviral Treatment Problem for the Insurant

The expected cost of the insurant depends on whether s/he eventually gets vaccinated or infected. Table 5.1 presents the notation used in the Stackelberg vaccination game. Even though an individual insurant makes a vaccination decision before the treatment-seeking decision, treatment-seeking problem is first presented since the expected cost from treatment decision should be factored into the vaccination decision. It is assumed the insurants make the vaccination decision in every period (i.e., if the insurant does not take vaccine this period, s/he faces the same decision in the next period). It is assumed the treatment-seeking is a one-time decision for infected insurants. Insurants choose whether or not to seek treatment based on their respective expected costs with and without medical care. The expected cost of an untreated insurant with age a and risk type r is

$$E_U^{a,r} = (1 - p_{H|F}^a - p_{D|F}^a)(c_{med,U}^a + c_{ind,U}^a) + p_{H|F}^a(f_H c_{med,H}^{a,r} + c_{ind,H}^{a,r}) + p_{D|F}^a(f_D c_{med,D}^{a,r}), \quad (5.1)$$

where $c_{med,U}^a + c_{ind,U}^a$, $f_H c_{med,H}^{a,r} + c_{ind,H}^{a,r}$, and $f_D c_{med,D}^{a,r}$ are an untreated insurant's cost for no medical care, hospitalization, and death, respectively.

The expected cost of a treated insurant is

$$E_T^{a,r} = x_{sharing} c_{med,T}^{a,r}. \quad (5.2)$$

If $E_T^{a,r} < E_U^{a,r}$, then the insurant decides to seek medical care.

Insurants choose whether or not to vaccinate based on their respective expected costs with and without vaccination. If the insurant does not get vaccinated today, s/he can still receive a vaccination on the following time steps. The expected cost of an unvaccinated insurant with age a and risk type r

staying in compartment S is

$$E_S^{a,r} = p_{F|S}^a \min\left\{E_U^{a,r}, E_T^{a,r}\right\}, \tag{5.3}$$

where $p_{F|S}^a$ is the infection probability for a susceptible insurant, and $\min\left\{E_U^{a,r}, E_T^{a,r}\right\}$ is infected insurant' minimum expected cost, involving treatment-seeking behavior. At each time period (e.g., per day or per week), the infection probability $p_{F|S}^a$ changes according to the health state of an insurant's neighbors.

The expected cost of a vaccinated insurant is,

$$\begin{aligned} E_V^{a,r} &= (1 - p_{F|V}^a)(c_{ind,V} - x_{reward}) + p_{F|V}^a(c_{ind,V} - x_{reward} + \min\{E_U^{a,r}, E_T^{a,r}\}) \\ &= c_{ind,V} - x_{reward} + p_{F|V}^a \min\{E_U^{a,r}, E_T^{a,r}\} \end{aligned} \tag{5.4}$$

where $p_{F|V}^a$ is the infection probability for a vaccinated insurant, and x_{reward} is the monetary incentive given to vaccinated insurant. Since an insurant's immune response to flu vaccine is not perfect, the following infection cost, $\min\{E_U^{a,r}, E_T^{a,r}\}$, is added and $p_{F|V}^a$ is the corresponding infection probability for vaccinated insurant. Comparing to $E_S^{a,r}$ in (5.3), $E_V^{a,r}$ in (5.4) contains incentive for vaccination x_{reward} and indirect cost for vaccination $c_{ind,V}$, assuming the direct cost is free. If $E_V^{a,r} < E_S^{a,r}$, the insurant decides to receive a flu vaccine; otherwise, s/he may still receive a flu vaccine in the future using the same decision rule.

Incentive Policy Setting Problem for the Insurer

The insurer's goal is to minimize the total expected cost from both vaccination and flu infection while considering the social benefit, i.e., public health outcome of the population. Let ϕ_V and ϕ_F be the proportion of the population that is vaccinated and infected, respectively, with response to vaccination reward x_{reward} and coinsurance rate $x_{sharing}$. Let $\phi_{T|F}^{a,r}$, $\phi_{H|F}^{a,r}$, and $\phi_{D|F}^{a,r}$ be the proportion of influenza-attributable cases that lead to outpatient visits, hospitalization, and death, respectively. The insurer's objective as

$$\min_{x_{reward}, x_{sharing}} E_{Insurer}, \tag{5.5}$$

where

$$
\begin{aligned}
E_{Insurer} =& \phi_V x_{reward} \\
&+ \phi_F \sum_{a,r} \left[\phi_{T|F}^{a,r} \left(1 - x_{sharing}\right) c_{med,T}^{a,r} + \phi_{H|F}^{a,r} \left(1 - f_H\right) c_{med,H}^{a,r} + \phi_{D|F}^{a,r} \left(1 - f_D\right) c_{med,D}^{a,r} \right] \\
&+ w\phi_F,
\end{aligned}
$$

and w is the weight of social benefit that converts the proportion of infected population to a monetary unit. If the insurer is the public payer, the weight of social benefit should be higher because public healthcare insurance can act like a welfare subsidy. The first term of the equation is the average incentive cost per insurant, the second term is the average medical cost per insurant that includes outpatient visit cost, hospitalization cost, and death cost, and the last term is the cost of public health.

A coupling relationship exists between x_{reward} and $x_{sharing}$. A high x_{reward} will potentially increase the total vaccination reward cost, while a low x_{reward} may cause a lower vaccination rate, leading to a higher infection rate. On the other hand, increasing $x_{sharing}$ will discourage the infected insurants to seek treatment, leading to higher hospitalization and mortality cases, while decreasing $x_{sharing}$ increases the insurer's medical payment.

The Stackelberg vaccination game is Incorporated in the agent-based stimulation logic as follows. The simulation proceeds over iterations and one iteration can be seen as one decision-making period (e.g., daily or weekly). The behavior of insurants is modeled with two decisions, i.e., vaccination and medical treatment seeking dynamics. In the first decision, each insurant decides whether or not to vaccinate based on Equations (5.3) and (5.4) every decision period. Once an insurant is infected, the second decision is made whether or not to receive medical treatment based on Equations (5.1) and (5.2). The insurants only interact with their connecting neighbors, i.e., neighboring nodes in the network. The insurer's incentive policy x_{reward} and coinsurance $x_{sharing}$ are determined at the beginning of the simulation. An insurant's decision is determined via comparing expected cost based on the probability of being infected, the indirect cost of vaccination, the incentive and their neighboring nodes' health states, etc. At the end of each iteration, the health states of all insurants are updated to simulate the spread of flu. The process continues until the end of the flu season.

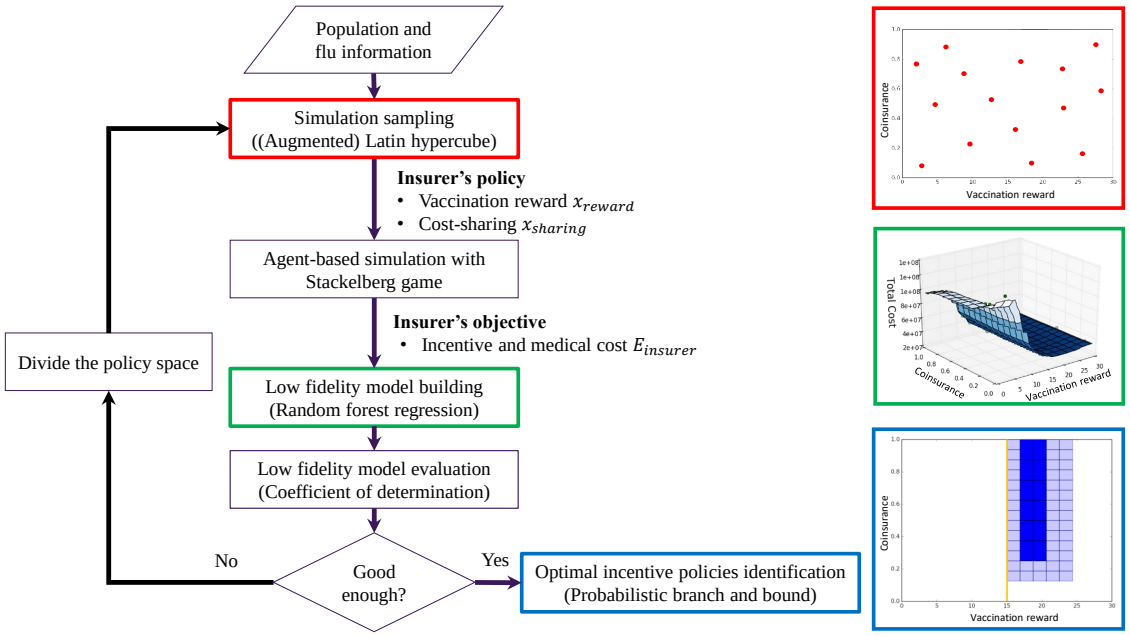


Figure 5.2: Simulation optimization and machine learning framework in the agent-based simulation and Stackelberg game.

5.1.3 Simulation Optimization and Machine Learning to Analyze Healthcare Incentive Policies

The agent-based simulation with the Stackelberg game provides a realistic representation of a population during the flu season. Moreover, it can be used to predict the population's behavior under different insurance policies and optimize such policies. However, fully exploring the original simulation for incentive policy analysis requires expensive computation due to consideration of all individuals' decision-making behaviors in the model. To efficiently identify an optimal incentive policies with improved computation burden, the advantages of simulation optimization algorithm and machine learning approach are combined.

From insurer's perspective, the proposed agent-based simulation with embeded Stackelberg vaccination game in Sections 5.1.1 and 5.1.2 can be seen as a black box, with vaccination reward x_{reward} and coinsurance $x_{sharing}$ as the input, and the observed cost $E_{insurer}$ as the output. A low-fidelity model, similar to surrogate modeling or meta-modeling approaches, is developed to reduce the complexity of the proposed agent-based simulation to a less complex, more computa-

tionally tractable model for large population-based study. This research applies the random forest regression to construct a validate low-fidelity model. Further, a low-fidelity model with statistical and global optimum-seeking techniques are coupled. This research incorporates the random forest regression model into Probabilistic Branch-and-Bound (PBnB) algorithm ([165]) to analyze the incentive policies and corresponding population behavior. The proposed methodology is presented in Figure 5.2 and consists of four main steps:

Simulation sampling using Latin hypercube design

First, a space-filling Latin Hypercube sampling design ([137]) is used to take several sampling points (i.e., x_{reward} and $x_{sharing}$) in the initial set as input to the agent-based simulation. As illustrated in Figure 5.2 at the upper-right position, Latin Hypercube sampling design ensures that each sampling space dimension is evenly sampled. In the first iteration, the number of sampling points for the initial space of incentive policies is specified. If more sampling data are needed, augmented Latin Hypercube sampling design is adopted to collect more sampling points in later iterations ([9]). Augmented Latin Hypercube sampling design adds more points to the original Latin Hypercube sampling design within each subspace, while still maintaining the Latin properties.

Low-fidelity model design using random forest regression

After running agent-based simulation at each sampling point, a data set consisting of policies (inputs) along with its corresponding total cost (outputs) is generated. Using the set of incentive policies (x_{reward} and $x_{sharing}$) as the independent variables, and the overall cost ($E_{insurer}$) as dependent variable, machine-learning techniques (specifically, random forest regression) is used to build a low-fidelity model capturing the input-output relationship in the black box. An example of the low-fidelity model is illustrated in Figure 5.2 at the center-right position. Random forests are one of the widely-used ensemble machine learning techniques ([77]). It is useful for regression tasks due to its simplicity and quality of the fitting even without hyper-parameter tuning. To understand infectious disease models, other techniques like the active-learning approach are also used to build a surrogate model ([159]). These approaches give a good approximation of the response variables when the whole model parameters are given, however, the learning process is untraceable if the model

involves a large number of parameters.

Low-fidelity model evaluation using coefficient of determination

The model fit is measured by the coefficient of determination (R^2). The R^2 value is calculated by the proportion of the variance in the overall cost that is predictable from the given incentive policies. A R^2 score value close to 1 indicates that the random forest regression can replicate the observed response well. The R^2 measurement is used to evaluate the performance of the random forest regression as the low-fidelity model. If the R^2 is smaller than a preassigned threshold, the policy space is dived into two subsets, and use Augmented Latin Hypercube sampling design to sample more points and build a low-fidelity model using the random forest regression again. This continues until the R^2 value achieves an input threshold for all subsets.

Optimal incentive policies identification using probabilistic branch-and-bound

PBnB is applied on the low-fidelity model for analyzing the incentive and cost-sharing policy. Using the property of partition-based algorithm, PBnB can approximate a level set of good solutions with statistical confidence (e.g., estimate top 10 percent quantile with %90 confidence intervals) by the partitioned solution space, instead of identifying the global optimal. PBnB is thus especially useful for solving healthcare management problem since the optimal solution is often impractical, and decision-makers require more flexibility in implementing healthcare policy.

The PBnB algorithm has two primary components. The first component focuses on estimating the target quantile with confidence interval using ordered sample function values. The number of sample points and replications mainly depends on the desired confidence level. Processing subregions of the solution space based on the confidence interval, the second component in PBnB labels each subregion with maintained, pruned, or undecided. In each iteration, PBnB statistically confirms the maintained and pruned subregions, and branches the rest of the subregions as the undecided for further exploring. For examle, the figure at the lower-right position in Figure 5.2 shows an example of PBnB analysis, where the deep blue (dark) regions are maintained regions representing a set of top 10 percent quantile solutions.

5.2 Numerical Result

5.2.1 Experiment setting

Agent-based simulation modeling of influenza and game model

This section demonstrates the mechanism on a stochastic influenza epidemic simulation model - FluTE ([35]), which has been used in many policy-making studies to compare the effectiveness of interventions ([102, 85]). The FluTE is written in C++ and simulates the stochastic spread of influenza across age- and risk- structured population of individuals interacting in known population groups. The Stackelberg game with realistic incentive policies is Incorporated into FluTE with a natural history of influenza infection.

The FluTE model is first extended by (i) adding two more health states, i,e, flu-associated hospitalizations and flu-associated death, which CDC uses to estimate influenza disease burden in the United States each year; (ii) replacing the policy maker or government-led compulsory interventions, i.e., vaccination intervention and antivirus treatment, with individual decision-making governed by self-interested behavior (Equations (5.1)-(5.4) in Section 5.1.2).

Hospitalization and death have severe impacts on public health in flu seasons. The hospitalization and death states (absorbing status compartments H and D) are incorporated into FluTE based on flow diagram developed in Section 5.1.1 (Figure 5.1). Death is modeled as an absorbing state. Flu-associated hospitalization is also modeled as an absorbing state because hospitalization causes huge medical cost and once the insurant is hospitalized due to flu complications, the insurants' next health state has no impact on the policy maker's incentive settings.

