Manuscript

Improving access to community-based chronic care through improved capacity allocation

(Authors' names blinded for peer review)

This paper studies a model of *community-based* health care delivery for *chronic diseases* in a *non-profit* setting. In this setting, patients periodically access the health care delivery system, which influences their disease progression and consequently their health outcomes. We investigate how the provider can maximize community-level health outcomes through better operational decisions pertaining to capacity allocation across different patients. To do so, we develop an integrated capacity allocation model that incorporates clinical (disease progression) and operational (capacity constraint) aspects. Specifically, we model the provider's problem as a finite horizon stochastic dynamic program, where the provider decides which patients to schedule at the beginning of each period. Therapy is provided to scheduled patients, which may improve their health states. Patients that are not seen follow their natural disease progression. We derive a quantitative measure for comparison of patients' health states and use it to design an easy-to-implement myopic heuristic that is provably optimal in a simple setting of the problem. We employ the myopic heuristic in a more general setting and test its performance using operational and clinical data obtained from Mobile C.A.R.E. Foundation, a community-based provider of pediatric asthma care in Chicago. We find that the myopic heuristic provides a performance that is very close to that of the optimal policy with an average optimality gap of 0.4%. Furthermore, the myopic heuristic can improve the health gains at the community-level by up to 15% over the current policy with the potential benefit being greater when capacity is tighter. The benefit is driven by the integrated nature of our myopic heuristic that ranks patients based on our quantitative measure of relative health states and dynamically alters the duration between visits for patients with different states depending on the tightness of capacity and health state of the entire patient population.

Key words: capacity allocation, chronic disease, mobile care, disease progression, appointment scheduling

---

1. Introduction

There has been a recent increase in operations research models for health care delivery (see Brandeau et al. 2004). Most models can be broadly partitioned based on their focus: (i) improvement in efficiency of the health care delivery system (reduce cost, increase revenue, reduce waiting time) and (ii) improvement in effectiveness of clinical decisions for individual patients (optimal time to initiate treatment, transplant an organ). In this paper, we consider a novel setting of *community-based* health care delivery for *chronic diseases*, where it is necessary to integrate these two approaches and investigate how operational decisions can improve population-level health outcomes.

Community-based chronic care is growing in North America fueled by an ageing population and growing disparities in access to care. These factors necessitate outreach to disadvantaged population groups in their communities rather than waiting for them to access health care in conventional settings. For example, Project Dulce run by Scripps Health provides diabetes care and management to thousands of ethnically diverse and low-income patients in San Diego County (http://www.scripps.org/services/diabetes/project-dulce). The Diabetes Integration Project in Manitoba provides care through mobile teams (http://www.diabetesintegrationproject.ca/). Similar programs have been initiated to target homeless populations (Post 2007).

Several features of these settings render conventional health care planning and scheduling models inapplicable due to the interaction of operational decisions and health outcomes. First, objective functions in most health care planning and scheduling models relate to efficiency or patient satisfaction; i.e., either cost minimization, revenue maximization, waiting time minimization, etc. (Gupta and Denton 2008), and the impact on health outcomes is not modeled explicitly. In community-based care, the most natural objective is to maximize aggregate health outcomes of the target population subject to resource constraints due to the non-profit and/or public model of delivery. Second, primary care or surgical procedures, which are the focus of most operational models, involve episodic access of the health care system by patients. Chronic care requires repeated interaction with patients whose disease states evolve over time as a function of prior care. Time between consecutive visits and health state at the prior visit have a direct impact on disease progression.

In this paper, we focus on the decision of allocating scarce capacity (appointment slots) among different patient segments (depending on their health states) to maximize aggregate health benefits measured in Quality Adjusted Life Years (QALYs) gained. We propose an integrated capacity allocation approach that combines clinical and operational decisions. Specifically, we formulate a finite horizon stochastic dynamic program comprising patients with different health states that compete for limited appointment slots in each period. We analyze a special case of the model that allows us to derive structural properties of the optimal capacity allocation policy and a quantitative measure to compare patients' health states. We then use the insights gained from this special case to develop an easily implementable heuristic for a more general version of the problem.

We calibrate our model and apply our results to data collected from Mobile C.A.R.E. Foundation (MCF), a non-profit organization that provides school-based asthma care for children in Chicago. Given a roster of active schools and a schedule of visits to these schools, MCF staff determines the capacity allocation at each school through daily patient schedules. The schedules are based on the medically recommended treatment duration of the patients that are modified based on the available capacity at each school. Results of our extensive numerical analysis using clinical and operational data from MCF show that our heuristic provides near-optimal performance (with optimality gaps less than 3% in all instances) and provides significant improvement (up to 15%) in health outcomes over current practice.

Our paper makes several contributions to the health care operations literature. First, we consider a novel setting of community-based health care delivery for chronic diseases in a non-profit setting, which necessitates the integration of clinical (disease progression) and operational (capacity constraint) decisions into a single decision model to maximize community-level health outcomes. Second, we propose a quantitative method of ordering patients with different health states in the absence of perfect observability since each patient is not seen each period due to capacity constraints. Third, we use this quantitative characterization to devise a simple and implementable solution approach, prove its optimality for a special case of our problem and demonstrate close-to-optimal performance for the general version. Fourth, we calibrate the model with data from a mobile provider of childhood asthma care and demonstrate that community level health outcomes can be improved significantly by making more effective operational decisions pertaining to capacity allocation without changing therapeutic decisions. Thus, our work highlights the significant potential of improving access to community-based chronic care programs through application of operations research methods to improve capacity allocation decisions.

The remainder of the paper is organized as follows. The clinical context of childhood asthma and related community-based programs are described in §2. Relevant streams of operational and health care literature are discussed in §3. The analytical model is developed in §4 and the optimal policy is characterized in §5. The myopic policy is described and analyzed in §6. Operational decision making at MCF is described in detail in §7, followed by model calibration. Results of our numerical study evaluating the improvement over current practice and associated policy implications are discussed in §8. Overall conclusions are summarized in §9 along with a discussion of potential extensions. All proofs are available in the Appendix.

2. Childhood asthma and community based programs

We focus on community-based care for childhood asthma due to the magnitude of disease burden, recent growth of such delivery models, lack of analytical models to inform operations decisions and MCF's willingness to collaboratively develop and implement these models.

Childhood asthma, a chronic respiratory disease characterized by periods of difficulty in breathing, affects about 6.5 million children in the U.S.. Childhood asthma places severe burden on families and society at large through utilization of health care resources as asthma management requires regular monitoring by health care providers even during asymptomatic periods. Absence of appropriate care can lead to asthma attacks. In 2004, such attacks resulted in 12.8 million missed school days, 750,000 emergency department visits, and 198,000 hospitalizations in the U.S. (Akinbami 2006). Severe disparities exist in prevalence of childhood asthma (higher), the resultant health care utilization (lower), and mortality (higher) among many ethnic minorities such as Native American, African American and Puerto Rican compared to white children (Marder et al. 1992). These disparities have been attributed to lack of health insurance coverage and awareness among parents and caregivers. Studies have found that children in inner city neighborhoods have higher prevalence of asthma likely due to factors such as allergens, tobacco smoke, pollution and poorly ventilated houses (Kitch et al. 2000) and family income levels (Gupta et al. 2008).

Community-based programs have been shown to reduce these disparities by improving access to health services for the underserved populations, including Chicago, Los Angeles, Baltimore, and St. Louis. These programs typically consist of a mobile clinic with the following common characteristics: (i) a school-based model, (ii) patient recruitment with involvement of parents and school nurses, and (iii) continuous patient follow-up using appointment scheduling and periodic school visits. Recent studies demonstrate the success of such programs in controlling asthma (Jones et al. 2007). While a return on investment calculator for mobile clinics has been developed (Oriol et al. 2009), there are no mathematical models to explicitly account for the impact of operational decisions on population-level health outcomes. In this paper, we aim to fill this gap by investigating how improved capacity allocation can improve health outcomes for a population of asthmatic children and apply our findings to MCF as a case study.

3. Literature review

We discuss existing work in four streams of relevant literature. The first two streams focus, separately, on two aspects of health care delivery–appointment scheduling and medical decision making–that we combine. The third stream, literature on restless multiarmed bandit problems, focuses on developing solution methods for problems with similar structure as ours. The fourth stream comprises models of machine repair that have structural similarity with ours.

Appointment Scheduling literature. Gupta and Denton (2008) review health care appointment scheduling challenges in three common systems, namely primary care (outpatient facilities), and specialty clinics and elective surgeries (inpatient facilities). Models of primary care appointment systems consider patients that arrive punctually for their scheduled appointment time to a single physician and experience stochastic service times that are identically and independently distributed. The key decisions involved are length of the appointment slot (Cayirli and Veral 2006), appointment times (Robinson and Chen 2003), and number of appointment requests to accept including overbooking (Gupta and Wang 2008). The objective is to minimize the total cost of direct patient waiting (inside the clinic), physician idleness and/or overtime (Denton and Gupta 2003). Recent papers analyze open access systems (Murray and Tantau 1999) where some slots are reserved for same day arrivals (Liu et al. 2010, Robinson and Chen 2010). The key tradeoff here is between reduction of indirect waiting time (between a request for appointment and the actual appointment) and a possible increase in direct waiting time and physician overtime.

