A branch and bound algorithm for dynamic resource allocation in population disease management

TingYu Ho, Shan Liu, and Zelda B. Zabinsky

Department of Industrial and Systems Engineering, University of Washington, Seattle WA

Abstract

We embed an approximate dynamic program into a branch-and-bound framework to solve sequential resource allocation problems in population disease management. The proposed algorithm is capable of providing an optimality guarantee and getting bounds on the optimality gap of healthcare interventions. A numerical study on screening and treatment policy implementation for chronic hepatitis C virus (HCV) infection provides useful insights regarding HCV elimination for baby boomers.

Keywords: Dynamic resource allocation, Branch-and-bound, Population-based disease, Hepatitis C

2010 MSC: 00-01, 99-00

1. Introduction

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 [1]. In our previous research [2], we proposed a Multi-Fidelity Rollout Algorithm (MF-RA) 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. In this paper, we design 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. Our 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.

Our 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 we eliminate those paths that are known to be sub-optimal, 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.

One stream of literature closely related to ours is the literature on finding optimal intervention using dynamic programming. Dynamic programming is widely used to solve chronic disease management problems when chronic disease models are formulated using MDPs. Policy and value iterations, and linear programming have been commonly used to solve MDPs with small sizes of the state and action spaces [1]. Some research makes use of structural properties of individual-level models to ensure the existence of control-limit optimal policies [3, 4]. Population-level disease management typically involves a resource-limited setting, which results in an exponential growth of problem size due to sequential resource allocation. Approximate dynamic programming (ADP) approaches can be applied to reduce the curse of dimensionality and computational effort [5]. As an example, ADP was used to determine treatment allocation under different classes of policies regarding discontinuation of treatment for HIV control [6].

[1]Corresponding author, TingYu Ho, add address and email later.

We use one of the ADP approaches called a rollout algorithm to estimate future disease progression and population evolution [7].

Another stream of literature related to ours is the application of the branch-and-bound approach. Branch-and-bound has been used in many applications, e.g, data analysis [8], multi-objective functions [9], vehicle routing [10], video search [11], and scheduling [12]. Linear programming and various approximate and heuristic algorithms (e.g., genetic algorithm) are also used to find feasible solutions [13, 14, 15]. In this research, we develop an approach to embed an approximate dynamic program (rollout algorithm) in a branch and bound procedure.

Our contribution improves the existing MF-RA algorithm [2] by adding an optimality guarantee and optimality gap to allow decision makers to trade off computation with solution quality in large-scale dynamic resource allocation problem, which provides insights on long-term population disease management. Moreover, we provide numerical results using HCV elimination as an example. In particular, we evaluate the performance of long-term disease control under different annual budgets by quantifying the improvement of quality-adjusted life-years (QALYs) and computation time.

2. Resource Allocation Problem in Population Disease Management and RA-BnB

2.1. Basic model

We consider a disease with multiple health conditions in a closed population. As in [2], we divide the population in a healthcare management system into different groups based on the intervention they need or health conditions they are in. The group is indexed by $g \in \mathcal{G}$, where $\mathcal{G}$ is the set of all groups. Let $\widetilde{\mathcal{G}} \subset \mathcal{G}$ denote the set of groups waiting for interventions. The individuals in the same group $g \in \widetilde{\mathcal{G}}$ are eligible to receive the same intervention. We define a group-based health category $s \in \mathcal{S}_{HS} = \cup_{g\in\mathcal{G}}\mathcal{S}_g$, where $\mathcal{S}_{HS}$ is the set of all group-based health categories and $\mathcal{S}_g$ is the set of group-based health categories in group g. Let N_t^s be the population size in each category s at the beginning of time t, where t is the time index, and let $\mathbf{N}_t$ be the 1-by-$|\mathcal{S}_{HS}|$ row vector of population size. The decision variable $x_g^{(t)}$ is the fraction of budget for implementing intervention policy in group g during time period t.