Then, the insurants' vaccination and treatment-taking decision-making processes (Equations (5.1)-(5.4)), and the insurer's total cost calculation ($E_{Insurer}$ in Equation (5.5)) are added into FluTE. The cost and infection probability of the population in the agent-based simulation model are associated with age and risk type. Table F.1 in Appendix A lists the data that are used in obtaining the optimal solutions and simulation inputs. To account for variation in costs by age and risk of severe complications from influenza, the cost is estimated for five age groups: <5, $5-17$, $18-49$, $50-64$, and ≥ 65 years. In addition, the cost is varied for high-risk and non-high-risk groups. The probability of influenza-attributable hospitalizations and deaths and the corresponding medical cost and the

value of the lost productive day (indirect cost) are from [110], which estimated the annual economic burden of influenza epidemics. The age 65+ group has higher hospitalization and death probability than other groups. Healthcare costs and productivity losses are greater in high-risk groups than in non-high-risk groups.

The use of the proposed mechanism and optimization approach are illustrated by simulating epidemics in metropolitan Seattle with a population of $563,441$ agents. The agents are randomly generated using the U.S.-wide family-size distribution from the 2000 Census. The FluTE model has been calibrated to past influenza pandemics - Asian A (H2N2) and 2009 novel influenza A (H1N1) influenza, so that outcomes are consistent with these influenza viruses. The default values of the flu and Seattle population parameters are adopted in the model but further extend it for more comprehensive infection phases for the Stackelberg vaccination game.

In the configuration file, the number of agents initially infected at the beginning of simulation is set to 100. No healthcare intervention, e.g., compulsory pre-vaccination, is implemented before the epidemic. Fractions of individuals in each of the five age groups who are at high risk of complications from influenza are set to 0.089 0.089, 0.212, 0.212, and 0.9. Fractions of individuals in each of the five age groups who are pregnant are 0, 0, 0.02539, 0.01896, and 0.

Random forest regression setting

Random forest regression in scikit-learn, which is a free software machine learning library for the Python programming language, is used. In the parameter setting, the number of trees in the forest (n_estimators) is set to 100 and the maximum depth of the tree (max_depth) is set to eight. Each subregion necessitates 100 simulation sampling points in each iteration to build its low-fidelity regression model.

Probabilistic branch-and-bound setting

In the probabilistic branch-and-bound algorithm, the results are presented in two dimensions (vaccination reward and coinsurance) to visualize the partitioning and level set approximation. Two key user-defined parameters in PBnB are closeness parameter δ and error rate α, where δ is used for determining the target solution threshold in terms of desired level quantile of the objective function's

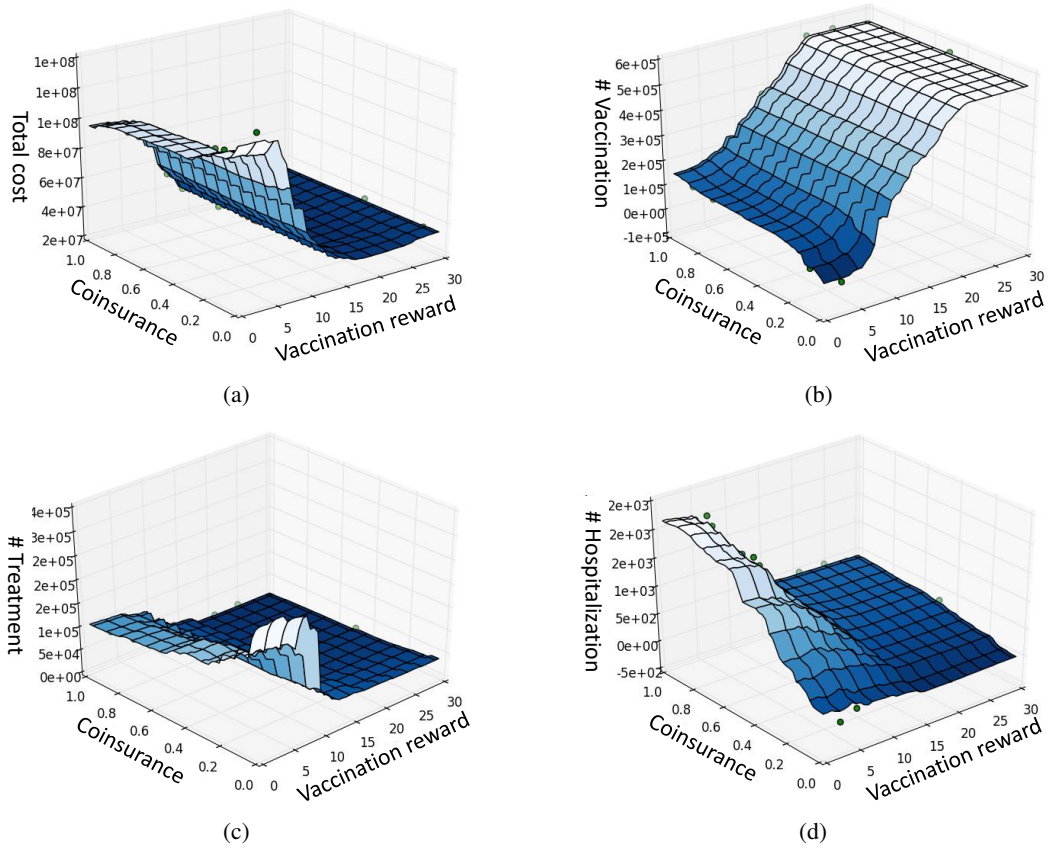


Figure 5.3: Base case: (a) total cost (b) vaccinated population size (c) treated population size (d) hospitalized population size with respect to vaccination reward and coinsurance rate

range distribution, and α is used for making probability bounds for the quality of the approximation. PBnB provides a $1-\alpha$ confidence interval of the desired quantile level δ. δ is set to 0.1 and α is set to 0.05. That is, it is found that the best 10 percent solution with 95% confidence interval.

5.2.2 Sensitivity Analysis

Base case

The base case sets public cost weight to \$200, vaccine efficacy to 60%, and attack rate to 1.6. Figure 5.3 presents the total cost function and numbers of vaccination, treatment, and hospitalization derived from the random forest regression with respect to vaccination reward and coinsurance. The

100 dots indicate the observed values.

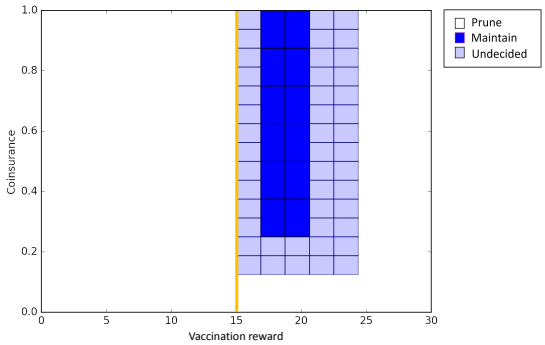
From the regression outputs in the base case, it is observed that the effect of vaccination reward on optimal healthcare policies outweighs the effect of coinsurance, indicating that the policy maker should put prevention instead of cure and treatment in a higher priority. However, the coinsurance rate has a significant influence on the number of hospitalization and number of vaccination when the incentive is low. Higher coinsurance rate will see as penalty for the infected insurants, so it will encourage more vaccination behavior (see Figure 5.3b). However, it also discourages treatment-seeking behavior due to exorbitant medical care cost (see Figure 5.3c), resulting in high hospitalization rate (see Figure 5.3d).

Figure 5.4a presents the PBnB output with random forest regression model of total cost as input. The result indicates that low vaccination reward and low coinsurance result in high total cost (prune region: light blue color/light gray) due to high infected case, hospitalization, and death. Set incentive too high also leads to high total cost since almost all population are vaccinated and the thus the marginal reduction of medical cost cannot catch up with increasing incentive cost. Low total cost region (maintained region: deep blue color/dark gray) appears in the medium coinsurance rate side. To give more specific policy recommendations, the result of PBnB in Figure 5.4a identifies the top 5% incentive policies with the base case setting. The coinsurance rate should be in $25\% - 100\%$ and the incentive ranges from $\$16.975 - \$20.625$.

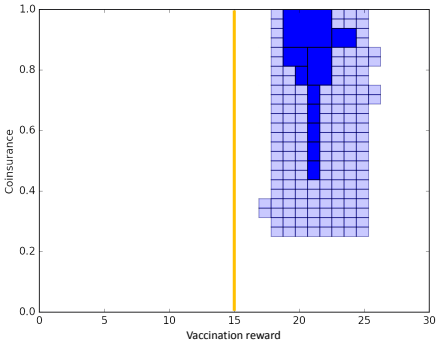
Sensitivity of the incentive policies to the following parameters is examined: 1) the vaccine efficacy, 2) the flu attack rate, and 3) the weight of public health w. Three experiments where values of vaccine efficacy, attack rate, and cost weight are selected according to Table 5.2 are conducted. Note that the values in the central column (vaccine efficacy 60%, attack rate 1.6, and public cost weight \$200) are the parameter settings in the base case. One parameter value is changed in each case, e.g., low case, therefore, there are a total of seven cases(one base case, three low cases, and three high cases).

Experiment 1 Vaccine efficacy

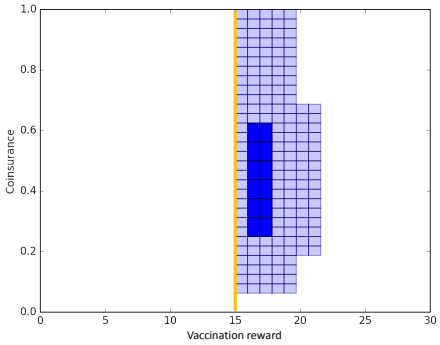
The first sensitivity analysis studies how different values for vaccine efficacy affect the healthcare management policy. Figures 5.4b, 5.4a, and 5.4c show a one-way sensitivity analysis of total cost



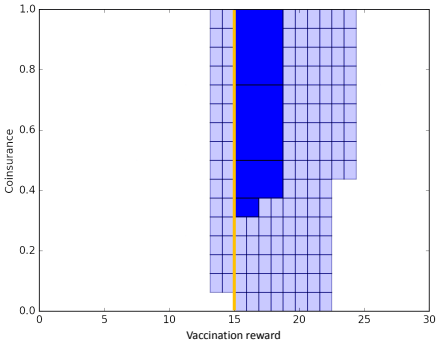
(a) Base case (vaccine efficacy = 30%, attack rate = 1.6, and w = \$200)



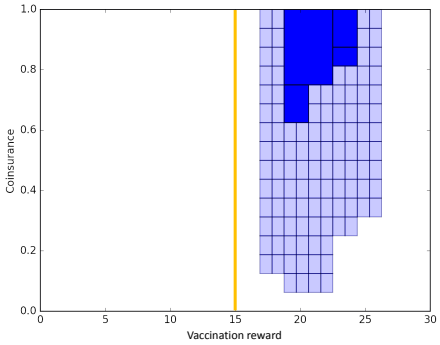
(b) Low vaccine efficacy (30%)



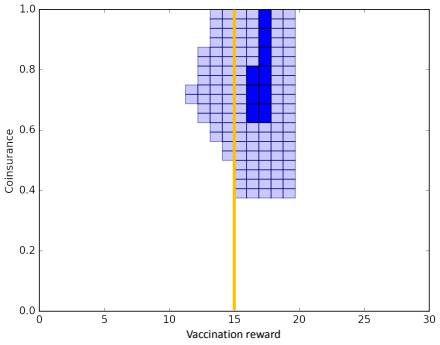
(c) High efficacy (90%)



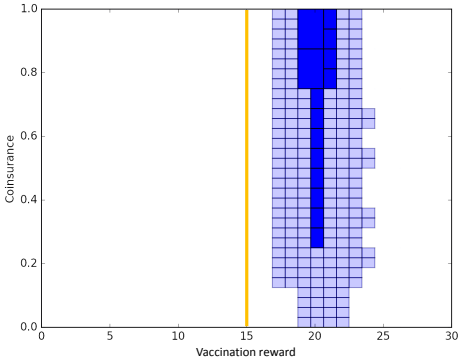
(d) Low attack rate (1.2)



(e) High attack rate (2)



(f) Low public cost weight (\$0)



(g) High public cost weight (\$400)

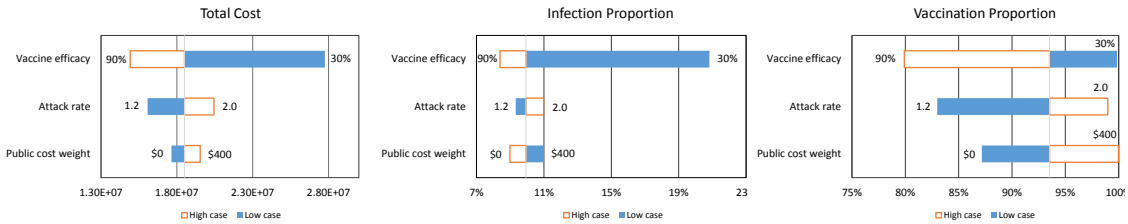
Figure 5.4: Sensitivity analysis

($E_{Insurer}$) to values of vaccine efficacy 30%, 60%, 90%, respectively, with fixed base values for attack rate and public cost weight. It is observed that the optimal maintained region gradually shifts to higher incentive setting as vaccine efficacy improves. The reason is that higher vaccine efficacy leads to lower infection probability ($p^a_{F|V}$). The threshold value of x_{reward} which makes $E_s^{a,r} > E_v^{a,r}$ for each individual agent is therefore lower.

It is also observed that Figures 5.4b, 5.4a, and 5.4c demonstrate the advantage of PBnB, which provides policy makers more flexibility to implement the healthcare policy even with uncertainty or have limited information about vaccine efficacy. If the vaccine efficacy is between 30% and 60%, the overlapping maintained region between Figures 5.4a and 5.4b gives set of optiaml incentive policies, i.e., coinsurance rate above 81.25% and incentive setting between \$18.75 and \$20.625; on the other hand, the results of PBnB in Figures 5.4a and 5.4c suggest implementing healthcare policies with vaccination reward between \$16.875 and \$17.8125 and coinsurance between 25% and 75% when the vaccine efficacy is in the range between 60% and 90%.

Experiment 2 Attack rate

The second sensitivity analysis studies a fundamental parameter that captures how infectious the flu is, i.e., illness attack rate. The attack rate is the cumulative incidence of infection in a group of people observed over a period of time during an epidemic. Higher attack rate indicates that the higher number of new cases of the disease in the population during a specified period. The age-specific illness attack rates in FluTE directly depend on the reproduction number, R_0, the mean number of new infections from a single infected individual in a susceptible population. R_0 is used to represent the attack rate. Thus, tuning the value of R_0, different illness attack rates can be simulated. It is observed from Figures 5.4a, 5.4d, and 5.4e that more vaccination reward should be offered to encourage vaccination behaviors when the attack rate is high. It is also observed in the base and high attack rate cases, the subregions with low incentive are pruned with statistical confidence in the early iteration due to relatively high hospitalization and death cost.