Our work is distinct due to our focus on chronic care whereby patients return for follow-up visits instead of unidentifiable appointment requests as in the above models. Gupta and Denton (2008) highlight the paucity of planning and scheduling models for chronic conditions. Lee and Zenios (2009) is an exception but they consider exogenous flows of patients between different system compartments. In contrast, we explicitly model the impact of capacity allocation on the health state of patients. Thus, our objective is maximization of aggregate quality of life for patient population rather than cost minimization or revenue maximization.

Medical decision making literature. Several OR models have been proposed recently to optimize timing of clinical decisions such as liver transplants (Alagoz et al. 2005) and initiation of HIV therapy (Shechter et al. 2008). While these models employ disease progression models, they typically consider individual patient outcomes in contrast to our focus on population health outcomes. Some models of organ transplant consider population level dynamics (e.g., Zenios et al. 2000) but differ from our model because of the one-time nature of their intervention as opposed to periodic follow-up in our case. There is a vast literature on cost-effectiveness of medical interventions (Sonnenberg and Beck 1993) including asthma (Paltiel et al. 2006) that analyzes disease progression at the population level. However, these models consider theoretical cohorts of patients, do not model capacity constraints in the health care system and do not optimize over decisions (Brennan et al. 2006). In contrast, we explicitly model the health care delivery system and develop models to generate qualitative insights and provide decision support on the ground.

Restless multi-armed bandit problem literature. Our model belongs to a general class of dynamic resource allocation problems: *restless multiarmed bandit (RMAB)* problem (Whittle 1988). In each period, the decision maker chooses the number of arms to activate (the number of projects to work on) subject to a capacity constraint with the objective of maximizing the total discounted return over a long time horizon. Return from each project depends on its current state (active or passive). The state of each project evolves stochastically; the transition probabilities depend on whether the project was active or passive in that period. RMAB problem is proven to be intractable in general (Papadimitriou and Tsitsiklis 1994) and hence the majority of work in this stream has focused on developing heuristics. One of the most widely used heuristics is Whittle's index that is based on Lagrangian relaxation of the capacity constraint linking the projects. The index corresponds to a (state-dependent) subsidy that makes the decision maker indifferent between keeping a project passive and activating it. The projects are ranked based on the index and the ones with the highest index, subject to the capacity constraint, are chosen to be active.

Weber and Weiss (1990) prove that Whittle's index is asymptotically optimal when the ratio of the active number of bandits to the total number is fixed. Most papers (e.g., Ansell et al. 2003, Glazebrook et al. 2005a, 2006) focus on proving indexability before proceeding to compute Whittle's index and numerically solving the problem. Instead, we take an alternate approach. We prove that a much simpler myopic policy is optimal for a special case of our problem and use it to devise a heuristic for the general formulation. The performance of this heuristic is comparable to Whittle's index but the structure is much simpler thus making it more attractive from a practical perspective. Liu and Zhao (2009) also shows the optimality of myopic policy but their model is a special case of ours. In their model, action in each period corresponds to only *inspection* of active arms. In our model, action corresponds to *inspection* of active arms combined with *treatment*, which complicates the state transition dynamics.

Machine repair literature. Our stochastic dynamic program is related to models of machine repair and maintenance with multiple units (e.g., Wang 2002). At an abstract level, our model is most closely related to Glazebrook et al. (2005b), which studies the optimal allocation of repairmen to machines which deteriorate under usage. However, there are two major differences. First, we generalize their assumption of "perfect repair"; i.e., machines return to the best possible state immediately after repair. In our model, the treatment matrix is imperfect and directly estimated from clinical data. Second, Glazebrook et al. (2005b) assume that the states of machines are known at every decision epoch. In our model, since a patient's health state is not observable without a diagnosis during the visit and every patient is not seen each period, we assume that the decision maker only knows the probability distribution over health states. Pham and Wang (1996) consider "imperfect repair" but characterize a machine's status as a binary variable; i.e., good or bad. We consider a more detailed model of disease progression comprising multiple health states. Rosenfield (1976) studies the optimal policy for deteriorating processes with imperfect information that is governed by a discrete time Markov process in an unconstrained setting. Our model also differs from machine maintenance models with imperfect information (Valdez-Flores and Feldman 1989) in that inspection (diagnosis) and repair (treatment) are inseparable in our setting.

4. Model

We formulate the capacity allocation problem as a discrete-time, finite-horizon, discounted Markov Decision Process (MDP) model with the objective of maximizing the quality adjusted life years (QALY) for the entire patient cohort. We propose an integrated approach to capacity allocation that combines both operational and clinical decisions. In contrast, the current policy of capacity allocation at MCF, as discussed in Section 7.2, makes operational and clinical decisions in a decoupled manner in two stages. Assuming a fixed schedule of equally-spaced visits to a school, the sequence of events in a period are as follows:

- Payoff (in terms of QALY values) of the previous period's decision is collected
- Capacity allocation decisions for the current period are made
- Appointments occur for patients scheduled in the current period
  - — True health state is diagnosed
  - — Treatment is applied and takes effect immediately
- Natural disease progression occurs for all patients

The description of the model is divided into two parts: Sections 4.1 through 4.3 present a discrete-time Markov chain for individual patient disease progression and Section 4.4 presents a constrained Markov Decision Process for allocating resources among all patients.

4.1. Individual Patient Model: Disease Progression and Treatment

Consistent with the disease progression models of asthma in the medical literature (Paltiel et al. 2006, Shahani et al. 1994), we consider a homogenous patient population of I patients, whose disease progression is governed by a Markov process over discrete health states $0, 1, \ldots, K$, where 0 represents the best and K represents the worst health state, respectively.

At the beginning of Period $t \in \{1, 2, \ldots, T\}$, Patient i's health state is given by a tuple, $\vec{s}_{i,t} = (h_{i,t}, n_{i,t})$, where $h_{i,t} \in \{0, 1, \ldots, K\}$ represents the health state at the last appointment, and $n_{i,t} \in \{1, 2, ..., T-1\}$ represents the time since last appointment measured by number of periods.

The natural disease progression, without any medical intervention, is characterized by a per-period transition matrix $\mathcal{P}$, where p_{ij}, the $(i,j)^{th}$ element of matrix $\mathcal{P}$, is the probability that a patient who is in health state i be in health state j after one period, ($\sum_{j=0}^{K} p_{ij} = 1, \forall i \in \{0, 1, \ldots, K\}$ and $p_{ij} \geq 0, \forall i, j \in \{0, 1, \ldots, K\}$). In the presence of medical intervention, the state transition is governed by both treatment and natural disease progression, which we explain next.

A typical clinical appointment for chronic care involves *diagnosis* of patient's current health state and adjustment of future *treatment* plan. Treatment for chronic conditions consists of a clinical component (modifying prescribed drugs and/or their dosages) and a behavioral component (counseling/education of patients and caregivers). Clearly, the effect of treatment is strongest immediately following the appointment but wanes over time until the next appointment. However, explicitly modeling this temporal effect introduces non-stationarity and severely hampers the analytical tractability of our model. To simplify this complex situation, we assume that the treatment effect occurs immediately after the appointment, improving the patient's current health state. This is modeled by a lower triangular treatment matrix $\mathcal{Q}$, where the $(i,j)^{th}$ element in matrix $\mathcal{Q}$ is denoted by q_{ij} ($\sum_{j=0}^{K} q_{ij} = 1, \forall i \in \{0, 1, \ldots, K\}$ and $q_{ij} \geq 0, \forall i, j \in \{0, 1, \ldots, K\}$). After the treatment effect occurs, the patient's disease progression is again governed by $\mathcal{P}$. We assume the disease progression process and treatment process to be independent of each other. The effective state transition of patients can be interpreted as a Markov chain with two transition rates that depend on the time since the last visit. The patient transition matrix is $\mathcal{QP}$ for the first period after the visit and $\mathcal{P}$ for all subsequent periods until the next appointment. Note that, while the isolated effect of treatment is to only improve patient health, the combined effect of $\mathcal{Q}$ and $\mathcal{P}$ between two visits might be such that a patient under treatment can transition to a worse health state before the next visit.

4.2. Information vector

At the time of making capacity allocation decisions, the current health state of the patients is not known with certainty. We use random variable $\mathbf{x}_{i,t}$ to denote Patient i's true health state at the beginning of Period t and $x_{i,t} \in \{0, 1, \ldots, K\}$ as its realizations. Then, for Patient i with state $\vec{s}_{i,t} = (h_{i,t}, n_{i,t})$, the distribution of $\mathbf{x}_{i,t}$ is given by

$$\vec{\pi}_{i,t} = \vec{e}_{h_{i,t}} \mathcal{Q} \mathcal{P}^{n_{i,t}},$$

where $\vec{e}_k$ is a row vector of $K+1$ zeros with a one in its $(k+1)^{th}$ element. We refer to this as the *information vector* of Patient i at the beginning of Period t. It represents the health care provider's belief about Patient

i's true health state at the beginning of Period t before a capacity allocation decision is made and the patient is seen. Note that for an information vector with $K+1$ elements, the first K elements are sufficient to define the entire vector since all elements sum to 1.