A specific example is chronic HCV screening and treatment, which is in general cost-effective depending on treatment efficacy, drug cost, patients' initial disease stage and adherence to treatment, and care delivery strategy in the healthcare system [16, 17, 18, 19, 20, 21]. In the progression of chronic HCV infection, time period t is set to correspond to three months (a quarter). We use A, B, C, D, and M to denote five groups ($\mathcal{G} = \{A, B, C, D, M\}$), where group A indicates individuals have no knowledge of their HCV status, group B indicates those are identified HCV+, group C indicates those who are either screened healthy or recovered from HCV, group D indicates those who have untreatable HCV, and group M indicates those who are dead. Screening for population in group A and treatment for population in group B are recommended interventions for reducing the disease burden ($\widetilde{\mathcal{G}} = \{A, B\}$). Five sets of group-based health categories include target screening $\mathcal{S}_A$ (includes categories: healthy, no fibrosis, portal fibrosis with no septa, portal fibrosis with few septa, numerous septa without cirrhosis, compensated cirrhosis), target treatment $\mathcal{S}_B$ (includes categories: no fibrosis, portal fibrosis with no septa, portal fibrosis with few septa, numerous septa without cirrhosis, compensated cirrhosis), disease-free $\mathcal{S}_C$ (includes categories: healthy, recovered with history of mild fibrosis, recovered with history of moderate fibrosis, recovered with history of advanced fibrosis), end-stage liver disease $\mathcal{S}_D$ (includes categories: decompensated cirrhosis, hepatocellular carcinoma, liver transplant, after liver transplant), and dead $\mathcal{S}_M$ (includes one dead category). The total set of 20 group-based health categories is $\mathcal{S}_{HS} = \bigcup_{g\in A,B,C,D,M}\mathcal{S}_g$. In this HCV example, $\mathbf{N}_t$ is a 1-by-|20| row vector of population size for each $s \in \mathcal{S}_{HS}$. The decision variables are $x_A^{(t)}$ and $x_B^{(t)}$, where $x_A^{(t)}$ is the fraction of budget allocated to HCV screening tests for group A with health categories $s \in \mathcal{S}_A$ during time period t, and $x_B^{(t)}$ is the fraction of budget allocated to treatment for group B with health categories $s \in \mathcal{S}_B$.

The following objective function sums discounted QALYs over the entire lifetime horizon of the population given population size and intervention

$$QALY\left(\mathbf{N}_1, \left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}},t\in\{1,\ldots,n_T\}}\right) = \sum_{t=1}^{n_T+1}(1+\delta)^{-(t-1)}\mathbf{N}_t\mathbf{u}' = \sum_{t=1}^{n_T+1}\sum_{s\in\mathcal{S}_{HS}}(1+\delta)^{-(t-1)}u^s N_t^s, \tag{1}$$

where $\mathbf{N}_1$ is the 1-by-$|\mathcal{S}_{HS}|$ initial population size vector in the beginning of the first time period, $\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}},t\in\{1,\ldots,n_T\}}$ is the complete policy, i.e., budget allocation policy over n_T time periods in the entire lifetime horizon, $\mathbf{u}$ is a 1-by-$|\mathcal{S}_{HS}|$ 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 $s \in \mathcal{S}_{HS}$, and δ is the discount rate. Interventions are assumed to not decrease QALYs (either improve or keep the same QALYs).

As in [2], we consider the problem of maximizing the overall discounted QALYs for the population during budget planning under a lifetime horizon through healthcare interventions. The QALYs maximization problem is,

QALYs maximization problem

$$\max_{\mathbf{x}} QALY\left(\mathbf{N}_1, \left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}, t\in\{1,\ldots,n_T\}}\right) \tag{2}$$

$$\text{s.t.} \quad \mathbf{N}_{t+1} = \mathbf{N}_t\mathbf{P}_t\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}}\right),\ \forall t\in\{1,\ldots,n_T\} \tag{3}$$

$$\sum_{g\in\widetilde{\mathcal{G}}} x_g^{(t)} \le 1,\ \forall t\in\{1,\ldots,n_T\} \tag{4}$$

$$0 \le x_g^{(t)} \le 1,\ \forall g\in\widetilde{\mathcal{G}}, t\in\{1,\ldots,n_T\}\,. \tag{5}$$

In the objective function (2), calculating QALYs involves associating health categories with utility multiplier and the population size. Given population size in each group-based health category $s \in \mathcal{S}_{HS}$ and health intervention policy in each group $g \in \widetilde{\mathcal{G}}$, our transition matrix captures the dynamics of the disease progression and population evolution in Constraint (3), where $\mathbf{P}_t$ is the $|\mathcal{S}_{HS}|$-by-$|\mathcal{S}_{HS}|$ matrix of transition probabilities in time period t based on intervention policies $x_g^{(t)}$, $\forall g \in \widetilde{\mathcal{G}}$ and natural history of the disease. The natural history of disease progression and intervention policies drive the population evolution. The disease has a chance to progress every time period, as t ranges from 1 to n_T. The disease progression probabilities and mortality in the transition probability matrix $\mathbf{P}_t$ may be age- and gender-dependent, and thus $\mathbf{P}_t$ may be updated as the population ages. Constraint (4) indicates that the total budget in each time period t cannot exceed available budget during time period t. Constraint (5) indicates that the fraction of budget for group g is between 0 and 1.