(a) Total cost (b) Infection proportion (c) Vaccination proportion

Figure 5.5: Comparison of six sensitivity analysis cases and the base case

Experiment 3 Public cost weight

The last sensitivity analysis changes the value of the public cost weight ($0, $200, and $400) to see how decision-making is affected by how policy-maker values the population health outcome. The results in Figure 5.4g suggest the government-run health insurance agency such as a program run by U.S.. The insurer moves first byfederal, state, or local governments, may choose a high public cost weight, whereas the private insurance company may choose a low public cost weight. The decision maker with higher public cost weight tend to raise the vaccination reward to encourage vaccination behavior and avoid infection.

Comparing to the case of low public cost weight, PBnB prunes the low incentive subregion in the first iteration in the case of high public cost. The reason is because the total cost in Equation (5.5 is more sensitive to the changes in the flu-infected population size due to high weight.

5.2.3 Population outcome comparison

In order to analyze how the uncertainties of vaccine efficacy, attack rate, and public cost weight affect other population outcomes like flu infection and vaccination, the incentive settings in the maintained region (top 10% solution quality) are applied to calculate the average total cost, average infection proportion, and average vaccination proportion in seven cases. The average total cost, infection proportion, and vaccination proportion in the base case are 1.85×10^7, 9.93%, and 93.52%, respectively. Figure 5.5 compares the average total cost, infection proportion, and vaccination proportion in the base case to the other six cases. As seem in Figure 5.5a, low vaccine efficacy, high

attack rate, and high public cost weight increase the total cost from the base case. Especially the vaccine efficacy has a significant impact not only on total cost but also infection proportion (Figure 5.5b).

It is also observed that, even though the vaccinate proportion decreases 10% in the low attack rate case, the infection proportion still lower than its counterpart in the base case. In the high attack rate case, on the other hand, more than 5% population get vaccinated to keep infection proportion low (only 1.06% higher than its compartment in the base case).

It is also found that the high weight case has a higher total cost because a higher incentive is given to improve vaccination behavior (results in Experiment 3). The vaccination proportion therefore increases 6.5% comparing to the base case (Figure 5.5c) to reduce the infection proportion (around 1% in Figure 5.5b).

5.3 Conclusions

This chapter considered a seasonal flu outbreak prevention problem where decision maker's goal is to design the insurance-based healthcare incentives to achieve minimum overall cost, including not only medical care cost but public benefit, and proposed a mechanism to address the problem.

In the mechanism, two effective healthcare incentive policies, vaccination reward and cost-sharing for antivirus treatment taking to prevent the flu outbreak are combined. Momentary reward for the susceptible population was adopted as reimbursement to incentivize vaccination behavior. Coinsurance in healthcare insurance was selected as an out-of-pocket approach. An appropriate coinsurance charge encourages antivirus treatment-seeking behavior and avoids overuse of health care.

To have a more realistic representation of dynamic interaction between the decision maker and the population. The Stackelberg vaccination game, where the decision maker and population play as the leader and followers was proposed, for its advantage of the individual-level mathematical description of a social situation in which two or more individuals, interact. Healthcare incentives and flu-related parameters, e.g., vaccine efficacy, have an influence on individual's vaccination and antivirus treatment taking behaviors during flu season and, as a result, affect the consequent effects at the population level.

To evaluate the effects of healthcare incentive policies with disease prevalence over time and simulate the propagation of influenza through a population, the proposed game-theoretic model was incorporated into the FLuTE model. Two health status-hospitalization and death are added, and the decision-making process and corresponding responses are integrated to FLuTE. Also, medical care cost, loss of productivity, hospitality and death probability for all age-specific and risk-type groups are added as the FLuTE parameters input and generated overall medical care cost and public welfare as a final output.

PBnB is implemented for its advantage over other optimization techniques as discussed in Section 5.1. In order to reduce computation time in large-scale simulations, we applied random forest regression for building a low-fidelity model and fed into PBnB algorithm instead of using the original agent-based model. Application of PBnB and machine learning techniques allow us to effectively analyze healthcare policies in a complicated flu transmission model.

Experimental results indicate that (1) both incentive reward and cost-sharing charge are effective approaches to encourage vaccination behavior and minimize the overall cost; and (2) the designed mechanism can motivate the insurants to maintain a low infection rate in the population with respect to different vaccine efficacy and attack rates while taking vaccination and medical cost paid by insurer and social benefit into consideration.

An incentive-based insurance mechanism was proposed to prevent the spread of seasonal influenza for a single health insurer and multiple insurants. The model captures four critical sources of uncertainty in the insurer's perspective such as uncertainty in population health status and behavior, the spread of flu and vaccine effectiveness. The combination of Stackelberg vaccination game and agent-based simulation approaches are resorted to provide a more systematic insight as to how private/public insurer would incentivize self-interest insurants to uptake vaccine and seek for antiviral treatment in flu season.

Table 5.1: Notation for Stackelberg vaccination game modeling

Indices	
a	index of insurant's age, $a \in \{0-4, 5-17, 18-49, 50-64, 65+\}$
r	index of insurant's risk type, $r \in \{\text{normal}, \text{high risk}\}$
Input parameters	
$c^{a}_{med,U}$	medical cost in the case of no medical attendance for an untreated insurant with age a
$c^{a,r}_{med,T}$	medical cost of outpatient visit for a treated insurant with age a and risk type r
$c^{a,r}_{med,H}$	medical cost of hospitalization for a hospitalized insurant with age a and risk type r
$c^{a,r}_{med,D}$	medical cost of death for a dead insurant with age a and risk type r
$c^{a}_{ind,U}$	indirect cost, i.e., lost productivity, not medically attended for an untreated insurant with age a
$c^{a,r}_{ind,H}$	indirect cost, i.e., lost productivity, for a hospitalized insurant with age a and risk type r
$c_{ind,V}$	indirect vaccination cost, e.g., the value of work loss time for vaccination
f_H	coinsurance rate for hospitalization in healthcare insurance
f_D	coinsurance rate for death in healthcare insurance
$p^{a}_{F\|S}$	probability that an unvaccinated insurant with age a will eventually be infected
$p^{a}_{F\|V}$	probability that a vaccinated insurant with age a will eventually be infected
$p^{a}_{H\|F}$	probability that an infected insurant with age a is hospitalized
$p^{a}_{D\|F}$	probability that an infected insurant with age a dies due to flu-related illness
w	weight of social benefit
Decision variables	
x_{reward}	monetary vaccination reward in healthcare insurance given to vaccinated insurant (unit: dollar)
$x_{sharing}$	coinsurance rate in healthcare insurance (unit: percentage)
Intermediate variables	
$E^{a,r}_{U}$	expected cost of insurant with age a and risk type r when not taking treatment
$E^{a,r}_{T}$	expected cost of insurant with age a and risk type r when taking treatment
$E^{a,r}_{S}$	expected cost of insurant with age a and risk type r when not vaccinated
$E^{a,r}_{V}$	expected cost of insurant with age a and risk type r when vaccinated
Output	
ϕ_V	proportion of vaccinated insurants
ϕ_F	proportion of infected insurants
$\phi^{a,r}_{T\|F}$	proportion of the infected insurants that receive antivirus treatment
$\phi^{a,r}_{H\|F}$	proportion of the infected insurants that are hospitalized
$\phi^{a,r}_{D\|F}$	proportion of the infected insurants that die

Table 5.2: Parameter setting in sensitivity analysis

		Low case	Base case	High case
Experiment 1	Vaccine efficacy	30%	60%	90%
Experiment 2	Attack rate	1.2	1.6	2.0
Experiment 3	Public cost weight	\$0	\$200	\$400

Chapter 6

CONCLUSIONS AND FUTURE RESEARCH

This dissertation focuses on three real-world healthcare delivery and insurance design applications: dynamic budget allocation for HCV elimination, healthcare insurance design for hypertension, and incentive mechanism design for preventing the outbreak of seasonal influenza. This research addresses two main research objectives for each of the three healthcare policy-making problems. First, to design sequential decision support models. Second, to develop corresponding optimization methodologies to efficiently discover policies/interventions.

The proposed models and methodologies provide improved healthcare outcomes and reduces costs under these three applications. The dissertation also presents new approaches to guarantee the intervention quality and identify a monotonic policy in computational intensive optimization problems

In Chapter 3, an efficient algorithm is introduced to solve high-dimensional dynamic resource allocation problems in population disease management. The allocation problem is formulated as a high-fidelity health outcomes maximization model with large-state and multiple-period dynamic decision-making. To address the computational challenges, a low-fidelity model is proposed by approximating the future disease progression and population evolution, recognizing the properties of absorbing Markov chain with stationary policies. The low-fidelity approximation is embedded into the high-fidelity optimization model and a rollout algorithm is developed to efficiently identify a good sequential intervention policy. Furthermore, a RA-BnB algorithm is then designed by integrating the rollout algorithm into the branch-and-bound framework, which provides an optimality gap between best possible policy and best found policy. Using the RA-BnB, 30 year-horizon could be evaluated to explore elimination of the disease, which previously was computationally intractable.

In Chapter 4, a game-theoretic framework, stochastic Stackelberg game, is applied to design an optimal cost-sharing plan in healthcare insurance, coupled with the long-term interaction between the insurer and the insurant with disease progression. The mechanism is then applied to a case of

hypertension medication use and blood pressure control. The impact of copayment on medication utilization for insurants is examined with different ages and incomes. The results complement related evidence in the literature and shed light on an important connection among dynamic health status, cost-sharing policy, and medical treatment adherence. Using the game-theoretic analysis has the potential to provide a customized and personalized health insurance design for targeted insurants with similar characteristic (age, income, and medical history).

In Chapter 5, an incentive-based insurance mechanism, including a reward for vaccination and cost-sharing for treatment, is proposed to prevent the outbreak of seasonal influenza. The Stackelberg vaccination game and agent-based modeling approaches are combined to provide an insight as to how private/public insurer would incent self-interest insurants to uptake vaccine and seek for antiviral treatment in the flu season. The model also captures critical sources of uncertainty in the insurer's perspective such as uncertainty in population health status and behavior, the spread of seasonal flu, and vaccine efficacy. In order to reduce computation time in large-scale simulations, instead of running the original ABM model, random forest regression is applied for building a low-fidelity model and fed into PBnB to analyze the healthcare interventions. Integration of machine learning techniques and PBnB algorithm allows us to effectively analyze healthcare interventions in a complicated flu transmission model under different attack rates, vaccine efficacy, and level of health and social benefits.

While the research presented in Chapter 3 aims to impact HCV care, it can be generalized to obtain other health policy decisions regarding resource allocation. The proposed algorithm is applicable to other diseases with quantifiable natural history, accepted screening intervention and effective but costly treatment intervention. For example, mitigating depression has become a national health priority as it affects 1 out of 10 American adults. Accurate assessment of the natural history of depression is well-studied [16]. The USPSTF recommended screening American adults for depression in 2016 [150] while antidepressant and psychotherapy are effective but costly treatment options. Population depression screening and treatment implementation is a major challenge in many healthcare systems and should be further studied. Another example is the overdose of prescription opioids; the related harm has now reached epidemic levels in America. However, self-reported misuse of prescription opioids is still low. U.S. Department of Health and Human Services (HHS) focuses the efforts on better prescription drug monitoring and improvement of addiction

treatment [75]. A qualitative study exploring the natural history of opioid misuse and abuse is also conducted [146]. The proposed approaches in this thesis have potential to examine the long-term resource allocation and management in depression and opioid.

The stochastic Stackelberg game applied in Chapter 4 can be extended to a scenario where the insurer has incomplete information on the insurant, e.g., race, medical record, and health information. Different types of insurants have different transition probabilities and immediate costs; this scenario can be solved by transforming the insurer's incomplete information regarding the insurant into an imperfect information game. That is, the problem can be mapped to a Bayesian Stackelberg game by considering different types of homogeneous insurants.

In Chapter 5, two health states (hospitalization and death) and decision-making processes (vaccination and treatment-taking) are integrated into a calibrated stochastic influenza epidemic simulation model. To build a realistic social network in flu transmission model, integrating geometric data and socioeconomics in the simulation process are required. This is a future research direction for expansion of the current study. Moreover, the proposed methodology has the potential to apply to the outbreak prevention of other infectious diseases. For example, the recent measles outbreaks due to dwindling vaccination rate in these two decades in the USA highlight the importance of fighting the disease through immunization. Determining the intervention setting of other vaccination promotion programs through communication and campaign resources (e.g., incentive or advertisements) for other infectious diseases could be a future area of research.

This dissertation provides insights in disease management and it warrants extensive research, which may pave the way for new and more efficient methods for future research in healthcare delivery and insurance design problem.

BIBLIOGRAPHY

[1] M. Akan, O. Alagoz, B. Ata, F. S. Erenay, and A. Said. A broader view of designing the liver allocation system. *Operations Research*, 60(4):757–770, 2012.

[2] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts. The optimal timing of living-donor liver transplantation. *Management Science*, 50(10):1420–1430, 2004.

[3] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts. Determining the acceptance of cadaveric livers using an implicit model of the waiting list. *Operations Research*, 55(1):24–36, 2007.

[4] S. S. Alistar, E. F. Long, M. L. Brandeau, and E. J. Beck. HIV epidemic control-A model for optimal allocation of prevention and treatment resources. *Health Care Management Science*, 17(2):162–181, 2014.

[5] American Public Health Association. Public health and chronic disease cost savings and return on investment, 2014. Available via `https://www.apha.org/~/media/files/pdf/factsheets/chronicdiseasefact_final.ashx`. [accessed: 5 May 2017].

[6] R. Andersen and J. F. Newman. Societal and individual determinants of medical care utilization in the united states. *The Milbank Quarterly*, 83(4):Online–only, 2005.

[7] K. J. Arrow. Uncertainty and the welfare economics of medical care. *The American Economic Review*, 53(5):941–973, 1963.

[8] K. J. Arrow. *Welfare analysis of changes in health coinsurance rates*. National Bureau of Economic Research, 1976.

[9] augmentLHS. augmentLHS: Augment a Latin hypercube design, 2019. Available via `https://rdrr.io/cran/lhs/man/augmentLHS.html/` [accessed: 10 June 2019].

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

[11] T. Ayer, C. Zhang, A. Bonifonte, A. C. Spaulding, and J. Chhatwal. Prioritizing hepatitis C treatment in US prisons. *Operations Research*, 2019.

[12] S. Barrett. The smallpox eradication game. *Public Choice*, 130(1):179–207, 2007.

[13] C. T. Bauch and D. J. D. Earn. Vaccination and the theory of games. *Proceedings of the National Academy of Sciences of the United States of America*, 101(36):13391–13394, 2004.

[14] C. T. Bauch, A. P. Galvani, and D. J. D. Earn. Group interest versus self-interest in smallpox vaccination policy. *Proceedings of the National Academy of Sciences of the United States of America*, 100(18):10564–10567, 2003.

[15] C. Beauchemin, J. Samuel, and J. Tuszynski. A simple cellular automaton model for influenza a viral infections. *Journal of Theoretical Biology*, 232(2):223–234, 2005.