The system state at the beginning of Period t can be characterized by every patient's health state at the previous visit and the time since then, which is $(\vec{s}_{1,t}, \vec{s}_{2,t}, \ldots, \vec{s}_{I,t})$ or $\big((h_{1,t}, n_{2,t}), (h_{2,t}, n_{2,t}), \ldots, (h_{I,t}, n_{I,t})\big)$, or equivalently the information vectors; i.e., the distribution of the true state of each patient, which is $(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t})$ (Rosenfield 1976, Lovejoy 1987).

4.3. Reward

Let b_k denote the quality of life (QoL) score associated with health state k. In any Period t, the expected health reward for Patient i with state $(h_{i,t}, n_{i,t})$ and corresponding information vector $\vec{\pi}_{i,t}$ is given by:

$$\phi(\vec{\pi}_{i,t}) = \sum_{k=0}^{K} (\vec{\pi}_{i,t})_k b_k \tag{1}$$

where $(\vec{\pi}_{i,t})_k$ is the $(k+1)^{th}$ element in vector $\vec{\pi}_{i,t}$, representing the probability that Patient i is in health state k at the beginning of Period t before a decision is made. Without loss of generality, we assume that $b_0 \geq b_1 \geq \ldots \geq b_K$ since higher index denotes worse health state.

4.4. Capacity allocation

We represent capacity allocation decisions by binary decision variables $a_{i,t}$, where $a_{i,t} = 1$ if Patient i is scheduled in Period t, and $a_{i,t} = 0$ otherwise. Denoting capacity in each period by C and assuming that all scheduled patients attend their appointments, the capacity constraint is given by:

$$\sum_{i=1}^{I} a_{i,t} \leq C, \qquad \forall t \in \{1, 2, \ldots, T-1\} \tag{2}$$

Note that since the payoff is collected at the beginning of every period, for a T period problem, only decisions in the first $T-1$ periods impact the objective function. Since treatment is beneficial (i.e., the treatment matrix $\mathcal{Q}$ is lower triangular), the capacity constraint (2) is binding at optimality. Therefore, we consider only allocations that use all capacity. Let $\mathbf{N}$ denote the set containing all selections or subsets from index vector $\{1, 2, \ldots, I\}$ with a cardinality of C. An element $\mathbf{n}_t \in \mathbf{N}$ represents a feasible capacity allocation rule in Period t.

We define the value function $u_t^*(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t})$ as the total optimal discounted expected payoff from Period t until the end of the horizon if the information vector for all patients at the beginning of Period t is $(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t})$. Puterman (1994) shows that, for the discounted MDP, for any optimal policy, its induced value function must satisfy the following optimality equations:

$$u_t^*(\vec{\pi}_{1,t}, \ldots, \vec{\pi}_{I,t}) = \sum_{i=1}^{I} \phi(\vec{\pi}_{i,t}) + \beta \max_{\mathbf{n}_t \in \mathbf{N}} \Big\{ \mathbb{E}\big[u_{t+1}^*(\vec{\boldsymbol{\pi}}_{1,t+1}\ ,\ \ldots\ ,\ \vec{\boldsymbol{\pi}}_{I,t+1}) \big| \mathbf{n}_t\big] \Big\}, t \in \{1, \ldots, T-1\}, \tag{3a}$$

$$u_T^*(\vec{\pi}_{1,T}, \ldots, \vec{\pi}_{I,T}) = \sum_{i=1}^{I} \phi(\vec{\pi}_{i,T}), \tag{3b}$$

where $\beta \in (0,1)$ is the discounting factor. The information vector for Patient i at the beginning of Period $t+1$ is given by:

$$\vec{\boldsymbol{\pi}}_{i,t+1} = \begin{cases} \vec{e}_k \mathcal{Q}\mathcal{P} \text{ with probability } (\vec{\pi}_{i,t})_k, \forall k \in \{0,\dots,K\}, & \text{if Patient } i \text{ is scheduled in Period } t \\ \vec{\pi}_{i,t}\mathcal{P}, & \text{otherwise} \end{cases} \tag{4}$$

5. Analysis of the optimal policy

In this section, we establish the structural properties of the MDP model formulated in (3a)–(3b), which falls under the category of restless multi-armed bandit (RMAB) problems (Whittle 1988). The general RMAB problem is known to be PSPACE-complete (Papadimitriou and Tsitsiklis 1994), which makes it analytically intractable. Consequently, most of the effort in the literature has been devoted to developing heuristics, usually an index policy (e.g. Whittle 1988, Bertsimas and Nino-Mora 2000). In contrast to this approach, we use the characteristics of our specific operational setting to derive the structure of the optimal policy and obtain managerial intuition.

5.1. Two health states

With two health states ($k \in \{0,1\}$), the information vector of Patient i at the beginning of Period t can be expressed as $\vec{\pi}_{i,t} = (\pi_{i,t}, 1-\pi_{i,t})$, and $\pi_{i,t}$ is sufficient to characterize $\vec{\pi}_{i,t}$. Therefore, to simplify notation, in this section, we define $\pi_{i,t} = (\vec{\pi}_{i,t})_0$, i.e., the first element in the information vector and call it, simply, the *information* for patient i. Consequently, for our two health state model, the value function $u_t^*(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \dots, \vec{\pi}_{I,t})$ is simplified to $u_t^*(\pi_{1,t}, \pi_{2,t}, \dots, \pi_{I,t})$.

For further simplicity, we denote the disease progression matrix $\mathcal{P}$ and treatment matrix $\mathcal{Q}$ as:

$$\mathcal{P} = \begin{bmatrix} p_0 & 1-p_0 \\ p_1 & 1-p_1 \end{bmatrix} \tag{5}$$

$$\mathcal{Q} = \begin{bmatrix} 1 & 0 \\ q & 1-q \end{bmatrix} \tag{6}$$

In the following proposition, we establish the monotonicity and componentwise convexity properties of value function $u_t^*(\pi_{1,t}, \pi_{2,t}, \dots, \pi_{I,t})$.

PROPOSITION 1. *If $p_0 > p_1$, then for all $i \in \{1,2,\dots,I\}$ and for all $t \in \{1,2,\dots,T\}$: (i) $u_t^*(\pi_{1,t}, \pi_{2,t}, \dots, \pi_{I,t})$ is nondecreasing in $\pi_{i,t}$ and (ii) $u_t^*(\pi_{1,t}, \pi_{2,t}, \dots, \pi_{I,t})$ is componentwise convex in $\pi_{i,t}$.*

Observe that in the two health state case, all patients in the community can be ranked by their *information* completely. Also, recall that information $\pi_{i,t}$ for Patient i is the probability that Patient i's true health state is 0 at the beginning of Period t. The larger the information, the greater the probability that a patient is in a better health state. Therefore, intuitively it is more beneficial to give a lower priority to such patients. This intuition is formally proven to be correct in Theorem 1 below.

THEOREM 1. *Given any Period* $t \in \{1,2,\ldots,T-1\}$*, without loss of generality, rank all patients by their information such that* $\pi_{1,t} \le \pi_{2,t} \le \cdots \le \pi_{I,t}$*. It is optimal to schedule Patients* $1,2,\ldots,C$*.*

Thus, the optimal policy allocates capacity preferentially to patients that are more likely to be in worse health state.

5.2. Multiple health states

For the multiple health state case, due to the intractability of the problem we focus on a special situation where the treatment is very effective. This is referred to as "perfect treatment". Mathematically, this corresponds to a treatment matrix $\mathcal{Q}$, in which $q_{i1} = 1\ \forall i$ and $q_{ij} = 0\ \forall i,\ \forall j \neq 1$.

Next, we establish a proposition that is parallel to Proposition 1, characterizing the monotonicity properties of value function $u_t^*(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t})$. Because function $u_t^*(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t})$ maps information vectors to $\mathbb{R}_{(+)}$, we use first order stochastic dominance $\succ_{s.t.}$ to define an order over information vectors such that monotonicity is properly defined. Formally, $\vec{\pi}_{i,t} \succ_{s.t.} \vec{\pi}_{j,t}$ implies that, at the beginning of Period t, Patient i has a higher or equal probability than Patient j to be in any health state k or worse. For ease of exposition, we refer to this by saying that Patient i's health state is *worse* than that of Patient j.

We also assume that any patient starting in a worse health state remains in a worse health state in the next period after treatment. Formally, $x_{i,t} > x_{j,t}$ and $a_{i,t} = a_{j,t} = 1$ implies that $\vec{\pi}_{i,t+1} \prec_{s.t.} \vec{\pi}_{j,t+1}$. Defining $\vec{\gamma}_k = \vec{e}_k \mathcal{Q}\mathcal{P}$, from (4), this is equivalent to requiring that $\vec{\gamma}_{x_{i,t}} \prec_{s.t.} \vec{\gamma}_{x_{j,t}}$. Similarly, for two patients that are not scheduled in Period t, we assume that $\vec{\pi}_{i,t} \succ_{s.t.} \vec{\pi}_{j,t}$ implies that $\vec{\pi}_{i,t+1} \succ_{s.t.} \vec{\pi}_{j,t+1}$. Defining $z(\vec{\pi}_{i,t}) = \vec{\pi}_{i,t}\mathcal{P}$ and using (4), this is equivalent to requiring that $z(\vec{\pi}_{i,t}) \succ_{s.t.} z(\vec{\pi}_{j,t})$, i.e., Patient i remains worse than Patient j in the next period, if neither is seen. We say that $z(\cdot)$, or equivalently, the corresponding $\mathcal{P}$ that defines $z(\cdot)$, *preserves stochastic order*.