The amount of budget allocated to each group $g \in \widetilde{\mathcal{G}}$ is $x_g^{(t)}B_t$, where B_t is the available budget during time period t. Note that the budget $x_g^{(t)}B_t$ may be allocated again within the group $g \in \widetilde{\mathcal{G}}$ based on its priority allocation rules. For example, in HCV management case, given fractions of budget for screening $x_A^{(t)}$ and treatment $x_B^{(t)}$, the budget allocation rules within groups A and B are different. Since the health status is unknown for candidates in group A, it is reasonable that the screening budget is uniformly allocated to screening fractions in each group-based health category $s \in \mathcal{S}_A$, e.g., population fraction in moderate-fibrosis health category who is offered screening is the same as population fraction in mild-fibrosis health category who is offered screening. However, the treatment allocation rule in group B prioritizes patients with more severe liver damage, e.g., the patients with moderate fibrosis have higher priority for treatment than the patients with mild fibrosis. Budget allocation rules within each group $g \in \{A, B\}$ determine the population size of receiving intervention in each group-based health category $s \in \mathcal{S}_A \cup \mathcal{S}_B$ and thus trigger the population evolution by transition probability $\mathbf{P}_t\left(x_A^{(t)}, x_B^{(t)}\right)$.

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. In the following subsections, we introduce how to efficiently compute upper bounds and an incumbent. We then present the details of the RA-BnB algorithm.

2.2. Upper bound computation

At time period t, given previous allocation policies before time period t, $\{x_g^{(t')}\}_{g\in\widetilde{\mathcal{G}}, t'\in\{1,\ldots,t-1\}}$, 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

$$\max_{\left\{x_g^{(t'')}\right\}_{g\in\widetilde{\mathcal{G}}, t''\in\{t,\ldots,n_T\}}} QALY\left(\mathbf{N}_1, \left\{x_g^{(t')}\right\}_{g\in\widetilde{\mathcal{G}}, t'\in\{1,\ldots,t-1\}}, \left\{x_g^{(t'')}\right\}_{g\in\widetilde{\mathcal{G}}, t''\in\{t,\ldots,n_T\}}\right) \tag{6}$$

$$\text{s.t.} \quad 0 \le x_g^{(t'')} \le 1, \forall g\in\widetilde{\mathcal{G}},\ \forall t''\in\{t,\ldots,n_T\} \tag{7}$$
$$\text{and (3).}$$

In the upper bound optimization problem, Constraint (4) in the QALYs maximization problem is eliminated. 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. Eliminating Constraint (4) essentially allows allocation of the same amount of budget B_t to every group $g \in \widetilde{\mathcal{G}}$. The complexity of evaluating different budget allocation policies among different groups is therefore eliminated. In addition, the $QALY$ function (1) is monotonically non-decreasing in $x_g^{(t)}$ due to the model structure, where interventions have a non-negative impact on QALYs. Thus, we can easily solve the upper bound optimization problem by setting $x_g^{(t'')}$ to its maximum value of one, i.e., the optimal budget allocation in the upper bound optimization is $x_g^{(t'')ub} = 1, \forall g \in \widetilde{\mathcal{G}}$ 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.

2.3. Low fidelity approximation

The Low Fidelity Approximation algorithm is similar to the rollout algorithm MF-RA in [2]. 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, we can quickly calculate the matrix of expected number of time periods spent in each state until absorption, beginning at time period t,

$$\mathbf{E}_t\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}}\right) = \left(\mathbf{I} - \mathbf{Q}_t\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}}\right)\right)^{-1} \tag{8}$$

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 $\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}}$ is expressed as the Low Fidelity Model (LFM),

$$\begin{aligned} LFM\left(\mathbf{N}_t, \left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}}\right) &= \mathbf{N}_t\mathbf{E}_t\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}}\right)\mathbf{u}' \\ &= \sum_{s\in\mathcal{S}_{HS}}\sum_{v\in\mathcal{S}_{HS}} N_t^s u^v E_t^{s,v}\left(\left\{x_g^{(t)}\right\}_{g\in\widetilde{\mathcal{G}}}\right) \end{aligned} \tag{9}$$

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 [2].