[16] A. T. Beekman, S. W. Geerlings, D. J. Deeg, J. H. Smit, R. S. Schoevers, E. de Beurs, A. W. Braam, B. W. Penninx, and W. van Tilburg. The natural history of late-life depression: A 6-year prospective study in the community. *Archives of General Psychiatry*, 59(7):605–611, 2002.

[17] D. P. Bertsekas. Rollout algorithms for discrete optimization: A survey. In *Handbook of Combinatorial Optimization*, pages 2989–3013. Springer, 2013.

[18] D. P. Bertsekas and D. A. Castanon. Rollout algorithms for stochastic scheduling problems. In *Decision and Control, 1998. Proceedings of the 37th IEEE Conference on*, volume 2, pages 2143–2148. IEEE, 1998.

[19] D. Bertsimas, V. F. Farias, and N. Trichakis. Fairness, efficiency, and flexibility in organ allocation for kidney transplantation. *Operations Research*, 61(1):73–87, 2013.

[20] D. Bertsimas, A. O'Hair, S. Relyea, and J. Silberholz. An analytics approach to designing combination chemotherapy regimens for cancer. *Management Science*, 62(5):1511–1531, 2016.

[21] D. P. Bertsimas and R. Demir. An approximate dynamic programming approach to multidimensional knapsack problems. *Management Science*, 48(4):550–565, 2002.

[22] J. R. Betancourt, A. R. Green, J. E. Carrillo, et al. *Cultural competence in health care: Emerging frameworks and practical approaches*, volume 576. Commonwealth Fund, Quality of Care for Underserved Populations New York, NY, 2002.

[23] C. Betsch, R. Bohm, and G. B. Chapman. Using behavioral insights to increase vaccination policy effectiveness. *Policy Insights from the Behavioral and Brain Sciences*, 2(1):61–73, 2015.

[24] B. A. Briesacher, J. H. Gurwitz, and S. B. Soumerai. Patients at-risk for cost-related medication nonadherence: A review of the literature. *Journal of General Internal Medicine*, 22(6):864–871, 2007.

[25] P. A. Briss, L. E. Rodewald, A. R. Hinman, A. M. Shefer, R. A. Strikas, R. R. Bernier, V. G. Carande-Kulis, H. R. Yusuf, S. M. Ndiaye, and S. M. Williams. Reviews of evidence regarding interventions to improve vaccination coverage in children, adolescents, and adults. *American Journal of Preventive Medicine*, 18(1 Suppl):97–140, 2000.

[26] E. T. Bronchetti, D. B. Huffman, and E. Magenheim. Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination. *Journal of Economic Behavior and Organization*, 116:270–291, 2015.

[27] R. Brook, J. Ware, W. Rogers, E. Keeler, A. Davies, C. Sherbourne, and et al. The effect of coinsurance on the health of adults: Results from the RAND Health Insurance Experiment, Report R-3055-HHS, ISBN 0-8330-0614-2, Santa Monica, CA, 1984.

[28] M. Brunström and B. Carlberg. Effect of antihypertensive treatment at different blood pressure levels in patients with diabetes mellitus: Systematic review and meta-analyses. *British Medical Journal*, 352:i717, 2016.

[29] Cardiovascular Business. Treating uncontrolled hypertension makes dollars and sense, 2017. Available via `http://www.cardiovascularbusiness.com/topics/healthcare-economics/treating-uncontrolled-hypertension-makes-dollars-and-sense`. [accessed: 28 Aug. 2017].

[30] J. Catania, L. G. Que, J. A. Govert, J. W. Hollingsworth, and C. R. Wolfe. High intensive care unit admission rate for 2013–2014 influenza is associated with a low rate of vaccination. *American Journal of Respiratory and Critical Care Medicine*, 189(4):485–487, 2014.

[31] Centers for Disease Control and Prevention. National health and nutrition examination survey NHANES, 1999-2010. Available via `http://www.cdc.gov/nchs/nhanes/`. [accessed: April 1, 2018].

[32] Centers for Disease Control and Prevention (CDC). Vital signs: Awareness and treatment of uncontrolled hypertension among adults–United States, 2003-2010. *Morbidity and Mortality Weekly Report*, 61:703, 2012.

[33] Centers for Disease Control and Prevention (CDC). High Blood Pressure Facts, 2015. Available via `http://www.cdc.gov/bloodpressure/facts.htm/`. [accessed: 28 Aug. 2017].

[34] E. Chak, A. H. Talal, K. E. Sherman, E. R. Schiff, and S. Saab. Hepatitis C virus infection in USA: An estimate of true prevalence. *Liver International*, 31(8):1090–1101, 2011.

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

[36] L.-W. Chao. A comparison of consensus and nonconsensus approaches to modeling contraceptive choice behavior. *Health Economics*, 11(7):599–622, 2002.

[37] B. Chen and V. Y. Fan. Strategic provider behavior under global budget payment with price adjustment in taiwan. *Health Economics*, 24(11):1422–1436, 2015.

[38] M. E. Chernew, A. B. Rosen, and A. M. Fendrick. Value-based insurance design. *Health Affairs*, 26(2):w195–w203, 2007.

[39] S. E. Chick, S. Hasija, and J. Nasiry. Information elicitation and influenza vaccine production. *Operations Research*, 65(1):75–96, 2016.

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

[41] A. V. Chobanian, G. L. Bakris, H. R. Black, W. C. Cushman, L. A. Green, J. L. Izzo, D. W. Jones, B. J. Materson, S. Oparil, J. T. Wright, et al. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *Hypertension*, 42(6):1206–1252, 2003.

[42] V. Conitzer and T. Sandholm. Computing the optimal strategy to commit to. In *Proceedings of the 7th ACM Conference on Electronic Commerce*, pages 82–90, 2006.

[43] CostHelper. High blood pressure treatment cost, 2017. Available via `http://health.costhelper.com/treating-high-blood-pressure-cost.html`. [accessed: 28 Aug. 2017].

[44] D. M. Cutler and R. J. Zeckhauser. The anatomy of health insurance. *Handbook of Health Economics*, 1:563–643, 2000.

[45] M. M. Denniston, R. M. Klevens, G. M. McQuillan, and R. B. Jiles. Awareness of infection, knowledge of hepatitis C, and medical follow-up among individuals testing positive for hepatitis C: National Health and Nutrition Examination Survey 2001-2008. *Hepatology*, 55(6):1652–1661, 2012.

[46] S. Deo, S. Iravani, T. Jiang, K. Smilowitz, and S. Samuelson. Improving health outcomes through better capacity allocation in a community-based chronic care model. *Operations Research*, 61(6):1277–1294, 2013.

[47] S. Deo, K. Rajaram, S. Rath, U. S. Karmarkar, and M. B. Goetz. Planning for HIV screening, testing, and care at the Veterans Health Administration. *Operations Research*, 63(2):287–304, 2015.

[48] S. Deo and M. Sohoni. Optimal decentralization of early infant diagnosis of HIV in resource-limited settings. *Manufacturing & Service Operations Management*, 17(2):191–207, 2015.

[49] J. B. Dunham. An agent-based spatially explicit epidemiological model in mason. *Journal of Artificial Societies and Social Simulation*, 9(1), 2005.

[50] M. T. Eaddy, C. L. Cook, K. O'Day, S. P. Burch, and C. R. Cantrell. How patient cost-sharing trends affect adherence and outcomes: A literature review. *Pharmacy and Therapeutics*, 37(1):45, 2012.

[51] S. S. El-Kamary, R. Jhaveri, and M. D. Shardell. All-cause, liver-related, and non-liver-related mortality among HCV-infected individuals in the general US population. *Clinical Infectious Diseases*, 53(2):150–157, 2011.

[52] W. J. Elliott. The economic impact of hypertension. *The Journal of Clinical Hypertension*, 5(3):3–13, 2003.

[53] R. P. Ellis, S. Jiang, and W. G. Manning. Optimal health insurance for multiple goods and time periods. *Journal of Health Economics*, 41:89–106, 2015.

[54] R. P. Ellis and W. G. Manning. Optimal health insurance for prevention and treatment. *Journal of Health Economics*, 26(6):1128–1150, 2007.

[55] E. J. Emanuel and V. R. Fuchs. The perfect storm of overutilization. *Journal of the American Medical Association*, 299(23):2789–2791, 2008.

[56] B. J. Epstein. Improving blood pressure control rates by optimizing combination antihypertensive therapy. *Expert Opinion on Pharmacotherapy*, 11(12):2011–2026, 2010.

[57] J. Epstein, D. Cummings, and S. Chakravarty. Toward a containment strategy for smallpox bioterror: an individual-based computational approach. *Generative social science: Studies in agent-based computational modeling*, 2006.

[58] R. Feldman and B. Dowd. A new estimate of the welfare loss of excess health insurance. *The American Economic Review*, 81(1):297–301, 1991.

[59] J. Filar and K. Vrieze. *Competitive Markov decision processes*. Springer Science & Business Media, 2012.

[60] P. A. Fishman, V. Ding, R. Hubbard, E. J. Ludman, C. Pabiniak, C. Stewart, L. Morales, and G. E. Simon. Impact of deductibles on initiation and continuation of psychotherapy for treatment of depression. *Health Services Research*, 47(4):1561–1579, 2012.

[61] P. A. Fishman, V. Ding, R. Hubbard, E. J. Ludman, C. Pabiniak, C. Stewart, L. Morales, and G. E. Simon. Impact of deductibles on initiation and continuation of psychotherapy for treatment of depression. *Health Services Research*, 47(4):1561–1579, 2012.

[62] FluTE. FluTE, an influenza epidemic simulation model, 2010. Available via `https://www.cs.unm.edu/~dlchao/flute/`. [accessed: 5 April 2019].

[63] P. J. Francis. Optimal tax/subsidy combinations for the flu season. *Journal of Economic Dynamics and Control*, 28(10):2037–2054, 2004.

[64] F. Fu, D. I. Rosenbloom, L. Wang, and M. a. Nowak. Imitation dynamics of vaccination behaviour on social networks. *Proceedings of the Royal Society B: Biological Sciences*, 278(1702):42–49, 2010.

[65] A. P. Galvani, T. C. Reluga, and G. B. Chapman. Long-standing influenza vaccination policy is in accord with individual self-interest but not with the utilitarian optimum. *Proceedings of the National Academy of Sciences of the United States of America*, 104(13):5692–7, 2007.

[66] S. Ganzfried and T. Sandholm. Computing equilibria by incorporating qualitative models? In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1*, pages 183–190, Toronto, Canada, May 2010. International Foundation for Autonomous Agents and Multiagent Systems.

[67] J. Gerteis, D. Izrael, D. Deitz, L. LeRoy, R. Ricciardi, T. Miller, and B. Jayasree. Multiple chronic conditions chartbook. Rockville (MD): Agency for healthcare research and quality, 2014.

[68] A. Ghate. Lecture notes 5 in IND E 599B: Markov Decision Processes, 2017.

[69] D. Goldman and G. Joyce. Prescription drug cost sharing. *Jama*, 298(1):61–69, 2007.

[70] D. Goldman and T. Philipson. Integrated insurance design in the presence of multiple medical technologies. Technical report, National Bureau of Economic Research, 2007.

[71] D. P. Goldman, G. F. Joyce, and Y. Zheng. Prescription drug cost sharing: Associations with medication and medical utilization and spending and health. *Journal of the American Medical Association*, 298(1):61–69, 2007.

[72] T. J. Gordon. A simple agent model of an epidemic. *Technological Forecasting and Social Change*, 70(5):397–417, 2003.

[73] J. J. Grefenstette, S. T. Brown, R. Rosenfeld, J. DePasse, N. T. Stone, P. C. Cooley, W. D. Wheaton, A. Fyshe, D. D. Galloway, A. Sriram, et al. Fred (a framework for reconstructing epidemic dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations. *BMC Public Health*, 13(1):940, 2013.

[74] J. E. Helm, M. S. Lavieri, M. P. Van Oyen, J. D. Stein, and D. C. Musch. Dynamic forecasting and control algorithms of glaucoma progression for clinician decision support. *Operations Research*, 63(5):979–999, 2015.

[75] HHS. About the U.S. opioid epidemic, 2016. Available via `https://www.hhs.gov/opioids/about-the-epidemic/`. [accessed: April 1, 2018].

[76] T. Ho, P. A. Fishman, and Z. B. Zabinsky. Using a game-theoretic approach to design optimal health insurance for chronic disease. *IISE Transactions on Healthcare Systems Engineering*, (just-accepted):1–30, 2019.

[77] T. K. Ho. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282. IEEE, 1995.

[78] T.-Y. Ho, P. A. Fishman, and Z. B. Zabinsky. Designing and analyzing healthcare insurance policies to reduce cost and prevent the spread of seasonal influenza. In *Proceedings of the 2016 Winter Simulation Conference*, pages 2018–2029. IEEE, 2016.

[79] T.-Y. Ho, S. Liu, and Z. B. Zabinsky. Prevention of seasonal influenza outbreak via healthcare insurance. Working paper.

[80] T.-Y. Ho, S. Liu, and Z. B. Zabinsky. A multi-fidelity rollout algorithm for dynamic resource allocation in population disease management. *Health Care Management Science*, pages 1–29, 2018.

[81] T.-Y. Ho, S. Liu, and Z. B. Zabinsky. A branch and bound algorithm for dynamic resource allocation in population disease management. *Operations Research Letters*, 2019. Under review.

[82] J. Hoadley. Cost containment strategies for prescription drugs: Assessing the evidence in the literature, prepared for the Kaiser Family Foundation. *Health Policy Institute Georgetown University*, 2005.

[83] M. Joffres, E. Falaschetti, C. Gillespie, C. Robitaille, F. Loustalot, N. Poulter, F. A. McAlister, H. Johansen, O. Baclic, and N. Campbell. Hypertension prevalence, awareness, treatment and control in national surveys from England, the USA and Canada, and correlation with stroke and ischaemic heart disease mortality: A cross-sectional study. *British Medical Journal*, 3(8):e003423, 2013.

[84] D. Kalathil, V. S. Borkar, and R. Jain. Approachability in stackelberg stochastic games with vector costs. *Dynamic Games and Applications*, pages 1–21, 2014.

[85] P. Kasaie and W. D. Kelton. Simulation optimization for allocation of epidemic-control resources. *IISE Transactions on Healthcare Systems Engineering*, 3(2):78–93, 2013.

[86] W. O. Kermack and A. G. McKendrick. A Contribution to the Mathematical Theory of Epidemics. In *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, volume 115, pages 700–721. The Royal Society, 1927.

[87] A. Khademi, D. R. Saure, A. J. Schaefer, R. S. Braithwaite, and M. S. Roberts. The price of nonabandonment: HIV in resource-limited settings. *Manufacturing & Service Operations Management*, 17(4):554–570, 2015.

[88] N. Kong, A. J. Schaefer, B. Hunsaker, and M. S. Roberts. Maximizing the efficiency of the us liver allocation system through region design. *Management Science*, 56(12):2111–2122, 2010.