PROPOSITION 2. *If* $\vec{\gamma}_0 \prec_{s.t.} \vec{\gamma}_1 \prec_{s.t.} \ldots \prec_{s.t.} \vec{\gamma}_K$*, and* $z(\cdot)$ *preserves stochastic ordering, then* $u_t^*(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t})$ *is componentwise decreasing in the sense that for all* $\vec{\pi}_{i,t} \succ_{s.t.} \vec{\pi}'_{i,t}$ *and for all* $t \in \{1,2,\ldots,T\}$*, we have* $u_t^*(\vec{\pi}_{1,t}, \ldots, \vec{\pi}_{i,t}, \ldots, \vec{\pi}_{I,t}) \le u_t^*(\vec{\pi}_{1,t}, \ldots, \vec{\pi}'_{i,t}, \ldots, \vec{\pi}_{I,t})$*.*

The above proposition implies that if two systems have exactly the same states at the beginning of Period t except for the state of Patient i, where one is worse than the other (i.e., $\vec{\pi}_{i,t} \succ_{s.t.} \vec{\pi}'_{i,t}$), then the system in which Patient i is worse ($\vec{\pi}_{i,t}$) should have a lower aggregated QALY under the optimal policy. Given this proposition, if Patient i is worse than Patient j, it is more beneficial to schedule Patient i than Patient j. This intuition is formally proven in Theorem 2.

THEOREM 2. *Let* $(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{C,t})$ *be the information vectors of patients who are scheduled and* $(\vec{\pi}_{C+1,t}, \vec{\pi}_{C+2,t}, \ldots, \vec{\pi}_{I,t})$ *be the information vectors of patients who are not scheduled under optimal policy in any given Period* $t \in \{1,2,\ldots,T-1\}$*. If*

1. *the treatment is perfect, and*

2. *$z(\cdot)$ preserves stochastic ordering;*

then $\forall i \in \{1, 2, \ldots, C\}$ and $j \in \{C+1, C+2, \ldots, I\}$ such that $\vec{\pi}_{i,t}$ and $\vec{\pi}_{j,t}$ are comparable using first order stochastic dominance, it must be that $\vec{\pi}_{i,t} \succ_{s.t.} \vec{\pi}_{j,t}$. In other words, any patient in the scheduled patient group must have an information vector stochastically dominating the information vector of any patient in the not-scheduled patient group, if a stochastic ordering can be achieved.

Theorem 2 extends the intuition of prioritizing "sicker" patients to a model with multiple health states by appropriately modifying the quantitative characterization of the term "sicker" using first order stochastic dominance. Note that Theorem 2 is a partial characterization of the optimal policy, in that we only specify priority between any two patients when their information vectors can be ranked stochastically. However, it is quite likely that two patients' information vectors cannot be ranked stochastically, in which case Theorem 2 does not specify their relative priority.

6. Myopic policy

The partial characterization of the optimal policy in Theorem 2 provides insights into the design of a relatively efficient heuristic. We define *myopic policy* as the policy that schedules patients to maximize the community's aggregated QALY only in the immediately next period after the scheduling decision. Formally, at the beginning of Period t, myopic policy $\Pi(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t})$ determines a selection $\mathbf{n}_t \in \mathbf{N}$ to be seen in Period t, by solving the following static optimization problem:

$$\max_{\mathbf{n}_t \in \mathbf{N}} \mathbb{E}\Big[\sum_{i=1}^{I} \phi(\vec{\boldsymbol{\pi}}_{i,t+1}) \Big| \vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{I,t}\Big] \tag{7}$$

THEOREM 3. *Let $(\vec{\pi}_{1,t}, \vec{\pi}_{2,t}, \ldots, \vec{\pi}_{C,t})$ be the information vectors of patients who are scheduled and $(\vec{\pi}_{C+1,t}, \vec{\pi}_{C+2,t}, \ldots, \vec{\pi}_{I,t})$ be the information vectors of patients who are not scheduled under myopic policy Π in any given Period $t \in \{1, 2, \ldots, T-1\}$. If*

1. *the treatment is perfect, and*
2. *$z(\cdot)$ preserves stochastic ordering;*

then for any $i \in \{1, 2, \ldots, C\}$ and $j \in \{C+1, C+2, \ldots, I\}$ such that $\vec{\pi}_{i,t}$ and $\vec{\pi}_{j,t}$ are comparable using first order stochastic dominance, it must be that $\vec{\pi}_{i,t} \succ_{s.t.} \vec{\pi}_{j,t}$.

Theorem 3 shows that the structure of the myopic policy is indeed similar to that of the optimal policy under perfect treatment assumption and the conditions that are required for Theorem 2. Moreover, with only two health states, all patients can be ranked stochastically, and the optimal policy is exactly the same as the myopic policy. Theorem 4 formalizes this observation.

THEOREM 4. *With only two health states, in any Period $t \in \{1, \ldots, T-1\}$, without loss of generality, rank patients by their information such that $\pi_{1,t} \leq \pi_{2,t} \leq \cdots \leq \pi_{I,t}$. There exists an optimal myopic policy Π that schedules patients $1, 2, \ldots, C$.*

We use Theorems 3 and 4 to develop a *myopic heuristic* for more general cases with multiple health states and imperfect repair. The myopic heuristic can be interpreted as an index policy. Patients are ranked at the beginning of Period t based on indices given by $\Psi_{i,t+1}(\vec{\pi}_{i,t}) = \sum_{k=0}^{K}(\vec{\pi}_{i,t})_k \phi(\vec{\gamma}_k) - \phi(z(\vec{\pi}_{i,t}))$, which correspond to the marginal benefit of scheduling Patient i in Period $t+1$. We formally describe the heuristic below:

Myopic heuristic for scheduling patients in Period $t \in \{1, 2, \ldots, T-1\}$

Step 0: Calculate the information vector $\vec{\pi}_{i,t}$ based on $h_{i,t}$ and $n_{i,t}$ for all patients $i \in \{1, 2, \ldots, I\}$.
Step 1: Calculate the marginal benefit $\Psi_{i,t}(\vec{\pi}_{i,t})$ for all patients $i \in \{1, 2, \ldots, I\}$.
Step 2: Rank all patients by $\Psi_{i,t}$ in descending order.
Step 3: Schedule the first C patients from the above ranking.

7. Model Calibration

In this section, we calibrate our model using operational and clinical data from Mobile C.A.R.E. Foundation (MCF) regarding their school-based asthma care children. Founded in 1998, MCF has partnered with over 100 schools and Head Start programs, assisted in the screening of about 50,000 children for asthma and treated nearly 5,000 patients. We describe the data in Section 7.1 and provide a qualitative and quantitative characterization of the current capacity allocation process at MCF in Section 7.2. In Section 7.3, we estimate the disease progression and treatment matrices of our model. These form the basic building blocks for calculating the potential benefit of the integrated allocation approach over the current practice.

7.1. Data

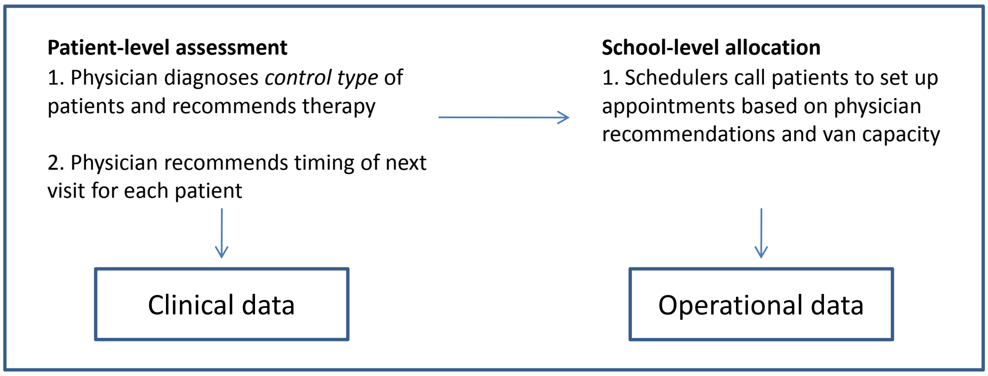
We utilize a comprehensive dataset consisting of data that are routinely collected during patient visits to the vans using an electronic medical records system called Asma-Trax® (see Jones et al. 2007). It comprises 29,745 observations (appointments/visits) for 5,041 individual patients, over ten years between 1999 and 2009. During each patient visit, the medical staff records a variety of clinical indicators self-reported by patients and caregivers, including the occurrence of day and night time symptoms, $FEV_1\%$ (Forced Expiratory Volume in 1 second) measured using spirometry, number of school days missed and number of hospital visits. These data are used to characterize patient health along two dimensions—*severity* and *control*. Severity captures the inherent intensity of the disease process in a patient, which typically remains unchanged across visits. MCF, consistent with the broader clinical literature, maintains four categories: *mild intermittent*, *mild persistent*, *moderate persistent*, and *severe persistent*. Control indicates how well asthma-related symptoms are currently controlled in a patient. MCF maintains four categories: *controlled*, *improved*, *unchanged* and *worsened*, where the last three fall under the umbrella of *uncontrolled*.