The Low Fidelity Approximation provides an incumbent solution which is easily computed. The incumbent solution and the QALYs upper bound calculated in Section 2.2 are incorporated into a branch-and-bound framework which is detailed in the following section.

2.4. RA-BnB algorithm

We introduce a decision period index d 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, $x_g^{(t)}$, $\forall t \in \mathcal{T}_d$ are the same, where $\mathcal{T}_d = \{n_B(d-1)+1, \ldots, n_B d\}$ is the set of time period indices in decision period d. That is, we keep the same percentages of budget for each intervention policy in every n_B time periods.

We discretize the continuous solution space into finite policies by discretizing the budget into m_B pieces, so $x_g^{(d)}$ is uniformly discretized and has $m_B + 1$ possible values, i.e., $x_g^{(d)} \in \{0, \frac{1}{m_B}, \ldots, \frac{m_B-1}{m_B}, 1\}$. We therefore have $m_P = C(|\widetilde{\mathcal{G}}| + m_B - 1, |\widetilde{\mathcal{G}}| - 1)$ different combinations[2] 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 $\mathbf{x}_{k_1,\ldots,k_d}^{(d)} =$

[2] $C(x, y)$ is often read as "x choose y" meaning the number of ways y can be chosen from x.

$\{x^{(d)}_{1,k_1,\ldots,k_d}, \ldots, x^{(d)}_{|\widetilde{\mathcal{G}}|,k_1,\ldots,k_d}\}$ denote the set of the k_dth combination of discretized budget in decision period d for all groups in $\widetilde{\mathcal{G}}$.

To illustrate, the example in Figure 1(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, we have 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 2.2, 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 2.3, 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 (1).

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

$$Q_U^{(d)} - Q^*. \tag{10}$$

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), and proceed to Step 1 for the next decision period, $d = d + 1$.

Theorem 1. *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 (1) 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, $Q^* = Q_{opt}$. □

2.5. An Example Problem

We solve an example problem to illustrate the working of the RA-BnB. We again use the example in Figure 1. In the first decision period (Figure 1(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, we assume 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, we assume 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 we 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 $\frac{Q_U^{(1)} - Q^*}{Q^* - Q_{no\,budget}} > \delta$, we go to Step 1 and proceed to the second decision period, $d = 2$ (Figure 1(b)).

Figure 2 presents an illustration of the Low Fidelity Approximation in Steps 4-1 through 4-5, using $\mathbf{x}_1^{(1)}$ as an example. Our 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 2(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 2(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 1(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