[89] J. R. Kramer, F. Kanwal, P. Richardson, T. P. Giordano, L. A. Petersen, and H. B. El-Serag. Importance of patient, provider, and facility predictors of hepatitis C virus treatment in veterans: A national study. *The American Journal of Gastroenterology*, 106(3):483–491, 2011.

[90] D. Lakić, G. Petrova, N. Bogavac-Stanojević, Z. Jelić-Ivanović, and M. Kos. The cost-effectiveness of hypertension pharmacotherapy in Serbia: A Markov model. *Biotechnology & Biotechnological Equipment*, 26(3):3066–3072, 2012.

[91] S. Larney, H. Kopinski, C. G. Beckwith, N. D. Zaller, D. D. Jarlais, H. Hagan, J. D. Rich, B. J. Bergh, and L. Degenhardt. Incidence and prevalence of hepatitis C in prisons and other closed settings: Results of a systematic review and meta-analysis. *Hepatology*, 58(4):1215–1224, 2013.

[92] J. C. Lauffenburger, C. L. Mayer, R. L. Hawke, K. L. Brouwer, M. W. Fried, and J. F. Farley. Medication use and medical comorbidity in patients with chronic hepatitis C from a US commercial claims database: High utilization of drugs with interaction potential. *European Journal of Gastroenterology & Hepatology*, 26(10):1073, 2014.

[93] B. Y. Lee, S. T. Brown, P. C. Cooley, R. K. Zimmerman, W. D. Wheaton, S. M. Zimmer, J. J. Grefenstette, T. M. Assi, T. J. Furphy, D. K. Wagener, and D. S. Burke. A Computer Simulation of Employee Vaccination to Mitigate an Influenza Epidemic. *American Journal of Preventive Medicine*, 38(3):247–257, 2010.

[94] C. P. Lee, G. M. Chertow, and S. A. Zenios. Optimal initiation and management of dialysis therapy. *Operations Research*, 56(6):1428–1449, 2008.

[95] J. Letchford, L. MacDermed, V. Conitzer, R. Parr, and C. L. Isbell. Computing optimal strategies to commit to in stochastic games. In *Processing of Twenty-Sixth National Conference on Artificial Intelligence*, 2012.

[96] Y. Li, H. Huang, Z. B. Zabinsky, and S. Liu. Optimizing implementation of hepatitis C birth-cohort screening and treatment strategies: Model-based projections. *MDM Policy & Practice*, 2(1), 2017.

[97] S. Liu, M. L. Brandeau, and J. D. Goldhaber-Fiebert. Optimizing patient treatment decisions in an era of rapid technological advances: The case of hepatitis C treatment. *Health Care Management Science*, 20(1):16–32, 2017.

[98] S. Liu, L. E. Cipriano, M. Holodniy, and J. D. Goldhaber-Fiebert. Cost-effectiveness analysis of risk-factor guided and birth-cohort screening for chronic hepatitis C infection in the United States. *PLoS One*, 8(3):e58975, 2013.

[99] S. Liu, L. E. Cipriano, M. Holodniy, D. K. Owens, and J. D. Goldhaber-Fiebert. New protease inhibitors for the treatment of chronic hepatitis C: A cost-effectiveness analysis. *Annals of Internal Medicine*, 156(4):279–290, 2012.

[100] S. Liu, M. Schwarzinger, F. Carrat, and J. D. Goldhaber-Fiebert. Cost effectiveness of fibrosis assessment prior to treatment for chronic hepatitis C patients. *PLoS One*, 6(12):e26783, 2011.

[101] S. Liu, D. Watcha, M. Holodniy, and J. D. Goldhaber-Fiebert. Sofosbuvir-based treatment regimens for chronic, genotype 1 hepatitis C virus infection in US incarcerated populations: A cost-effectiveness analysis. *Annals of Internal Medicine*, 161(8):546–553, 2014.

[102] I. M. Longini, A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D. A. Cummings, and M. E. Halloran. Containing pandemic influenza at the source. *Science*, 309(5737):1083–1087, 2005.

[103] K. Lovibond, S. Jowett, P. Barton, M. Caulfield, C. Heneghan, F. R. Hobbs, J. Hodgkinson, J. Mant, U. Martin, B. Williams, et al. Cost-effectiveness of options for the diagnosis of high blood pressure in primary care: A modelling study. *The Lancet*, 378(9798):1219–1230, 2011.

[104] B. S. Mann, L. Barnieh, K. Tang, D. J. T. Campbell, F. Clement, B. Hemmelgarn, M. Tonelli, D. Lorenzetti, and B. J. Manns. Association between Drug Insurance Cost Sharing Strategies and Outcomes in Patients with Chronic Diseases: A Systematic Review. *PLoS One*, 9(3):e89168, 2014.

[105] N. K. Martin, A. B. Pitcher, P. Vickerman, A. Vassall, and M. Hickman. Optimal control of hepatitis C antiviral treatment programme delivery for prevention amongst a population of injecting drug users. *PLoS One*, 6(8):e22309, 2011.

[106] J. E. Mason, D. A. England, B. T. Denton, S. A. Smith, M. Kurt, and N. D. Shah. Optimizing statin treatment decisions for diabetes patients in the presence of uncertain future adherence. *Medical Decision Making*, 32(1):154–166, 2012.

[107] G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems. *Mathematical Programming*, 10(1):147–175, 1976.

[108] R. V. Milani and C. J. Lavie. Health care 2020: Reengineering health care delivery to combat chronic disease. *The American Journal of Medicine*, 128(4):337–343, 2015.

[109] R. Mojtabai and M. Olfson. Medication costs, adherence, and health outcomes among medicare beneficiaries. *Health Affairs*, 22(4):220–229, 2003.

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

[111] A. E. Moran, M. C. Odden, A. Thanataveerat, K. Y. Tzong, P. W. Rasmussen, D. Guzman, L. Williams, K. Bibbins-Domingo, P. G. Coxson, and L. Goldman. Cost-effectiveness of hypertension therapy according to 2014 guidelines. *New England Journal of Medicine*, 372(5):447–455, 2015.

[112] C. V. Nguyen. The impact of voluntary health insurance on health care utilization and out-of-pocket payments: New evidence for vietnam. *Health Economics*, 21(8):946–966, 2012.

[113] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic game theory*, volume 1. Cambridge University Press Cambridge, 2007.

[114] Y. Ostchega, S. S. Yoon, J. Hughes, and T. Louis. Hypertension awareness, treatment, and control–continued disparities in adults: United states, 2005-2006. *NCHS Data Brief*, (3):1–8, 2008.

[115] M. V. Pauly and F. E. Blavin. Moral hazard in insurance, value-based cost sharing, and the benefits of blissful ignorance. *Journal of Health Economics*, 27(6):1407–1417, 2008.

[116] L. Perez and S. Dragicevic. An agent-based approach for modeling dynamics of contagious disease spread. *International Journal of Health Geographics*, 8(1):50, 2009.

[117] L. Perez and S. Dragicevic. An agent-based approach for modeling dynamics of contagious disease spread. *International Journal of Health Geographics*, 8(1):50, 2009.

[118] A. Perisic and C. T. Bauch. Social contact networks and disease eradicability under voluntary vaccination. *PLoS Computational Biology*, 5(2), 2009.

[119] Prospective Studies Collaboration. Age-specific relevance of usual blood pressure to vascular mortality: A meta-analysis of individual data for one million adults in 61 prospective studies. *The Lancet*, 360(9349):1903–1913, 2002.

[120] M. S. Rauner, W. J. Gutjahr, K. Heidenberger, J. Wagner, and J. Pasia. Dynamic policy modeling for chronic diseases: Metaheuristic-based identification of pareto-optimal screening strategies. *Operations Research*, 58(5):1269–1286, 2010.

[121] T. C. Reluga and A. P. Galvani. A general approach for population games with application to vaccination. *Mathematical Biosciences*, 230(2):67–78, 2011.

[122] M. Rothschild and J. Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. In *Foundations of Insurance Economics*, pages 355–375. Springer, 1976.

[123] S. W. Roush, L. McIntyre, and L. M. Baldy. Manual for the surveillance of vaccine-preventable diseases,. *Centers for Disease Control and Prevention, Atlanta (GA)*, (4), 2008.

[124] J. A. Salomon, M. C. Weinstein, J. K. Hammitt, and S. J. Goldie. Cost-effectiveness of treatment for chronic hepatitis C infection in an evolving patient population. *Journal of the American Medical Association*, 290(2):228–237, 2003.

[125] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts. Modeling medical treatment using Markov decision processes. In *Operations Research and Health Care*, pages 593–612. Springer, 2005.

[126] G. J. Schell, W. J. Marrero, M. S. Lavieri, J. B. Sussman, and R. A. Hayward. Data-driven markov decision process approximations for personalized hypertension treatment planning. *MDM Policy & Practice*, 1(1):2381468316674214, 2016.

[127] P. H. T. Schimit and L. H. A. Monteiro. A vaccination game based on public health actions and personal decisions. *Ecological Modelling*, 222(9):1651–1655, 2011.

[128] N. Secomandi. A rollout policy for the vehicle routing problem with stochastic demands. *Operations Research*, 49(5):796–802, 2001.

[129] N. Secomandi. Analysis of a rollout approach to sequencing problems with stochastic routing applications. *Journal of Heuristics*, 9(4):321–352, 2003.

[130] S. M. Shechter, M. D. Bailey, A. J. Schaefer, and M. S. Roberts. The optimal time to initiate HIV therapy under ordered health states. *Operations Research*, 56(1):20–33, 2008.

[131] W. H. Shrank, T. Hoang, S. L. Ettner, P. A. Glassman, K. Nair, D. DeLapp, J. Dirstine, J. Avorn, and S. M. Asch. The implications of choice: Prescribing generic or preferred pharmaceuticals improves medication adherence for chronic conditions. *Archives of Internal Medicine*, 166(3):332–337, 2006.

[132] W. H. Shrank, H. N. Young, S. L. Ettner, P. Glassman, S. M. Asch, and R. L. Kravitz. Do the incentives in 3-tier pharmaceutical benefit plans operate as intended? Results from a physician leadership survey. *The American journal of managed care*, 11(1):16–22, 2005.

[133] S.-J. Sinnott, C. Buckley, D. ORiordan, C. Bradley, and H. Whelton. The effect of copayments for prescriptions on adherence to prescription medicines in publicly insured populations: A systematic review and meta-analysis. *PLoS One*, 8(5):e64914, 2013.

[134] F. A. Sonnenberg and J. R. Beck. Markov models in medical decision making: A practical guide. *Medical Decision Making*, 13(4):322–338, 1993.

[135] A. C. Spaulding and D. L. Thomas. Screening for HCV infection in jails. *Journal of the American Medical Association*, 307(12):1259–1260, 2012.

[136] L. N. Steimle and B. T. Denton. Markov decision processes for screening and treatment of chronic diseases. In *Markov Decision Processes in Practice*, pages 189–222. Springer, 2017.

[137] M. Stein. Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29(2):143–151, 1987.

[138] M. Stepanova, F. Kanwal, H. B. El-Serag, and Z. M. Younossi. Insurance status and treatment candidacy of hepatitis c patients: Analysis of population-based data from the united states. *Hepatology*, 53(3):737–745, 2011.

[139] L. Stone, B. Shulgin, and Z. Agur. Theoretical examination of the pulse vaccination policy in the SIR epidemic model. *Mathematical and Computer Modelling*, 31(4-5):207–215, 2000.

[140] X. Su and S. Zenios. Patient choice in kidney allocation: The role of the queueing discipline. *Manufacturing & Service Operations Management*, 6(4):280–301, 2004.

[141] X. Su and S. A. Zenios. Patient choice in kidney allocation: A sequential stochastic assignment model. *Operations Research*, 53(3):443–455, 2005.

[142] P. Sun, L. Yang, and F. De Véricourt. Selfish drug allocation for containing an international influenza pandemic at the onset. *Operations Research*, 57(6):1320–1332, 2009.

[143] K. Swartz. Cost-sharing: Effects on spending and outcomes. *The Synthesis project. Research synthesis report*, (20):42–45, 2010.

[144] D. A. Taira, K. S. Wong, F. Frech-Tamas, and R. S. Chung. Copayment level and compliance with antihypertensive medication: Analysis and policy implications for managed care. *American Journal of Managed Care*, 12(11):678–684, 2006.

[145] C. P. Thomas. *How prescription drug use affects health care utilization and spending by older Americans: A review of the literature*. Public Policy Institute, 2008.

[146] K. Trudeau, K. Manser, M. Behling, C. Brown, W. Niebler, and S. Budman. Natural history of prescription opioid misuse/abuse: A qualitative investigation. *The Journal of Pain*, 18(4):S32–S33, 2017.

[147] T. Ugajin, A. Hozawa, T. Ohkubo, K. Asayama, M. Kikuya, T. Obara, H. Metoki, H. Hoshi, J. Hashimoto, K. Totsune, et al. White-coat hypertension as a risk factor for the development of home hypertension: The Ohasama study. *Archives of Internal Medicine*, 165(13):1541–1546, 2005.

[148] U.S. Census Bureau. 2010 Census Summary File, 2015. Available via `https://www.census.gov/prod/cen2010/doc/sf1.pdf`. [accessed April 1, 2018].

[149] US Census Bureau. Selected characteristics of people 15 years and over, by total money income, work experience, race, hispanic origin, and sex, 2017. Available via `http://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pinc/pinc-01.html/`. [accessed: 27 May 2017].

[150] USPSTF. U.S. preventive services task force: Screening for depression in adults. Available via `https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/depression-in-adults-screening1`. [accessed: April 1, 2018].

[151] VA. State of care for veterans with chronic hepatitis C, November 2010. Available via `http://www.hepatitis.va.gov/provider/policy/HCV-state-of-care-2010.asp`. [accessed: April 1, 2018].

[152] R. Vardavas, R. Breban, and S. Blower. Can influenza epidemics be prevented by voluntary vaccination? *PLoS Computational Biology*, 3(5):0796–0802, 2007.

[153] E. Vasudeva, N. Moise, C. Huang, A. Mason, J. Penko, L. Goldman, P. G. Coxson, K. Bibbins-Domingo, and A. E. Moran. Comparative cost-effectiveness of hypertension treatment in non-Hispanic Blacks and Whites according to 2014 guidelines: A modeling study. *American Journal of Hypertension*, 29(10):1195–1205, 2016.

[154] Y. Vorobeychik and S. P. Singh. Computing Stackelberg equilibria in discounted stochastic games. In *Processing of Twenty-Sixth National Conference on Artificial Intelligence*, 2012.

[155] E. H. Wagner, B. T. Austin, C. Davis, M. Hindmarsh, J. Schaefer, and A. Bonomi. Improving chronic illness care: translating evidence into action. *Health affairs*, 20(6):64–78, 2001.

[156] S. Wagner, T. S. Toftegaard, and O. W. Bertelsen. Challenges in blood pressure self-measurement. *International Journal of Telemedicine and Applications*, 2012:2, 2012.