7.2. Characterizing current policy at MCF

We present a qualitative overview of the current capacity allocation policy at MCF in Section 7.2.1, followed by its quantitative characterization in Section 7.2.2 using MCF data.

7.2.1. Qualitative overview of the current practice. The current MCF scheduling process is shown in Figure 1 and described in greater detail in Deo et al. (2009). Notably, scheduling is performed in two steps. First, physicians recommend a due date for the next visit that is primarily driven by a patient's control status. According to the MCF medical team the recommended interval between current and next visit is: 3 months for *controlled* category of patients and 1 month for the *uncontrolled* categories.

Figure 1 Current capacity allocation process at MCF.



Two to four weeks before a school visit, based on the physician assignments and available van capacity (typically 16 slots), schedulers at the MCF office develop a feasible allocation of capacity among the school population for that visit. In this scheduling process, priority is given in the following order: (1) patients who request appointments, (2) patients who could not be seen previously due to capacity limits, (3) patients who have missed prior appointments, (4) all other patients who are due for an appointment, and (5) new patients.

This two-step process with disjoint operational and clinical considerations is driven by a lack of a systematic framework to integrate capacity constraints directly in the recommended intervals between visits in addition to an inadequate IT system. MCF management is interested in understanding if this disjoint decision making process leads to stable patients being scheduled more frequently than necessary thereby resulting in reduced program effectiveness.

7.2.2. Quantitative analysis of the current practice. Based on our qualitative understanding of the current practice, we investigate how the duration between visits quantitatively depends on two key dimensions of patient severity and control type. We conduct an OLS regression of the square root transformation of our dependent variable (duration between visits) against these independent variables and control for whether the previous appointment was kept or not and for the year of the visit. For this analysis, we consider data from the years 2006 to 2009 consisting of 15,859 *appointments* of 2,840 patients since records in this period distinguish between scheduled appointments and actual visits. Our results confirm two main aspects of the qualitative characterization. First, a worse control category is consistently associated with smaller time between appointments. This is consistent with physicians' medically recommended intervals. Second, the time to the next appointment is smaller if the previous appointment was missed than if it was kept. This is consistent with the MCF schedulers' prioritization.

7.2.3. A baseline characterization of the current practice. While the statistical significance for the key independent variables in our OLS regression confirm our qualitative findings, low value of R^2 precludes direct application of the regression results to calculate the duration between visits. Therefore, we construct the following baseline version of MCF's current practice. At each visit, the patient is assigned a follow-up date consistent with the medically recommended intervals. These follow-up dates are modified to ensure feasibility and to capture the current prioritization method at MCF as follows. If the number of patients due for a visit exceeds capacity on a particular day, priority is given to patients who were due back in prior periods but not scheduled due to capacity constraints followed by patients who are due back in the current period. Within each group, patients are selected randomly if the group size exceeds available capacity. If the capacity exceeds the number of patients due for a visit, patients who are due back in future periods are scheduled to use all capacity. In what follows, we refer to this as the *fixed duration* policy.

7.3. Characterizing disease progression

We investigate the drivers of disease progression to obtain an appropriate definition of health state (Section 7.3.1) and estimate transition probabilities over these health states (Section 7.3.2).

7.3.1. Drivers of state transition. We estimate a multinomial regression model to estimate the likelihood of patients moving from one control type to another depending on the underlying severity, time since last appointment, calendar year, season of the year, and whether the school is in session or not. To be consistent with Section 7.2, we focus on patient visits data from 2006 onwards consisting of 7,431 *visits* and 1,473 patients. We find that time since last appointment and severity type have a significant impact on the likelihood of transition among control types. This indicates that the transition probability matrices across control types should be estimated separately for different severity levels. We describe our estimation procedure next.

7.3.2. Estimation of the transition probability matrix. With common observation intervals for all patients, one can estimate the transition probability p_{ij} simply by the ratio of number of patients moving from state i to state j to the total number of patients moving out of state i during that observation interval. However, observation intervals in our data set vary across patients. This can be interpreted as a missing data situation; i.e., we do not observe the state of the patients in the intervening periods. We overcome this obstacle by using the Expectation-Maximization (EM) algorithm (Dempster et al. 1977) that iteratively estimates the missing data and then maximizes the likelihood of observing these data. Specifically, we adopt the approach in Craig and Sendi (2002) to obtain disease progression matrices of discrete time Markov chains as follows.

We fix one month as the *cycle length*, which is the underlying time scale over which state transitions occur. We round the actual observation intervals to integer multiples, n, of the cycle length. For a patient transitioning from state i to state j over n cycles, transitions occurring during the intervening cycles are

Table 1 **Transition probability matrix $\mathcal{P}$ for natural disease progression. C: "controlled", I: "improved, U: "unchanged", W: worsened refer to the control categories described in Section 7.1**

	Mild Intermittent				Mild Persistent				Moderate Persistent				Severe Persistent			
	C	I	U	W	C	I	U	W	C	I	U	W	C	I	U	W
C	0.97	0.01	0.01	0.02												
I	0.00	0.85	0.08	0.07												
U	0.00	0.00	1.00	0.00												
W	0.00	0.00	0.00	1.00												
C					0.93	0.03	0.01	0.02								
I					0.00	0.80	0.12	0.08								
U					0.00	0.00	0.93	0.07								
W					0.00	0.00	0.00	1.00								
C									0.93	0.03	0.02	0.02				
I									0.00	0.79	0.11	0.10				
U									0.00	0.00	0.94	0.06				
W									0.00	0.00	0.00	1.00				
C													0.90	0.04	0.03	0.04
I													0.00	0.81	0.08	0.11
U													0.00	0.00	0.94	0.06
W													0.00	0.00	0.00	1.00

Table 2 **Transition probability matrix $\mathcal{Q}$ for treatment. C: "controlled", I: "improved, U: "unchanged", W: worsened**

	Mild Intermittent				Mild Persistent				Moderate Persistent				Severe Persistent			
	C	I	U	W	C	I	U	W	C	I	U	W	C	I	U	W
C	1.00	0.00	0.00	0.00												
I	0.72	0.28	0.00	0.00												
U	0.70	0.21	0.09	0.00												
W	0.31	0.58	0.07	0.04												
C					1.00	0.00	0.00	0.00								
I					0.59	0.41	0.00	0.00								
U					0.38	0.46	0.17	0.00								
W					0.24	0.67	0.06	0.02								
C									1.00	0.00	0.00	0.00				
I									0.47	0.53	0.00	0.00				
U									0.35	0.45	0.20	0.00				
W									0.25	0.63	0.09	0.03				
C													1.00	0.00	0.00	0.00
I													0.57	0.43	0.00	0.00
U													0.42	0.42	0.16	0.00
W													0.35	0.53	0.09	0.02

unobserved and are treated as missing data. The *expectation* part of the EM algorithm imputes the states at these unobserved cycles and tallies the corresponding expected number of transitions. The *maximization* part of the EM algorithm uses the method of common observation intervals described above to estimate the transition probabilities. These steps are repeated iteratively until the estimates converge. We use different initial values to test the robustness of the algorithm's results.

We use the EM algorithm separately for each severity level to generate four matrices of transition probabilities $\mathcal{P}$ and four treatment matrices $\mathcal{Q}$. Consistent with our modeling assumptions, we restrict $\mathcal{P}$ to be upper triangular and $\mathcal{Q}$ to be a lower triangular matrix to enforce that (i) natural disease progression cannot improve the health state by itself, and (ii) treatment, in isolation, cannot worsen the patients' health status by itself. The estimates are shown below in Table 1 and Table 2. To confirm the Markovian property, we repeat the same analysis with a cycle length of two months and check that the two month transition matrix is similar to the square of the one month transition matrix.

8. Computational study

In this section, we report the results of our extensive computational study with three main objectives. First, we compare the performance of the myopic heuristic with that of the optimal policy (Section 8.2) and Whittle's index (Section 8.3). Second, we quantify the extent of improvement obtained by using the myopic heuristic over the fixed duration policy (Section 8.4). Third, we characterize how the myopic heuristic improves community access to care compared to the fixed duration policy through more effective capacity allocation (Sections 8.5 and 8.6). We start by describing the design of experiments and the associated parameter values in Section 8.1.

8.1. Design of experiments

For our computational experiments, we consider a finite horizon of two years comprising 24 one-month periods. A month is of the same order of magnitude as the typical duration between two successive van visits to a school. Moreover, the recommended interval between patient visits is typically an integer multiple of a month. We assume the discounting factor to be 1.

Our experiments consist of two parts: (i) comparison of the myopic heuristic with the optimal policy and the Whittle's index using dynamic programming and (ii) comparison of the myopic heuristic and the fixed duration policy using simulation. We first describe the choice of parameter values that are common to both parts followed by those that are different. All parameters are listed in Table 3.