Figure 1: Illustration example of RA-BnB

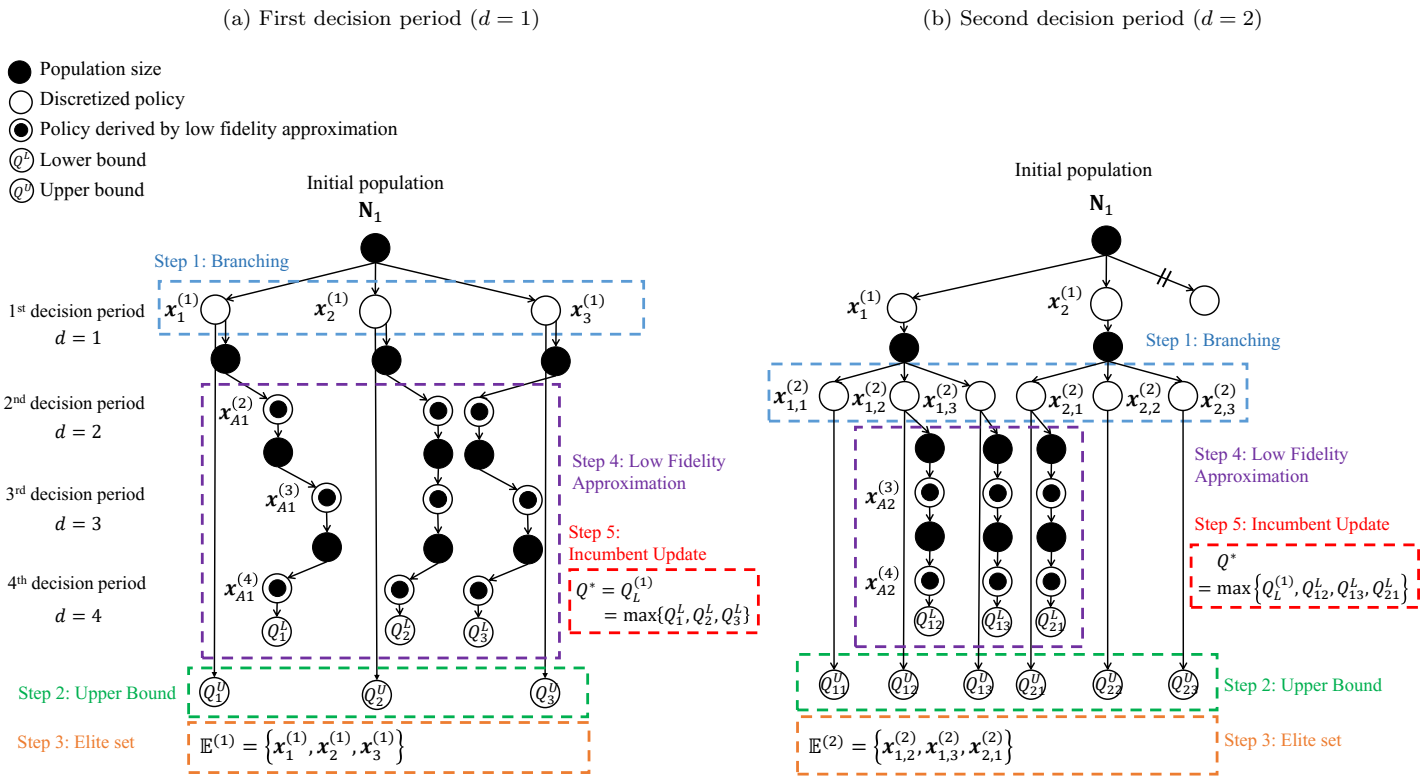
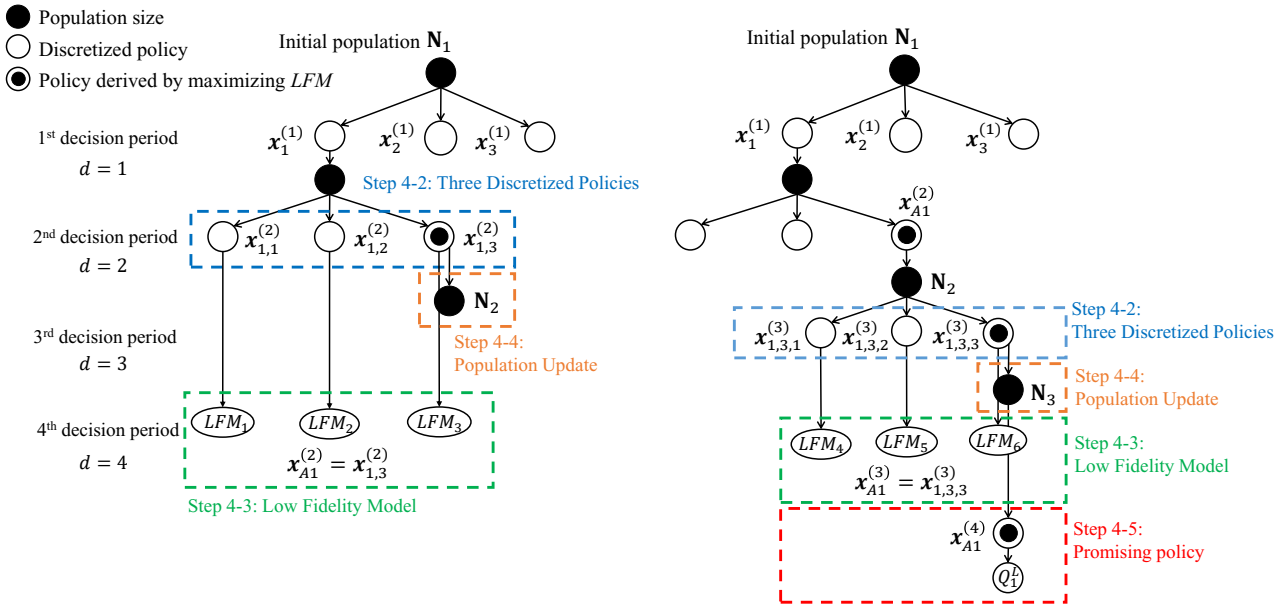