[157] Z. Wang, Y. Moreno, S. Boccaletti, and M. Perc. Vaccination and epidemics in networked populations—an introduction, 2017.

[158] S. H. White, A. M. Del Rey, and G. R. Sánchez. Modeling epidemics using cellular automata. *Applied Mathematics and Computation*, 186(1):193–202, 2007.

[159] L. Willem, S. Stijven, E. Vladislavleva, J. Broeckhove, P. Beutels, and N. Hens. Active learning to understand infectious disease models and improve policy making. *PLoS computational biology*, 10(4):e1003563, 2014.

[160] I. B. Wilson, W. H. Rogers, H. Chang, and D. G. Safran. Cost-related skipping of medications and other treatments among medicare beneficiaries between 1998 and 2000. *Journal of General Internal Medicine*, 20(8):715–720, 2005.

[161] W. L. Winston. *Introduction to probability models: Operations research. Volume two*. Brooks/Cole-Thomson Learning, 2003.

[162] M. D. Wong, R. Andersen, C. D. Sherbourne, R. D. Hays, and M. F. Shapiro. Effects of cost sharing on care seeking and health status: Results from the medical outcomes study. *American Journal of Public Health*, 91(11):1889–1894, 2001.

[163] B. Xie, D. M. Dilts, and M. Shor. The physician–patient relationship: The impact of patient-obtained medical information. *Health Economics*, 15(8):813–833, 2006.

[164] D. Yamin and A. Gavious. Incentives' effect in influenza vaccination policy. *Management Science*, 59(12):2667–2686, 2013.

[165] Z. Zabinsky and H. Huang. *A partition-based optimization approach for level set approximation: Probabilistic branch and bound*. Springer Nature, 2019. (forthcoming 2019).

[166] S. A. Zenios, G. M. Chertow, and L. M. Wein. Dynamic allocation of kidneys to candidates on the transplant waiting list. *Operations Research*, 48(4):549–569, 2000.

[167] Z. Zhu. *Optimal financing structure mechanism design for healthcare insurance*. PhD thesis, Purdue University, 2012.

[168] P. Zweifel and W. G. Manning. Moral hazard and consumer incentives in health care. In *Handbook of Health Economics*, volume 1, pages 409–459. Elsevier, 2000.

Appendix A

COMPUTATIONAL COMPLEXITY AND PROPERTIES OF MF-RA

A.0.1 Computational Complexity of MF-RA

There are two primary sources of computation in MF-RA: the progression of the population as in (3.2), and the computation of LFM in (3.7) involving a matrix inverse.

Evaluating a complete sequential policy for the whole lifetime horizon requires n_q population computations, i.e., population is evolved n_q times using Equation (3.2). An exhaustive search with n_p discretized policies, i.e., the grid search approach, would evaluate $(n_p)^{n_d}$ possible sequential policies, in total

$$n_q(n_p)^{n_d} \tag{A.1}$$

population computations, which is exponential in the number of decision periods n_d.

In contrast, MF-RA requires far fewer population computations. In each decision period $d \in \{1,\ldots,n_d\}$, there is a total of n_p discretized policies that need to be evaluated. For one discretized policy at decision d, we need $(n_q-(d-1)n_{q,dec})$ population computations to calculate QALYs from that time forward. Also, each discretized policy requires $n_p n_{q,dec}\,(n_d-d)$ population computations in low-fidelity approximation (Step 2 in MF-RA) to iteratively calculate LFM. Note that in the last decision period n_d, MF-RA does not need to compute LFM, so only population computations from n_d to n_q are needed.

As a result, the overall number of population computations in MF-RA is

$$\begin{aligned} & \sum_{d=1}^{n_d} n_p \left\{ \left[n_q-(d-1)\,n_{q,dec} \right] + n_p n_{q,dec}\,(n_d-d) \right\} \\ \leq\ & n_p n_d n_q + n_p^2 n_d^2 n_{q,dec}. \end{aligned} \tag{A.2}$$

In addition, the low-fidelity approximation conducts a matrix inversion, i.e., $(\mathbf{I}-\mathbf{Q}(I_t))^{-1}$,

$$\sum_{d=1}^{n_d} n_p^2 n_{q,dec}\,(n_d-d) \leq n_p^2 n_d^2 n_{q,dec} \tag{A.3}$$

times. Using Gauss-Jordan elimination, the computational complexity for each matrix inversion is $O(|\mathscr{S}_{HS}|^3)$; however, the matrix is very sparse.

To summarize, the number of population computations of MF-RA is quadratic in the number of decision periods n_d and the number of matrix inversions is quadratic in n_d.

A.0.2 Properties of MF-RA

We characterize the structure of the overall QALYs in budget planning years with respect to the general form of a rollout algorithm. We show the proposed MF-RA has the property of sequential QALYs improvement, which indicates that, using the low-fidelity approximation, the policy obtained in the decision period $d+1$ is no worse than the performance of the policies obtained in the decision period d.

A rollout algorithm is an iterative method that uses a heuristic to iteratively improve the objective function value of combinatorial optimization problems. Unlike dynamic programming, which calculates a recursive value function with boundary conditions, a rollout algorithm approximates the optimal value function by a designed heuristic approach [17]. To optimize healthcare interventions, we need more effective sequential actions. Thus, different from other applications of rollout algorithms by only adopting the rollout process once, e.g., stochastic scheduling [18], stochastic routing [128, 129], and multidimensional knapsack problem [21], MF-RA adopts two rollout processes by incorporating low-fidelity approximation into the high-fidelity model. The proposed MF-RA, combining discretized policies enumeration and low-fidelity approximation, satisfies the sequential QALYs improvement.

Proposition 1. ***Sequential QALYs Improvement of MF-RA***. *MF-RA improves overall QALYs in every decision period.*

Proof of Proposition 1

To prove MF-RA improves overall QALYs over time, we need to use two useful properties

$$\max_{k\in\{1,\ldots,n_p\}} QALY\left(\mathbf{N}_1, \left\{I^{(t)*}\right\}_{t\in\mathscr{T}_1\cup\cdots\cup\mathscr{T}_{d-1}}, \left\{I_k^{(t)}\right\}_{t\in\mathscr{T}_d}, \left\{I_{Ak}^{(t)}\right\}_{t\in\mathscr{T}_{d+1}\cup\cdots\cup\mathscr{T}_{n_d}}\right) \tag{A.4}$$

$$\max_{k\in\{1,\ldots,n_p\}} QALY\left(\mathbf{N}_1, \left\{I^{(t)*}\right\}_{t\in\mathscr{T}_1\cup\cdots\cup\mathscr{T}_{d}}, \left\{I_k^{(t)}\right\}_{t\in\mathscr{T}_{d+1}}, \left\{I_{Ak}^{(t)}\right\}_{t\in\mathscr{T}_{d+2}\cup\cdots\cup\mathscr{T}_{n_d}}\right) \tag{A.5}$$

$$\begin{aligned}
\mathbf{N}_{1+dn_{q,dec}} &= \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_d} \mathbf{P}\left(I_{Ak}^{(t)}\right) && \text{(A.6)}\\
\mathbf{N}_{1+(d+1)n_{q,dec}} &= \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_d\cup\mathscr{T}_{d+1}} \mathbf{P}\left(I_{Ak}^{(t)}\right) && \text{(A.7)}\\
&\vdots\\
\mathbf{N}_{1+n_d n_{q,dec}} &= \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_d\cup\mathscr{T}_{d+1}\cup\cdots\cup\mathscr{T}_{n_d}} \mathbf{P}\left(I_{Ak}^{(t)}\right) && \text{(A.8)}
\end{aligned}$$

developed in rollout algorithms: sequential consistency and sequential improvement. In our population health outcome maximization problem, sequential consistency means that, given current population size, MF-RA is able to generate the same population evolution afterward regardless of earlier population size. That is, given $\mathbf{N}_{1+(d-1)n_{q,dec}}$ at decision period d, if MF-RA generates the sequential policies $\left\{I^{(t)*}\right\}_{t\in\mathscr{T}_d\cup\mathscr{T}_{d+1}\cup\cdots\mathscr{T}_{n_d}}$ and associated population evolution $\{\mathbf{N}_t\}_{t\in\mathscr{T}_d\cup\mathscr{T}_{d+1}\cup\cdots\mathscr{T}_{n_d}}$, then given $\mathbf{N}_{1+(d)n_{q,dec}}$ at decision period $d+1$, MF-RA will generate the sequential policies $\left\{I^{(t)*}\right\}_{t\in\mathscr{T}_{d+1}\cup\cdots\cup\mathscr{T}_{n_d}}$ and associated population evolution $\{\mathbf{N}_t\}_{t\in\mathscr{T}_{d+1}\cup\cdots\cup\mathscr{T}_{n_d}}$. Sequential improvement in MF-RA means that the overall QALYs improves over time. That is, maximum QALYs calculated at decision period d (Equation (A.4)) is no better than maximum QALYs calculated at decision period $d+1$ (Equation (A.5)). Note that the low-fidelity approximation (step 2 in MF-RA) is also a rollout algorithm and holds these two properties.

We prove in (a) that the low-fidelity approximation satisfies sequential consistency, as well as MF-RA in (b). Then, according to the theorem in [17], the sequential consistency of MF-RA allows us to claim its sequential improvement.

(a) Sequential consistency of low-fidelity approximation

Proof of Proposition 1.

$$\begin{aligned}
\mathbf{N}_{1+(d+1)n_{q,dec}} &= \mathbf{N}_{1+dn_{q,dec}} \prod_{t\in\mathscr{T}_{d+1}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \\
&= \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_{d}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \prod_{t\in\mathscr{T}_{d+1}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \\
&= \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_{d}\cup\mathscr{T}_{d+1}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \qquad \text{(A.9)} \\
&\vdots \\
\mathbf{N}_{1+n_d n_{q,dec}} &= \mathbf{N}_{1+dn_{q,dec}} \prod_{t\in\mathscr{T}_{d+1}\cup\mathscr{T}_{d+2}\cup\cdots\cup\mathscr{T}_{n_d}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \\
&= \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_{d}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \prod_{t\in\mathscr{T}_{d+1}\cup\mathscr{T}_{d+2}\cup\cdots\cup\mathscr{T}_{n_d}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \\
&= \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_{d}\cup\mathscr{T}_{d+1}\cup\cdots\cup\mathscr{T}_{n_d}} \mathbf{P}\left(I_{Ak}^{(t)}\right) \qquad \text{(A.10)}
\end{aligned}$$

Starting from $\mathbf{N}_{1+(d-1)n_{q,dec}}$, maximizing *LFM* iteratively generates sequential policies

$$\left\{I_{Ak}^{(t)}\right\}_{t\in\mathscr{T}_{d}\cup\mathscr{T}_{d+1}\cup\cdots\cup\mathscr{T}_{n_d}}$$

and the corresponding sequence of population size

$$\left\{\mathbf{N}_{1+(d-1)n_{q,dec}},\ldots,\mathbf{N}_{1+(n_d)n_{q,dec}}\right\}$$

(shown in Equations (A.6)-(A.8)). Due to the fact that

$$\mathbf{N}_{1+dn_{q,dec}} = \mathbf{N}_{1+(d-1)n_{q,dec}} \prod_{t\in\mathscr{T}_{d}} \mathbf{P}\left(\mathbf{I}_{Ak}^{(t)}\right)$$

if starting from $\mathbf{N}_{1+dn_{q,dec}}$, then $\mathbf{N}_{1+(d+1)n_{q,dec}}\ldots\mathbf{N}_{1+(d-1)n_{q,dec}}$ is calculated by Equations (A.9) and (A.10).

Therefore, whenever low-fidelity approximation generates the sequence of population size

$$\left\{\mathbf{N}_{1+(d-1)n_{q,dec}},\mathbf{N}_{1+dn_{q,dec}},\ldots,\mathbf{N}_{1+n_d n_{q,dec}}\right\}$$

starting from $\mathbf{N}_{1+(d-1)n_{q,dec}}$, it also generates the sequence of population evolution

$$\left\{\mathbf{N}_{1+dn_{q,dec}}, \mathbf{N}_{1+(d+1)n_{q,dec}}, \ldots, \mathbf{N}_{1+n_d n_{q,dec}}\right\}$$

starting from $\mathbf{N}_{1+dn_{q,dec}}$. Thus, the low-fidelity approach is sequentially consistent, which helps us to prove that MF-RA is also sequentially consistent.

(b) Sequential consistency of MF-RA

With the same population size $\mathbf{N}_{1+(d-1)n_{q,dec}}$ and the budget constraints, MF-RA enumerates the discretized intervention policies in decision period d in the same discretized rule. In addition, due to the sequential consistency of the low-fidelity approximation, MF-RA examines the same promising sequential policies and thus derives the same best policy $I^{(t)*}$, $t \in \mathscr{T}_d$ and proceeds to the next population size $\mathbf{N}_{1+dn_{q,dec}}$. Therefore, MF-RA generates the same sequence of population evolution $\left\{\mathbf{N}_{1+(d-1)n_q d}, \mathbf{N}_{1+dn_{q,dec}}, \ldots, \mathbf{N}_{1+(n_d-1)n_{q,dec}}\right\}$ regardless of the population size before decision period d. That is, starting from any population size in the sequential population size generated by MF-RA, MF-RA is able to generate the same sequential population size afterward.

(c) Sequential improvement of MF-RA

By applying the theorem in [17], we can claim that MF-RA satisfies sequential improvement, and since MF-RA determined the best policy by calculating and comparing accumulating overall QALYs, it is capable of improving overall QALYs when rolling out to the next decision period.