Quality of Life (QoL) scores: We utilize the clinical literature to derive estimates for QoL scores. Specifically, we use QoL estimates from Briggs et al. (2006). They use validated questionnaires to obtain survey responses from patients or caregivers and then convert the questionnaire responses into utility scores for four different control types. Since their control types are slightly different from MCF's classification, we construct three alternate sets of QoL scores by fixing the utility levels for the healthiest state (state 0) and worst state (state 3) and interpolating between them. We label the sets as *convex, linear and concave* depending on how the QoL scores change over the health states.

Transition probabilities: For the disease progression $\mathcal{P}$ and treatment $\mathcal{Q}$ matrices, we use the estimated results from Section 7.3 for the four different severity levels.

Table 3 **Parameter values. Columns 3 and 4 reflect the fact that some parameter values might be different in the dynamic programming and simulation computations.**

Parameter	Cases	Dynamic program (Section 8.2)	Simulation (Sections 8.4—8.6)
Number of patients (I)		5	50
Capacity (C)	Low	1	5
	Medium	2	10
	High	3	15
Length of horizon (T)		24	24
History (N)		4	–
Initial distribution of control types	Best	{1, 1, 1, 2}	{10, 10, 10, 20}
	Medium	{0, 1, 1, 3}	{0, 10, 15, 25}
	Worst	{0, 0, 4, 1}	{0, 0, 40, 10}
Duration since last visit		4 for all patients	Distribution over [1,10]
Severity levels	Mild Intermittent	{5, 0, 0, 0}	{9, 23, 15, 3}
	Mild Persistent	{0, 5, 0, 0}	
	Moderate Persistent	{0, 0, 5, 0}	
	Severe Persistent	{0, 0, 0, 5}	
Disease progression matrix ($\mathcal{P}$)		Estimate from MCF data (Table 1)	
Treatment matrix ($\mathcal{Q}$)		Estimate from MCF data (Table 2)	
Quality of Life scores ($\overrightarrow{b}$)	Concave	{0.95, 0.90, 0.84, 0.73}	
	Linear	{0.95, 0.87, 0.80, 0.73}	
	Convex	{0.95, 0.82, 0.76, 0.73}	

History of past visits: In addition to the health state at the previous visit, another dimension of our state space is the time since last visit $n_{i,t} \in \{1, 2, \ldots, N\}$. For a finite time horizon of $T = 24$ periods, $N = T - 1 = 23$. However, due to memory constraints, we limit the history to be within $N = 4$ for our dynamic programming experiments. We do not impose any such restriction for simulation and let $N = T - 1$.

Capacity and cohort size: Due to the curse of dimensionality in dynamic programming, we consider a small cohort of 5 patients and choose capacity levels of 1, 2 and 3 to construct scenarios of varying tightness in capacity. For simulation, we consider a more realistic scenario with 50 patients and capacity levels of 5, 10 and 15 to reflect different patient—capacity ratios.

Initial state: The initial state consists of two main components—control state at the previous visit and time since the last visit. We construct three initial profiles of control types (best, medium, worst) depending on the number of patients in each state. For the dynamic program, we assume that all patients were seen 4 periods ago whereas for the simulation, we assume that patients were seen anywhere between 1 and 10 periods ago according to a general distribution.

Severity levels: For the dynamic program, we study four cohorts, each belonging to a different severity level. For the simulation, we consider one unified cohort with a severity profile that roughly matches the profile of all first visits at MCF.

8.2. Sub-optimality of the myopic heuristic

In this section, we report results of the dynamic program. We follow a three-step procedure to calculate the optimality gap for the myopic heuristic. We calculate the gross objective value (QALYs) corresponding to the optimal policy ($QALY_{opt}$), myopic heuristic $QALY_{myo}$ and a baseline of not scheduling any patients $QALY_{base}$. The net benefit of each policy (QALYs gained) is given by $QALY_{opt} - QALY_{base}$ and $QALY_{myo} - QALY_{base}$. We calculate the optimality gap as the percentage loss in net benefit by using myopic heuristic compared to the optimal policy, given by

$$\delta_1 = \frac{(QALY_{opt} - QALY_{base}) - (QALY_{myo} - QALY_{base})}{QALY_{opt} - QALY_{base}} = \frac{QALY_{opt} - QALY_{myo}}{QALY_{opt} - QALY_{base}}.$$

Table 4 shows various statistics of this optimality gap for 108 instances (36 instances for each capacity level) of the problem corresponding to different combinations of initial states, QoL scores and severity levels. This leads to several important observations.

Observation 1 (Overall performance) *Over all instances, the average optimality gap is 0.40%. Moreover, the maximum optimality gap is less than 3% and in more than 80% of the instances (88 out of 108), the optimality gap is less than 1%.*

Observation 2 (Impact of capacity) *The performance of the myopic heuristic improves as the capacity becomes less tight, as can be seen from the average and maximum gaps.*

Table 4 Measures of optimality gap for the myopic heuristic at different values of capacity. For each capacity level, the numbers reflect average over 36 instances corresponding to all combinations of 3 initial states and 3 set of QoL values and 4 severity levels.

Performance measure	Overall	Capacity 1	Capacity 2	Capacity 3
Average gap	0.40%	1.0%	0.16%	0.01%
Maximum gap	2.60%	2.60%	0.57%	0.05%
Instances with gap$\leq 5\%$	100%	100%	100%	100%
Instances with gap$\leq 2\%$	97.22%	91.67%	100%	100%
Instances with gap$\leq 1\%$	81.48%	44.44%	100%	100%

8.3. Comparison of the myopic heuristic with Whittle's index

As discussed in Section 3, our capacity allocation problem is an instance of the RMAB problem, which has been widely solved in the literature using Whittle's index. Hence, we benchmark the performance of the myopic heuristic with that of Whittle's index for the instances in Table 4. The average gap for Whittle's index is 0.22% across all instances, with average gaps of 0.24%, 0.30%, and 0.11%, for capacity levels 1, 2, and 3, respectively. Thus, the performances of the myopic heuristic and Whittle's index are quite comparable, with the latter being slightly better for smaller capacities and the former being slightly better for larger capacities.

Although yielding similar performance, Whittle's index is significantly more difficult to compute than the myopic heuristic. First, Whittle's index involves repeated solution of a (single patient) dynamic program through a value iteration algorithm, which requires an advanced computing environment such as MATLAB® or C++. On the other hand, the myopic heuristic can be easily implemented in Microsoft Excel®. Second, the time required to generate Whittle's index for all states is significantly larger than that for the myopic index, especially for problem instances with more than 10 health states and increases rapidly with the number of health states thereafter. As an example, for an instance with 16 health states and a time horizon of 24 months (periods), it takes 32 hours to generate Whittle's index whereas it takes less than 1 second to generate the myopic index. [1] These computational issues are likely to create barriers to implementation of Whittle's index in a nonprofit organization such as MCF that does not have the requisite monetary and human resources to manage this level of complexity. In addition to the computational simplicity, the easy interpretation of the myopic index–the benefit foregone for a particular patient by not seeing her in the next period–can further improve the chances of a successful implementation. Hence, we propose the myopic heuristic as a new approach to capacity allocation at MCF and evaluate its performance relative to a fixed duration policy in the subsequent sections.

8.4. Improvement over the fixed duration policy using the myopic heuristic

We report on the magnitude of potential improvement from implementing the myopic heuristic over the fixed duration policy for 27 combinations of capacity, QoL scores and initial states (Table 5). Similar to the optimality gap in Section 8.2, we calculate this improvement as:

$$\delta_2 = \frac{(QALY_{myo} - QALY_{base}) - (QALY_{fix} - QALY_{base})}{QALY_{fix} - QALY_{base}} = \frac{QALY_{myo} - QALY_{fix}}{QALY_{fix} - QALY_{base}}.$$

Observation 3 (Overall improvement) *MCF can obtain significant benefit by implementing the myopic heuristic, ranging between 3% and 16% for the parameter values considered.*

It is worth noting that potential improvement in health outcomes is driven solely by more effective operational decisions. Higher effectiveness is achieved in our approach by utilizing both clinical (disease progression) and operational (capacity constraint) data to make operational decisions in an integrated manner. This is in contrast to the decoupled approach followed currently at MCF described in Section 7.2. MCF might incur some nonmonetary adjustment costs in adopting the proposed policy and some upfront investment to modify their electronic medical records system. However, this change does not require additional recurring cost. Hence, the incremental cost-effectiveness of our recommendation, the standard criterion used to evaluate health care interventions, is quite attractive. Moreover, it does not require changes to the core clinical practice such as therapy, diagnosis etc. and hence would receive less resistance from physicians.

[1] All computations are performed on an AMD Opteron CPU core running at 2.8Ghz with 4 GB of memory.

Table 5 Performance improvement over the fixed duration policy obtained by using the myopic heuristic

QoL	Initial state	Capacity 5	10	15
Convex	Worst	11.33%	9.11%	5.40%
	Medium	9.86%	8.75%	5.39%
	Best	13.89%	11.02%	6.26%
Linear	Worst	12.74%	8.00%	4.06%
	Medium	12.34%	7.77%	3.97%
	Best	14.49%	8.98%	4.46%
Concave	Worst	14.28%	7.39%	3.30%
	Medium	14.22%	7.56%	3.20%
	Best	15.47%	8.09%	3.48%

Impact of capacity: Oowing to economic downturn and consequent reduction in funding, MCF has reduced the number of vans from 3 to 2 in 2009, thus significantly constraining their capacity relative to the patient population. Motivated by this event, we investigate the improvement in performance obtained from the myopic heuristic for different capacity levels.