Figure 2: Illustration example of Low Fidelity Approximation in the first decision period in RA-BnB



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

3. Case Study: HCV Birth-Cohort Screening and Treatment Strategies

We 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 $1945-1975$). We 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 our previous research [2, 17, 18, 19, 21]. 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, we 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 ($m_B = 10$, $\left|\widetilde{\mathcal{G}}\right| = 2$ and $m_P = C(2+10-1,\ 2-1) = 11$). We use a stopping threshold of zero ($\delta = 0$) in RA-BnB. All numerical computations were carried out on a Intel coreI7 processor at 3.6GHz with 16GB RAM and Matlab R2015b.

Table 1 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 $(Q^{(d)}_U - Q^*)/(Q^* - Q_{no\ budget})$ 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 allocation 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, in paper [2], we 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 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, we are confident that RA-BnB can provide useful insights for policymakers. Using the $40-49$ birth cohort as an example, Figure 4 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 5 shows the optimal fraction of population for treatment and screening. The result in Figure 5(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 5(b) and (c) show that screening begins earlier when the

Table 1: 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)

Birth cohort	Iterations	1 billion/year Time (s)	1 billion/year # prune	1 billion/year Performance measure	4 billion/year Time (s)	4 billion/year # prune	4 billion/year Performance measure	8 billion/year Time (s)	8 billion/year # prune	8 billion/year 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.

Figure 3: Upper bound and incumbent lower bound over 15 decision periods for 40-49 birth-cohort

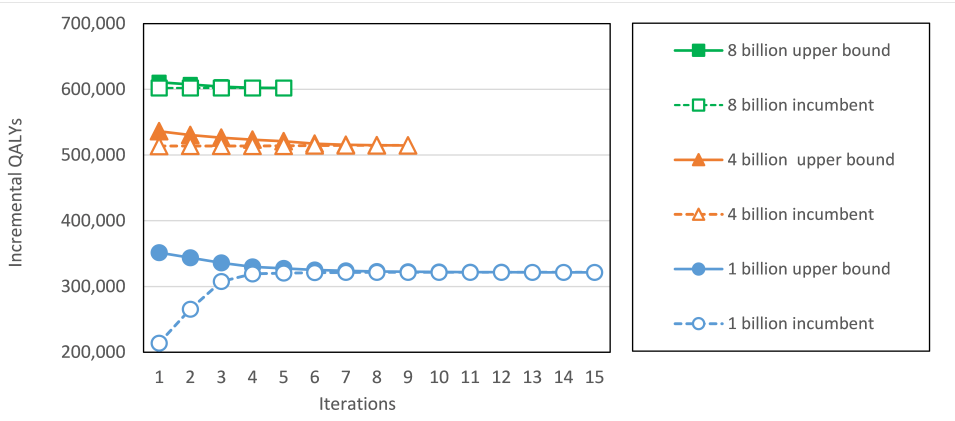


Figure 4: HCV control over 30 years with 15 decision periods for 40-49 birth-cohort

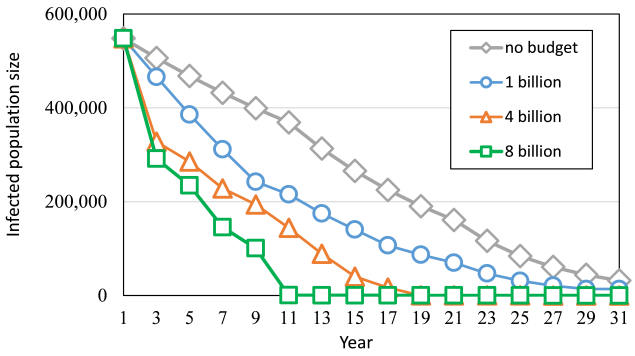
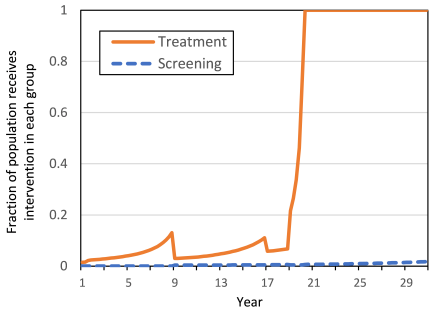
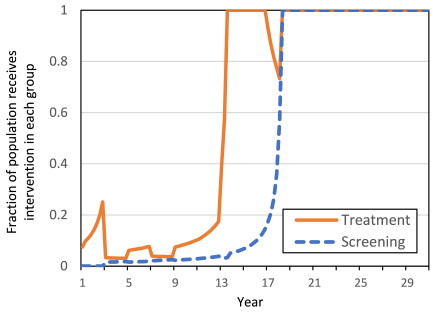