Appendix B

DETAIL OF THE TRANSITION PROBABILITY MATRIX FOR HCV PROGRESSION

We provide the details of the transition matrix for HCV progression. The transition probabilities from group A to other groups are functions of the decision variables

$$I_{A,t} = \left\{ I_{A,t}^{H_A}, I_{A,t}^{F0_A}, I_{A,t}^{F1_A}, I_{A,t}^{F2_A}, I_{A,t}^{F3_A}, I_{A,t}^{F4_A} \right\}.$$

The transition probabilities from group B to other groups are functions of the decision variables

$$I_{B,t} = \left\{ I_{B,t}^{F0_B}, I_{B,t}^{F1_B}, I_{B,t}^{F2_B}, I_{B,t}^{F3_B}, I_{B,t}^{F4_B} \right\}.$$

We denote $\theta_{a,r,waiting}^{ij}$ as the fraction of patient population of age $a \in \{40-49, 50-59, 60-69\}$ and gender $r \in \{female, male\}$ who are not going through treatment and transit from health status i to health status j, and $\theta_{a,r,ongoing}^{ij}$ as the fraction of patient population of age a and gender r who are going through treatment and transit from health status i to health status j, $i, j \in \{H,\ F0\text{-}F4,\ R1\text{-}R3,\ DC,\ HCC,\ LT,\ ALT,\ M\}$. The complete transition probability matrix is in Equation (B.1), where $\mathbf{P}_{a,female}(I_{A,t}, I_{B,t})$ and $\mathbf{P}_{a,male}(I_{A,t}, I_{B,t})$ are 20×20 sub-matrices of transition probabilities for female and male, respectively, $a \in \{40-49, 50-59, 60-69\}$. The transition probabilities are summarized in Fig. B.1 and Tables B.1-B.3.

$$\mathbf{P}(I_{A,t}, I_{B,t}) = \left[\begin{array}{c|c} \mathbf{P}_{a,female}(I_{A,t}, I_{B,t}) & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{P}_{a,male}(I_{A,t}, I_{B,t}) \end{array} \right] \tag{B.1}$$

Table B.1: List of transition probabilities from Group A to other groups

Transition probabilties from group-based health category $i \in \mathscr{S}_A$ to other state $j \in \mathscr{S}_{HS}$		
$A_1 = \theta^{HH}_{a,r,waiting}(1-I_{A,t})$	$A_{12} = \theta^{HF0}_{a,r,waiting}I_{A,t}$	$A_{22} = \theta^{HH}_{a,r,waiting}I_{A,t}$
$A_2 = \theta^{HF0}_{a,r,waiting}(1-I_{A,t})$	$A_{13} = \theta^{F0F0}_{a,r,waiting}I_{A,t}$	$A_{23} = \theta^{F4DC}_{a,r,waiting}$
$A_3 = \theta^{F0F0}_{a,r,waiting}(1-I_{A,t})$	$A_{14} = \theta^{F0F1}_{a,r,waiting}I_{A,t}$	$A_{24} = \theta^{F4HCC}_{a,r,waiting}$
$A_4 = \theta^{F0F1}_{a,r,waiting}(1-I_{A,t})$	$A_{15} = \theta^{F1F1}_{a,r,waiting}I_{A,t}$	$A_{25} = \theta^{F4LT}_{a,r,waiting}$
$A_5 = \theta^{F1F1}_{a,r,waiting}(1-I_{A,t})$	$A_{16} = \theta^{F1F2}_{a,r,waiting}I_{A,t}$	$A_{26} = \theta^{HM}_{a,r,waiting}$
$A_6 = \theta^{F1F2}_{a,r,waiting}(1-I_{A,t})$	$A_{17} = \theta^{F2F2}_{a,r,waiting}I_{A,t}$	$A_{27} = \theta^{F0M}_{a,r,waiting}$
$A_7 = \theta^{F2F2}_{a,r,waiting}(1-I_{A,t})$	$A_{18} = \theta^{F2F3}_{a,r,waiting}I_{A,t}$	$A_{28} = \theta^{F1M}_{a,r,waiting}$
$A_8 = \theta^{F2F3}_{a,r,waiting}(1-I_{A,t})$	$A_{19} = \theta^{F3F3}_{a,r,waiting}I_{A,t}$	$A_{29} = \theta^{F2M}_{a,r,waiting}$
$A_9 = \theta^{F3F3}_{a,r,waiting}(1-I_{A,t})$	$A_{20} = \theta^{F3F4}_{a,r,waiting}I_{A,t}$	$A_{30} = \theta^{F3M}_{a,r,waiting}$
$A_{10} = \theta^{F3F4}_{a,r,waiting}(1-I_{A,t})$	$A_{21} = \theta^{F4F4}_{a,r,waiting}I_{A,t}$	$A_{31} = \theta^{F4M}_{a,r,waiting}$
$A_{11} = \theta^{F4F4}_{a,r,waiting}(1-I_{A,t})$		

$$\mathbf{P}_{a,r}\left(I_{A,t}, I_{B,t}\right) =$$

	A						B						C				D				M
	A_1	A_2					A_{12}						A_{22}								A_{26}
		A_3	A_4				A_{13}	A_{14}													A_{27}
A			A_5	A_6				A_{15}	A_{16}												A_{28}
				A_7	A_8				A_{17}	A_{18}											A_{29}
					A_9	A_{10}				A_{19}	A_{20}										A_{30}
						A_{11}					A_{21}						A_{23}	A_{24}	A_{25}		A_{31}
							B_1	B_2					B_{10}	B_{11}							B_{19}
								B_3	B_4					B_{12}							B_{20}
B									B_5	B_6					B_{13}						B_{21}
										B_7	B_8				B_{14}						B_{22}
											B_9					B_{15}	B_{16}	B_{17}	B_{18}		B_{23}
		C_1											C_5								C_9
C		C_2												C_6							C_{10}
				C_3											C_7						C_{11}
						C_4										C_8					C_{12}
																	D_1	D_2	D_3		D_8
D																		D_4	D_5		D_9
																				D_6	D_{10}
																				D_7	D_{11}
M																					1

Figure B.1: Transition probability sub-matrix for $a \in \{40-49, 50-59, 60-69\}$ and $r \in \{female, male\}$

Table B.2: List of transition probabilities from Group B to other groups

Transition probabilties from group-based health category $i \in \mathscr{S}_B$ to other state $j \in \mathscr{S}_{HS}$
$B_1 = \theta^{F0F0}_{a,r,waiting}(1-I^{F0_B}_{B,t}) + \theta^{F0F0}_{a,r,ongoing} I^{F0_B}_{B,t}$
$B_2 = \theta^{F0F1}_{a,r,waiting}(1-I^{F0_B}_{B,t}) + \theta^{F0F1}_{a,r,ongoing} I^{F0_B}_{B,t}$
$B_3 = \theta^{F1F1}_{a,r,waiting}(1-I^{F0_B}_{B,t}) + \theta^{F1F1}_{a,r,ongoing} I^{F0_B}_{B,t}$
$B_4 = \theta^{F1F2}_{a,r,waiting}(1-I^{F1_B}_{B,t}) + \theta^{F1F2}_{a,r,ongoing} I^{F1_B}_{B,t}$
$B_5 = \theta^{F2F2}_{a,r,waiting}(1-I^{F2_B}_{B,t}) + \theta^{F2F2}_{a,r,ongoing} I^{F2_B}_{B,t}$
$B_6 = \theta^{F2F3}_{a,r,waiting}(1-I^{F2_B}_{B,t}) + \theta^{F2F3}_{a,r,ongoing} I^{F2_B}_{B,t}$
$B_7 = \theta^{F3F3}_{a,r,waiting}(1-I^{F3_B}_{B,t}) + \theta^{F3F3}_{a,r,ongoing} I^{F3_B}_{B,t}$
$B_8 = \theta^{F3F4}_{a,r,waiting}(1-I^{F3_B}_{B,t}) + \theta^{F3F4}_{a,r,ongoing} I^{F3_B}_{B,t}$
$B_9 = \theta^{F4F4}_{a,r,waiting}(1-I^{F4_B}_{B,t}) + \theta^{F4F4}_{a,r,ongoing} I^{F4_B}_{B,t}$
$B_{10} = \theta^{F0H}_{a,r,waiting}(1-I^{F0_B}_{B,t})$
$B_{11} = \theta^{F0R1}_{a,r,ongoing} I^{F0_B}_{B,t}$
$B_{12} = \theta^{F1R1}_{a,r,ongoing} I^{F1_B}_{B,t}$
$B_{13} = \theta^{F2R2}_{a,r,ongoing} I^{F2_B}_{B,t}$
$B_{14} = \theta^{F3R2}_{a,r,ongoing} I^{F3_B}_{B,t}$
$B_{15} = \theta^{F4R3}_{a,r,ongoing} I^{F4_B}_{B,t}$
$B_{16} = \theta^{F4DC}_{a,r,waiting}(1-I^{F4_B}_{B,t}) + \theta^{F4DC}_{a,r,ongoing} I^{F4_B}_{B,t}$
$B_{17} = \theta^{F4HCC}_{a,r,waiting}(1-I^{F4_B}_{B,t}) + \theta^{F4HCC}_{a,r,ongoing} I^{F4_B}_{B,t}$
$B_{18} = \theta^{F4LT}_{a,r,waiting}(1-I^{F4_B}_{B,t}) + \theta^{F4LT}_{a,r,ongoing} I^{F4_B}_{B,t}$
$B_{19} = \theta^{F0M}_{a,r,waiting}$
$B_{20} = \theta^{F1M}_{a,r,waiting}$
$B_{21} = \theta^{F2M}_{a,r,waiting}$
$B_{22} = \theta^{F3M}_{a,r,waiting}$
$B_{23} = \theta^{F4M}_{a,r,waiting}$

Table B.3: List of transition probabilities from Groups C and D to other groups

Transition probabilities from group-based health state $i \in \mathscr{S}_C \cup \mathscr{S}_D$ to other state $j \in \mathscr{S}_{HS}$	
$C_1 = \theta^{HF0}_{a,r,waiting}$	$D_1 = \theta^{DCDC}_{a,r,waiting}$
$C_2 = \theta^{R1F0}_{a,r,waiting}$	$D_2 = \theta^{DCHCC}_{a,r,waiting}$
$C_3 = \theta^{R2F2}_{a,r,waiting}$	$D_3 = \theta^{DCLT}_{a,r,waiting}$
$C_4 = \theta^{R3F4}_{a,r,waiting}$	$D_4 = \theta^{HCCHCC}_{a,r,waiting}$
$C_5 = \theta^{HH}_{a,r,waiting}$	$D_5 = \theta^{HCCLT}_{a,r,waiting}$
$C_6 = \theta^{R1R1}_{a,r,waiting}$	$D_6 = \theta^{LTALT}_{a,r,waiting}$
$C_7 = \theta^{R2R2}_{a,r,waiting}$	$D_7 = \theta^{ALTALT}_{a,r,waiting}$
$C_8 = \theta^{R3R3}_{a,r,waiting}$	$D_8 = \theta^{DCM}_{a,r,waiting}$
$C_9 = \theta^{HM}_{a,r,waiting}$	$D_9 = \theta^{HCCM}_{a,r,waiting}$
$C_{10} = \theta^{R1M}_{a,r,waiting}$	$D_{10} = \theta^{LTM}_{a,r,waiting}$
$C_{11} = \theta^{R2M}_{a,r,waiting}$	$D_{11} = \theta^{ALTM}_{a,r,waiting}$
$C_{12} = \theta^{R3M}_{a,r,waiting}$	

Appendix C

DETAILS ON INPUT DATA

Table C.1: Model parameters values

Variable	Base Case	Reference
Annual discount rate	0.03	[96]
HCV Natural History (annual probability)		[98, 101]
Proportion of patients with no fibrosis (F0) who do not progress	0.24	
Spontaneous remission from no fibrosis (F0) health category	0.012	
Fibrosis progression		
Males		
Age 40-49 y	0.05	
Age 50-59 y	0.12	
Age 60-69 y	0.2	
Age $\geq$ 70 y	0.26	
Females		
Age 40-49 y	0.03	
Age 50-59 y	0.06	
Age 60-69 y	0.11	
Age 70-79 y	0.14	
Age $\geq$ 80 y	0.2	
Cirrhosis to decompensated cirrhosis	0.04	
Cirrhosis (both F4 and decompensated cirrhosis) to HCC	0.02	
Liver transplant		
Decompensated cirrhosis to liver transplant	0.05	
HCC to liver transplant	0.15	
Liver related mortality		
Liver transplant	0.14	
After liver transplant	0.5	
Decompensated cirrhosis	0.26	
HCC	0.72	
Reinfection	0.0032	
Mortality Rate Ratio		[51]
Reduction factor on background mortality after successful treatment	0.9	
Mortality Hazard ratio, non-liver related	3.12	

Table C.1: Model parameters values (continued)

Variable	Base Case	Reference
Annual Quality of Life		[96]
HCV mild fibrosis (F0, F1)	0.98	
SVR after mild fibrosis (R1)	1.00	
HCV moderate fibrosis (F2, F3)	0.85	
SVR after moderate fibrosis (R2, R3)	0.93	
Compensated cirrhosis (F4)	0.79	
Decompensated cirrhosis, HCC	0.72	
Liver transplant, post-liver transplant	0.83	
Healthy	1.00	
Dead	0	
Screening and Treatment (USD$)		[96]
Quarterly screening cost		
HCV anti-body screening (ELISA)	$20	
Diagnosis (2 confirmatory ELISA, RIBA, and RNA test)	$210	
HCV genotyping	$369	
Liver biopsy	$1,340	
FibroTest	$240	
Treatment cost		
F0-F3	$40,320	
F4	$60,480	
Treatment effectiveness		
genotype 1&4	0.97	
genotype 2	0.97	
genotype 3	0.94	
Genotype prevalence		
genotype 1&4	0.69	
genotype 2	0.12	
genotype 3	0.19	

All the values are converted into quarterly probabilities from original annual data when applying MF-RA.

Table C.2: Cohort characteristics

Cohort Characteristics ([96])					
Number of infected people			3,100,000		
Fraction treatment eligible			100%		
U.S. Population ([148])					
Age	Male	Female			
40-49	21,603,062	21,996,493			
50-59	20,457,922	21,506,008			
50-59	13,929,977	15,323,140			
Proportion of HCV-infected Individuals in the U.S. Population ([31])					
Age	White Male	White Female	Black Male	Black Female	
40	2.79%	2.04%	0.21%	0.36%	
45	6.24%	4.60%	0.74%	0.70%	
50	9.92%	7.16%	1.66%	1.44%	
55	14.28%	6.33%	2.59%	2.39%	
60	12.78%	4.61%	2.90%	1.92%	
65	5.02%	1.71%	1.90%	0.84%	
Initial Fibrosis Stage Distribution ([45, 138, 92, 89])					
White male					
Age	F0	F1	F2	F3	F4
40	33.0%	34.6%	17.8%	12.4%	2.2%
45	27.0%	33.4%	22.8%	12.8%	4.1%
50	20.5%	28.8%	27.1%	15.8%	7.8%
55	15.3%	21.6%	28.3%	20.3%	14.5%
60	12.4%	13.6%	24.2%	23.1%	26.7%
65	11.8%	6.9%	16.9%	21.5%	42.9%
White female					
Age	F0	F1	F2	F3	F4
40	33.0%	34.6%	17.8%	12.4%	2.2%
45	29.6%	34.9%	21.1%	11.7%	2.7%
50	25.1%	33.5%	24.9%	12.5%	4.0%
55	20.4%	30.1%	28.4%	15.0%	6.2%
60	16.2%	24.5%	29.5%	18.9%	10.8%
65	13.2%	18.2%	28.2%	22.4%	17.9%
Black male					
Age	F0	F1	F2	F3	F4
40	33.0%	34.6%	17.8%	12.4%	2.2%
45	27.0%	33.4%	22.8%	12.8%	4.1%
50	20.5%	28.8%	27.1%	15.8%	7.8%
55	15.4%	21.6%	28.3%	20.3%	14.5%
60	12.4%	13.6%	24.3%	23.1%	26.6%
65	11.8%	7%	16.9%	21.5%	42.8%
Black female					
Age	F0	F1	F2	F3	F4
40	33.0%	34.6%	17.8%	12.4%	2.2%
45	29.6%	34.9%	21.1%	11.7%	2.7%
50	25.1%	33.5%	24.9%	12.5%	4.0%
55	20/5%	30.1%	28.4%	14.9%	6.2%
60	16.2%	24.5%	29.5%	18.9%	10.8%
65	13.2%	18.2%	28.3%	22.4%	17.9%