Observation 4 (Impact of capacity) *Potential benefit from implementing the myopic heuristic is higher for tighter capacity and decreases as capacity increases.*

This observation implies that, when capacity tightens, MCF can incur a significant loss of welfare with the fixed duration policy (recommended medical duration between visits). Notably, the gap between the myopic heuristic and the fixed duration policy is much more sensitive to changes in capacity than the gap between the optimal policy and myopic heuristic. This is driven by the fact that the myopic heuristic inherently adjusts the frequency of visits to varying capacity levels and prioritizes patients by control status when capacity is limited, whereas such an adjustment is *post facto* in the fixed duration policy only to ensure feasibility. We explore this issue further in Section 8.6.

Sensitivity to QoL scores: The three sets of QoL scores implicitly reflect different definitions of control types. For instance, the convex QoL set reflects the fact that the best control type enjoys a much higher quality of life than all other control types, which are more similar to each other. On the other extreme, the concave QoL set reflects the fact that the worst control type enjoys a much lower quality of life than all other control types.

Observation 5 (Impact of QoL scores) *Performance improvement for the concave QoL set is higher (than linear or convex set) for smaller capacity cases but lower for larger capacity cases.*

Our implementation of MCF's current policy recommends equal duration between visits for the three worse states and a much longer duration for the best state. Thus, we expect that the potential benefit should be higher under concave QoL set than convex or linear QoL set. The above observation confirms this intuition for small capacity cases.

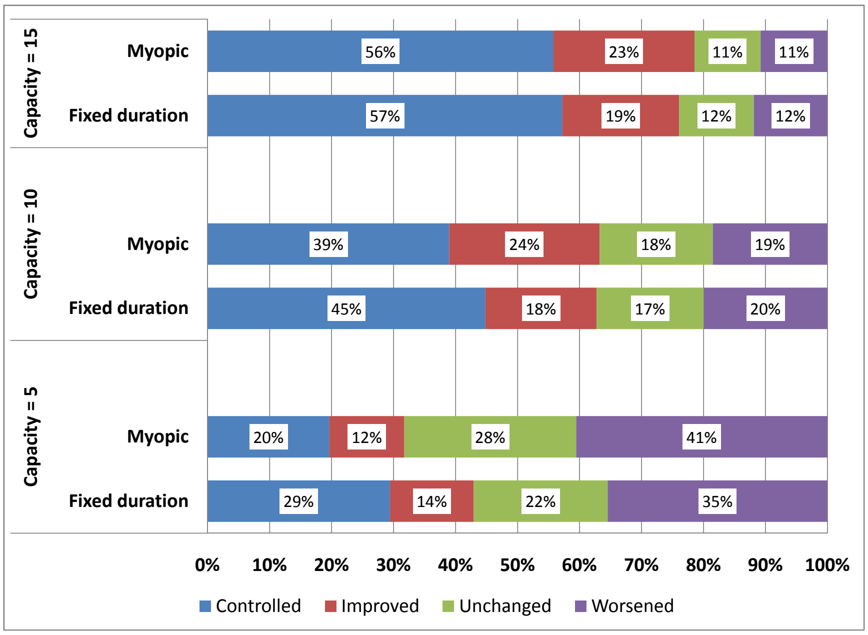
Sensitivity to initial state: Table 5 shows that the magnitude of potential improvement is similar across different initial states (control types) of the patient population. The general trend indicates that, per intuition, the magnitude of improvement is mostly higher if more patients start in a worse health state. This is summarized in the following observation.

Observation 6 (Sensitivity of initial state) *Performance improvement is robust to changes in the initial state; i.e., distribution of patients' health states at previous visits.*

8.5. Improved capacity allocation

To understand the drivers of potential improvement, we investigate how the fixed duration policy policy and the myopic heuristic allocate available capacity among different control types.

Figure 2 **Capacity allocation between patients with different control types (summed across severity levels) under myopic heuristic and the fixed duration policy**



Observation 7 (Capacity allocation) *The myopic heuristic allocates less capacity to patients in better health states and more to those in worse health states.*

Figure 2 illustrates one such set of results for convex QoL and worst initial state. When capacity is very tight, the myopic heuristic generates value by significantly increasing the capacity allocation to the worst control types—worsened and unchanged—and commensurately decreasing allocation to the best types. When capacity is relatively abundant, the difference in allocation between the myopic heuristic and fixed duration policy decreases.

8.6. Improved patient access to care

We further explore how the myopic heuristic achieves a more effective capacity allocation as described above. Intuitively, one might expect that compared to the fixed duration policy, the myopic heuristic results in higher frequency of visits for patients with worse health states and lower frequency for patients with better health states. Table 6 shows that this intuition is true when capacity is not very tight, as described in the observation below.

Table 6 Access characteristics of the myopic heuristic and the fixed duration policy

Capacity	Health state	Average duration between visits: Current (intended)	Current (actual)	Myopic	Fraction of patients not seen: Current	Myopic
5	Controlled	3	7.87	15.71	5.00%	0.10%
	Improved	1	6.34	11.35		
	Unchanged	1	6.30	9.06		
	Worsened	1	6.18	6.90		
10	Controlled	3	5.72	9.05	0.02%	0.00%
	Improved	1	3.91	4.44		
	Unchanged	1	3.95	3.28		
	Worsened	1	3.96	2.34		
15	Controlled	3	4.25	5.29	0.00%	0.00%
	Improved	1	2.38	2.00		
	Unchanged	1	2.42	1.75		
	Worsened	1	2.44	1.47		

Observation 8 (Access – large capacity) *When capacity is not very tight, the duration between visits under myopic heuristic is larger for better health states and smaller for worse health states compared to the fixed duration policy.*

This observation confirms the intuition of MCF management that the fixed duration policy results in more frequent visits for the healthier patients than might be necessary. However, the intuition needs to be modified when the capacity is very tight. In this case, the fixed duration policy fails to see a nontrivial fraction of the patients because of the rigidity in the recommended visit intervals. In contrast, the myopic heuristic accommodates these additional patients by increasing the visit interval for all patient segments.

Observation 9 (Access – small capacity) *When capacity is very tight, the fraction of patients without any appointment is lower under myopic heuristic compared to that under the fixed duration policy.*

In summary, we find that the myopic heuristic achieves near-optimal QALY benefits and has significant potential to improve access to care for MCF operations. This is particularly important in the light of the fact that MCF had to discontinue the operation of one van due to budgetary limitations thus making the capacity even tighter than before.

9. Conclusion and future work

In this paper, we analyze decisions related to the allocation of appointment slots in community-based health care for chronic diseases in a non-profit setting. We develop an integrated approach to capacity allocation that incorporates a clinical model of disease progression with the operational dynamics. We formulate this decision problem as a stochastic dynamic program, where the available capacity is allocated among patients of different health states whose transitions between states are governed by discrete-time Markov chains.

We show that the optimal policy for a special case of the problem has a relatively simple structure, and, based on these insights, we develop a myopic heuristic that performs very well for the more general version of the problem. In deriving the structure of the optimal policy, we provide a quantitative characterization of "better" and "worse" health states in the absence of perfect observability of each patient's health state in every period. Applying our results to data from a mobile health care provider shows that significant improvement in health outcomes can be obtained over the current practice of specifying fixed durations between visits for patients of different health states. Our proposed policy achieves this improvement by flexibly adjusting the visit durations to accommodate capacity constraints and health states of the entire patient population. In addition, this flexibility also improves access; i.e., it reduces the number of patients never seen. These results indicate that significant health benefits can be obtained by improving capacity allocation decisions in community-based chronic care settings.

Our model and results provide several interesting avenues for future research. In our model, we implicitly assume that a higher level allocation of capacity between new and returning patients has been made and hence we consider only returning patients. An extension of our model would be one that considers patients as two dimensional constructs characterized by health state and new/returning status. This would make the patient pool size endogenous to past decisions. Results from Jones et al. (2007) suggest that the Markovian properties of disease progression hold for returning patients, but not for new patients.

We assume that all scheduled patients show up according to their appointment schedule. Another extension would be one in which a fraction of patients do not show up, leading to wasted capacity. While some models in literature account for no-shows, none incorporate disease progression dynamics, which can influence no-show behavior (Deo et al. 2009).

The mobile care delivery model motivates an entire spectrum of operational questions that have not been explored in the literature. These include (i) joint capacity allocation between different locations and between different patient classes within a location, (ii) assignment of mobile units to different regions of the target population. These questions are gaining importance from the perspective of global health policy because of the growing popularity of community-based health care delivery in developing countries. Recently, one of the authors was approached to help design a mobile health care system to target chronic conditions in Ghana (Culhane et al. 2010).

In conclusion, our results highlight the significant potential to improve health outcomes by making better operational decisions. This adds a new dimension to the literature on health care operations management that has focused primarily on the efficiency gains obtained by improving patient flow in hospitals and appointment scheduling in conventional clinics (physician offices).

References

L. Akinbami. The state of childhood asthma, united states, 1980-2005. Technical report, National Center for Health Statistics, 2006.