Figure 5: Optimal fraction of population for treatment and screening for 40-49 birth-cohort

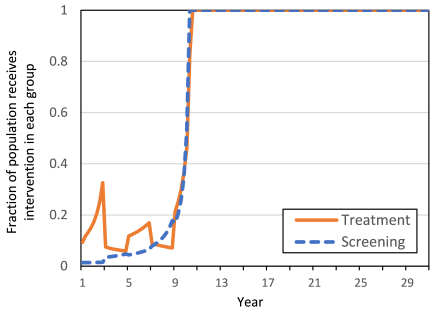
(a) $1 billion/year



(b) $4 billion/year



(c) $8 billion/year



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 4).

4. Concluding remarks

The success of the proposed RA-BnB algorithm, when applied to long-term budget allocation, can be attributed to the good qualities of a well-designed incumbent solution and upper bound, mitigating the curse of dimensionality. The numerical results show that $1 billion annual budget to address HCV is likely insufficient to have the desired impact on reducing the disease burden. To eliminate HCV for baby boomers, the government should consider increasing budget as well as implementing more effective screening and treatment interventions.

Acknowledgement

This work has been supported in part by the National Science Foundation grants CMMI−1632793 and CMMI−1824623.

References

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

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

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

[4] J. Chhatwal, O. Alagoz, E. S. Burnside, Optimal breast biopsy decision-making based on mammographic features and demographic factors, Operations Research 58 (2010) 1577–1591.

[5] W. B. Powell, Approximate Dynamic Programming: Solving the curses of dimensionality, volume 703, John Wiley & Sons, 2007.

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

[7] D. P. Bertsekas, Rollout algorithms for discrete optimization: A survey, in: Handbook of combinatorial optimization, Springer, 2013, pp. 2989–3013.

[8] M. J. Brusco, S. Stahl, Branch-and-bound applications in combinatorial data analysis, Springer Science & Business Media, 2006.

[9] T. Stidsen, K. A. Andersen, B. Dammann, A branch and bound algorithm for a class of biobjective mixed integer programs, Management Science 60 (2014) 1009–1032.

[10] P. Toth, D. Vigo, The vehicle routing problems, SIAM Monographs on Discrete Mathematics, vol. 9, SIAM Philadelphia, 2002.

[11] W.-S. Chu, F. De la Torre, J. F. Cohn, D. S. Messinger, A branch-and-bound framework for unsupervised common event discovery, International Journal of Computer Vision 123 (2017) 372–391.

[12] M. L. Pinedo, Scheduling: theory, algorithms, and systems, Springer, 2016.

[13] M. Fischetti, P. Toth, D. Vigo, A branch-and-bound algorithm for the capacitated vehicle routing problem on directed graphs, Operations Research 42 (1994) 846–859.

[14] R. Bouabda, B. Jarboui, M. Eddaly, A. Rebaï, A branch and bound enhanced genetic algorithm for scheduling a flowline manufacturing cell with sequence dependent family setup times, Computers & Operations Research 38 (2011) 387–393.

[15] M. Vila, J. Pereira, A branch-and-bound algorithm for assembly line worker assignment and balancing problems, Computers & Operations Research 44 (2014) 105–114.

[16] S. Liu, M. L. Brandeau, 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 (2017) 16–32.

[17] S. Liu, L. E. Cipriano, M. Holodniy, 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 (2013) e58975.

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

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

[20] S. Liu, D. Watcha, M. Holodniy, 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 (2014) 546–553.

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