Appendix D

SUPPLEMENTARY FIGURES AND TABLES FOR EXPERIMENTS

Table D.1: Experiment 1: Incremented QALYs over time (MF-RA: multi-fidelity rollout algorithm)

	1 billion/year		**4 billion/year**		**8 billion/year**	
	Time	**Incremental QALYs**	**Time**	**Incremental QALYs**	**Time**	**Incremental QALYs**
40-49 birth-cohort						
	0.70	146,879	0.57	440,735	0.53	598,791
	0.87	215,113	0.78	443,066	0.70	598,791
MF-RA	1.02	263,002	0.95	443,066	0.88	598,791
	1.15	287,883	1.08	443,066	1.01	598,791
	1.26	298,442	1.21	445,565	1.12	598,892
50-59 birth-cohort						
	0.52	215,585	0.64	476,211	0.60	892,971
	0.66	216,605	0.78	622,134	0.75	906,453
MF-RA	0.78	229,603	0.91	649,668	0.87	906,453
	0.89	258,120	1.05	650,555	0.97	907,062
	0.97	281,596	1.17	651,112	1.06	907,062
60-69 birth-cohort						
	0.63	132,022	0.69	319,901	0.70	496,523
	0.75	132,022	0.84	354,205	0.82	507,516
MF-RA	0.85	132,610	0.97	357,715	0.93	507,516
	0.94	132,610	1.05	357,715	1.01	507,516
	1.00	138,173	1.12	357,715	1.08	507,516

Table D.2: Experiment 3: Dollars (in billions) spent every two years (every decision period) for 50-59 birth-cohort

	Years									
Annual budget	1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20
1	2	2	2	2	2	2	2	2	2	2
4	8	8	8	8	8	8	8	8	6	0.014
8	16	16	16	16	16	6.56	0.035	0.026	0.019	0.015

Note 1: The total budgets that are used to control HCV over 20 years in the cases of $1, $4, and $8 billion/year are about $20, $70, and $87 billion, respectively.

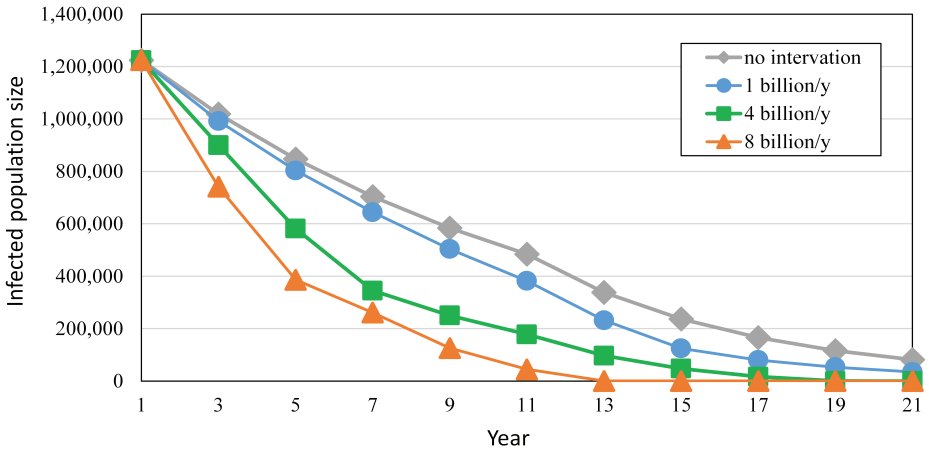


Figure D.1: Experiment 3: HCV control over 20 years for 50-59 birth-cohort

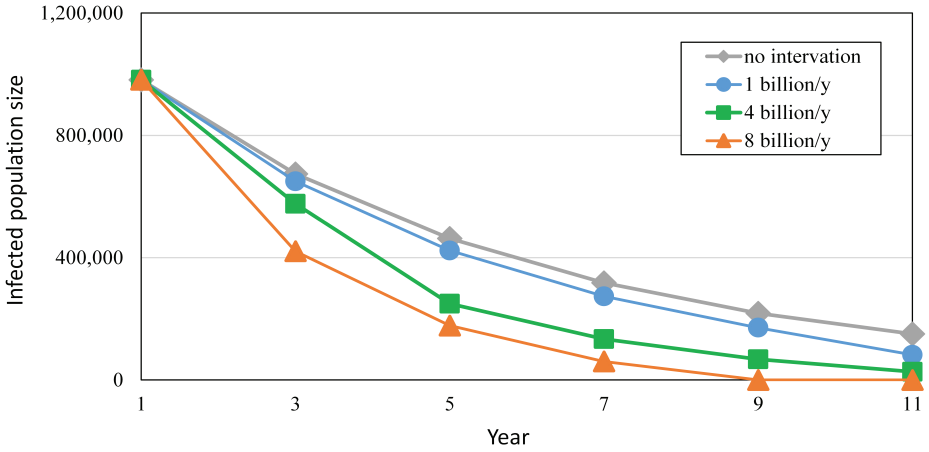


Figure D.2: Experiment 3: HCV control over 10 years for 60-69 birth-cohort

Table D.3: Experiment 3: Dollars (in billions) spent every two years (every decision period) for 60-69 birth-cohort

	Years				
Annual budget	1-2	3-4	5-6	7-8	9-10
1	2	2	2	2	2
4	8	8	8	8	8
8	16	16	16	11.8	0.025

Note 1: The total budgets that are used to control HCV over 10 years in the cases of $1, $4, and $8 billion/year are about $10, $40, and $60 billion, respectively.

Appendix E

MARKOV STATIONARY STACKELBERG EQUILIBRIUM APPROXIMATION

To approximately solve the MINLP in (4.5)−(4.11), $\pi(a_l|s)$ is discretized as in [66]. Let p_k be the kth probability value and let $\mathscr{K} = \{1,\ldots,K\}$ be the index set of discrete values. Define $d_{s,k}^{a_l}$ as binary variables such that $d_{s,k}^{a_l} = 1$ if and only if $\pi(a_l|s) = p_k$, and 0 otherwise. Thus, $\pi(a_l|s)$ can be approximated by $\pi(a_l|s) \approx \sum_{k\in K} p_k d_{s,k}^{a_l}$, $\forall s \in \mathscr{S}$ and $\forall a_l \in \mathscr{A}_L$. Constraints (4.6) and (4.7) become $\sum_{k\in K} d_{s,k}^{a_l} = 1$, $\forall s \in \mathscr{S}, a_l \in \mathscr{A}_L$, and $\sum_{a_l\in A_L}\sum_{k\in K} p_k d_{s,k}^{a_l} = 1$, $\forall s \in S$, respectively. Let $x_{s,k}^{a_l a_f} = d_{s,k}^{a_l}\sum_{s'\in S} T_{ss'}^{a_f} v_L\left(s'\right)$ and $y_{s,k}^{a_l a_f} = d_{s,k}^{a_l}\sum_{s'\in S} T_{ss'}^{a_f} v_F\left(s'\right)$ for the leader and follower, respectively. The leader's and the follower's expected costs, $R_F(s,a_f)$ and $R_L(s,a_f)$, are redefined as follows:

$$R_L(s,a_f) = \sum_{a_l\in A_L}\sum_{k\in\mathscr{K}} p_k\left(r_L(s,a_l,a_f)\, d_{s,k}^{a_l} + \gamma_L x_{s,k}^{a_l a_f}\right), \tag{E.1}$$

and

$$R_F(s,a_f) = \sum_{a_l\in A_L}\sum_{k\in\mathscr{K}} p_k\left(r_F(s,a_l,a_f)\, d_{s,k}^{a_l} + \gamma_F y_{s,k}^{a_l a_f}\right). \tag{E.2}$$

The mixed integer linear program (MILP) then becomes

$$\min_{\phi,v_L,v_F,x,y,d} \sum_{s\in S} \beta(s)\, v_L(s) \tag{E.3}$$

s.t.

$$d_{s,k}^{a_l} \in \{0,1\} \geq 0 \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L, k \in \mathscr{K} \tag{E.4}$$

$$\sum_{k \in \mathscr{K}} d_{s,k}^{a_l} = 1 \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L \tag{E.5}$$

$$\sum_{a_l \in A_L} \sum_{k \in \mathscr{K}} p_k d_{s,k}^{a_l} = 1 \qquad \forall s \in S \tag{E.6}$$

$$\phi\left(a_f|s\right) \in \{0,1\} \qquad \forall s \in \mathscr{S}, a_f \in \mathscr{A}_F \tag{E.7}$$

$$\sum_{a_f} \phi\left(a_f|s\right) = 1 \qquad \forall s \in \mathscr{S} \tag{E.8}$$

$$0 \leq R_F\left(s,a_f\right) - v_F\left(s\right) \leq \left(1-\phi\left(a_f|s\right)\right) Z \qquad \forall s \in \mathscr{S}, a_f \in \mathscr{A}_F \tag{E.9}$$

$$R_L\left(s,a_f\right) - v_L\left(s\right) \leq \left(1-\phi\left(a_f|s\right)\right) Z \qquad \forall s \in \mathscr{S}, a_f \in \mathscr{A}_F \tag{E.10}$$

$$x_{s,k}^{a_l a_f} \geq \sum_{s' \in S} T_{ss'}^{a_f} v_L\left(s'\right) - Z(1-d_{s,k}^{a_l}) \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L, a_f \in \mathscr{A}_F, k \in \mathscr{K} \tag{E.11}$$

$$x_{s,k}^{a_l a_f} \leq \sum_{s' \in S} T_{ss'}^{a_f} v_L\left(s'\right) + Z(1-d_{s,k}^{a_l}) \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L, a_f \in \mathscr{A}_F, k \in \mathscr{K} \tag{E.12}$$

$$-Z d_{s,k}^{a_l} \leq x_{s,k}^{a_l a_f} \leq Z d_{s,k}^{a_l} \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L, a_f \in \mathscr{A}_F, k \in \mathscr{K} \tag{E.13}$$

$$y_{s,k}^{a_l a_f} \geq \sum_{s' \in S} T_{ss'}^{a_f} v_F\left(s'\right) - Z(1-d_{s,k}^{a_l}) \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L, a_f \in \mathscr{A}_F, k \in \mathscr{K} \tag{E.14}$$

$$y_{s,k}^{a_l a_f} \leq \sum_{s' \in S} T_{ss'}^{a_f} v_F\left(s'\right) + Z(1-d_{s,k}^{a_l}) \qquad \forall s \in \mathscr{S}, a_l \in \mathscr{A}_L, a_f \in \mathscr{A}_F, k \in \mathscr{K} \tag{E.15}$$

$$-Z d_{s,k}^{a_l} \leq y_{s,k}^{a_l a_f} \leq Z d_{s,k}^{a_l} \qquad \forall s \in \mathscr{S}, a_l \in A_L, a_f \in A_F, k \in \mathscr{K} \tag{E.16}$$

We refer the reader to [154] for stochastic Stackelberg games formulation and details of Stackelberg equilibria computation in stochastic games. We examine the effect of copayment policies on the players in a SSG and determine an optimal design and behavior of insureds corresponding to a Stackelberg equilibrium. The MILP was solved by IBM CPLEX Optimizer.

Appendix F

PARAMETERS USED IN THE ANALYSIS OF AGENT-BASED SIMULATION

Table F.1: Model parameter values

Variable	**Base Case (Range)/Distribution**		**Reference**
Cost	**Medical cost**	**Loss productivity**	[110]
No medical attendance	$c^{a}_{med,U}$	$c^{a}_{ind,U}$	
All risk	log normal (mean, std)	Poisson (mean)	
Age 0-4	(3,2)	145	
Age 5-17	(3,2)	75	
Age 18-49	(3,2)	75	
Age 50-64	(3,2)	75	
Age 65+	(3,2)	145	
Treatment taking	$c^{a,r}_{med,T}$		
Normal	log normal (mean, std)		
Age 0-4	(167, 307)	145	
Age 5-17	(95, 258)	145	
Age 18-49	(125, 438)	145	
Age 50-64	(150, 766)	290	
Age 65+	(242, 1544)	435	
High risk	log normal (mean, std)		
Age 0-4	(574, 1266)	870	
Age 5-17	(649, 1492)	580	
Age 18-49	(725, 1717)	290	
Age 50-64	(733, 1307)	580	
Age 65+	(476, 1131)	1015	
Hospitalization	$c^{a,r}_{med,H}$	$c^{a,r}_{ind,H}$	
Normal	log normal (mean, std)	Poisson (mean)	
Age 0-4	(10880, 36189)	1160	
Age 5-17	(15014, 86804)	1305	
Age 18-49	(19012, 44636)	1740	
Age 50-64	(22304, 95727)	1885	
Age 65+	(11451, 23128)	1885	
High risk	log normal (mean, std)	Poisson (mean)	
Age 0-4	(81596, 123626)	4495	
Age 5-17	(41918, 50393)	3335	
Age 18-49	(47722, 85644)	3045	
Age 50-64	(41309, 74798)	3480	
Age 65+	(16750, 32091)	2610	

Table F.1: Model parameter values (continued)

Variable	Base Case (Range)/Distribution	Reference
Death	$c^{a,r}_{med,D}$	
Normal	log normal (mean, std)	
Age 0-17	(28818, 24483)	
Age 18-49	(76336, 91654)	
Age 50-64	(118575, 333879)	
Age 65+	(41948, 96467)	
High risk	log normal (mean, std)	
Age 0-17	(267954, 221130)	
Age 18-49	(75890, 65267)	
Age 50-64	(118842, 345973)	
Age 65+	(33011, 61904)	
Indirect vaccination cost	90% population has (5,20) uniform distribution	assume
$c_{ind,V}$	10% population has 0 cost for voluntary vaccination	assume
Probability		[62]
$p^a_{F\|S}$	intermediate parameter	
$p^a_{F\|V}$	intermediate parameter	
Probability (hospitalization|infection)	$p^a_{H\|F}$	
All risk	log normal (mean, std)	[110]
Age 0-4	(0.0141, 0.0047)	
Age 5-17	(0.0006, 0.0002)	
Age 18-49	(0.0042, 0.0014)	
Age 50-64	(0.0193, 0.0064)	
Age 65+	(0.0421, 0.014)	
Probability (death|infection)	$p^a_{D\|F}$	[110]
All risk	log normal (mean, std)	
Age 0-4	(0.00004, 0.00001)	
Age 5-17	(0.00001, 0)	
Age 18-49	(0.00009, 0.00003)	
Age 50-64	(0.00134, 0.00045)	
Age 65+	(0.0117, 0.0039)	
Population N	563,441	
Public cost weight N	200 $(0-400)$	

All the probability values are converted into daily probabilities from original flu season data when running ABS.