O. Alagoz, L. Maillart, A. Schaefer, and M. Roberts. The optimal timing of transplant living-donor liver transplantation. *Operations Research*, 53(3):443–455, 2005.

P. Ansell, K. Glazebrook, J. Nino-Mora, and M. O'Keeffe. Whittle's index policy for a multi-class queueing system with convex holding costs. *Mathematical Methods of Operations Research*, 57(1):21–39, 2003.

D. Bertsimas and J. Nino-Mora. Restless bandits, linear programming relaxations, and a primal-dual index heuristic. *Operations Research*, 48(1):80–90, 2000.

M. Brandeau, F. Sainfort, and W. Pierskalla. *Operatons research and health care: A handbook of methods and applications*. International series in operations research & management science. Kluwer Academic Publishers, Boston, MA, 2004.

A. Brennan, S. Chick, and R. Davies. A taxonomy of model structures for economic evaluation of health technologies. *Health Economics*, 15(12):1295–1310, 2006.

A. Briggs, J. Bousquet, M. Wallace, W. Busse, T. Clark, S. Pedersen, and E. Bateman. Cost-effectiveness of asthma control: an economic appraisal of the GOAL study. *Allergy*, 61(5):531–536, 2006.

T. Cayirli and H. Veral. Designing appointment systems for ambulatory care services. *Health Care Management Science*, 9(1):47–58, 2006.

B. Craig and P. Sendi. Estimation of the transition matrix of a discrete-time Markov chain. *Health Economics*, 11(1): 33–42, 2002.

D. Culhane, G. Dasovich, R. Farrell, and U. Moka. Mobile healthcare in ghana. Technical report, Northwestern University, 2010.

A. Dempster, N. Laird, D. Rubin, et al. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.

B. Denton and D. Gupta. A sequential bounding approach for optimal appointment scheduling. *IIE Transactions*, 35 (11):1003–1016, 2003.

S. Deo, K. Smilowitz, T. Asvatanakul, J. Kim, A. Kelly, J. Schieneson, and K. Stuewer. Mobile C.A.R.E. case study. *Kellogg School of Management*, 2009.

K. Glazebrook, H. Mitchell, and P. Ansell. Index policies for the maintenance of a collection of machines by a set of repairmen. *European Journal of Operational Research*, 165(1):267–284, 2005a.

K. Glazebrook, D. Ruiz-Hernandez, and C. Kirkbride. Some indexable families of restless bandit problems. *Advances in Applied Probability*, pages 643–672, 2006.

K. D. Glazebrook, H. M. Mitchell, and P. S. Ansell. Index policies for the maintenance of a collection of machines by a set of repairmen. *European Journal of Operational Research*, 165(1):267–284, 2005b.

D. Gupta and B. Denton. Appointment scheduling in health care: Challenges and opportunities. *IIE Transactions*, 40 (9):800–819, 2008.

D. Gupta and L. Wang. Dynamic and static job allocation for multiserver systems. *Operations Research*, 56:576–592, 2008.

R. Gupta, X. Zhang, L. Sharp, J. Shannon, and K. Weiss. Geographic variability in childhood asthma prevalence in Chicago. *Journal of Allergy and Clinical Immunology*, 121(3):639–645, 2008.

C. Jones, L. Clement, T. Morphew, K. Kwong, J. Hanley-Lopez, F. Lifson, L. Opas, and J. Guterman. Achieving and maintaining asthma control in an urban pediatric disease management program: the Breathmobile Program. *The Journal of Allergy and Clinical Immunology*, 119(6):1445–1453, 2007.

B. Kitch, G. Chew, H. Burge, M. Muilenberg, S. Weiss, T. Platts-Mills, G. OConnor, and D. Gold. Socioeconomic predictors of high allergen levels in homes in the greater Boston area. *Environmental Health Perspectives*, 108: 301–307, 2000.

D. Lee and S. Zenios. Optimal capacity overbooking for the regular treatment of chronic conditions. *Operations Research*, 57(4):852–865, 2009.

K. Liu and Q. Zhao. On the myopic policy for a class of restless bandit problems with applications in dynamic multichannel access. In *Proceedings of the 48th IEEE Conference on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009.*, pages 3592–3597. IEEE, 2009.

N. Liu, S. Ziya, and V. Kulkarni. Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. *Manufacturing & Service Operations Management*, 12(2):347–364, 2010.

W. S. Lovejoy. Some monotonicity results for partially observed markov decision processes. *Operations Research*, 35 (5):736–743, 1987.

D. Marder, P. Targonski, P. Orris, V. Persky, and W. Addington. Effect of racial and socioeconomic factors on asthma mortality in Chicago. *Chest*, 101:4265S–4429S, 1992.

M. Murray and C. Tantau. Redefining open access to primary care. *Managed Care Quarterly*, 7:45–55, 1999.

N. Oriol, P. Cote, A. Vavasis, J. Bennet, D. DeLorenzo, P. Blanc, and I. Kohane. Calculating the return on investment of mobile healthcare. *BMC Medicine*, 7(27), 2009.

A. Paltiel, K. Kuntz, and A. Weiss, S.T.and Fuhlbrigge. *Operations Research and Health Care*, chapter An Asthma Policy Model, pages 661–694. Springer, 2006.

C. Papadimitriou and J. Tsitsiklis. The complexity of optimal queueing network control. In *Proceedings of the Ninth Annual Structure in Complexity Theory Conference, 1994.*, pages 318–322. IEEE, 1994.

H. Pham and H. Wang. Imperfect maintenance. *European Journal of Operational Research*, 94(3):425–438, 1996.

P. Post. Mobile health care for homeless people: Using vehicles to extend care. Technical report, National Health Care for the Homeless Council, 2007.

M. Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. John Wiley & Sons, Inc. New York, NY, USA, 1994.

L. Robinson and R. Chen. Scheduling doctors' appointments: optimal and empirically-based heuristic policies. *IIE Transactions on Scheduling and Logistics*, 35:295–307, 2003.

L. Robinson and R. Chen. A comparison of traditional and open-access policies for appointment scheduling. *Manufacturing & Service Operations Management*, 12(2):330–346, 2010.

D. Rosenfield. Markovian deterioration with uncertain information. *Operations Research*, 24(1):141C155, 1976.

A. Shahani, N. Korve, K. Jones, and D. Paynton. Towards an operational model for prevention and treatment of asthma attacks. *Journal of the Operational Research Society*, 45(8):916–926, 1994.

M. Shaked and J. Shanthikumar. *Stochastic orders (Springer Series in Statistics)*. Springer Science+ Business Media, LLC, 2007.

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

F. Sonnenberg and J. Beck. Markov models in medical decision making. *Medical Decision Making*, 13(4):322, 1993.

C. Valdez-Flores and R. Feldman. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics*, 36(4):419–446, 1989.

H. Wang. A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139 (3):469–489, 2002.

R. Weber and G. Weiss. On an index policy for restless bandits. *Journal of Applied Probability*, pages 637–648, 1990.

P. Whittle. Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability*, 25:287–298, 1988.

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

Appendix

A. Proofs of theoretical results in Section 5.1

Before proving Proposition 1 and Theorem 1, we establish an auxiliary result and derive a more explicit version of the optimality equation in (3a) and (3b).

LEMMA 1. *If $b_0 > b_1$, the MDP defined by (3) is equivalent to the maximization problem where the single period reward is given by $\tilde{\phi}(\pi_{i,t}) = \pi_{i,t}$.*

Proof for Lemma 1

The result follows directly by noting that taking $b_0 = 1$ and $b_1 = 0$ gives exactly the same optimal solution. The proof is omitted here to save space. ∎

Establishing the optimality equations

Our objective here is to rewrite the optimality equation (3a) and (3b) in a more explicit form that is useful in proving Proposition 1 and Theorem 1.

Consider I patients with *information* $(\pi_{1,t}, \pi_{2,t}, \ldots, \pi_{I,t})$ at the beginning of Period $t \in \{1, 2, \ldots, T\}$. Let S denote the set of the indices of all patients that are scheduled in Period t and let $\overline{S}$ denote the set of indices of all patients who are not scheduled.

Without loss of generality assume $S = \{1, 2, \ldots, C\}$ and $\overline{S} = \{C+1, C+2, \ldots, I\}$. We first establish *information* $\pi_{i,t+1}$ at the beginning of Period $t+1$ for Patient i.

For $k \in \{0,1\}$, define γ_k such that

$$(\gamma_k, 1-\gamma_k) = \vec{e}_k \mathcal{QP}, \qquad \forall k \in \{0,1\} \tag{8}$$

using (4), we can interpret γ_k as the probability that Patient i is in health state k at the beginning of Period $t+1$, given that he is in health state k at the beginning of Period t and is scheduled in Period t. Using (8), given any realization of the real health state $x_{i,t}$, Patient i's *information* at the beginning of Period $t+1$ is

$$\pi_{i,t+1} = \gamma_0(1-x_{i,t}) + \gamma_1 x_{i,t} \tag{9}$$

For example, if Patient i's real heath state at the beginning of Period t is 0, then the *information* for Patient i is $\gamma_0 = \gamma_0 \times (1-0) + \gamma_1 \times 0$ at the beginning of Period $t+1$.

For an unscheduled Patient $i \in \overline{S}$ with *information* $\pi_{i,t}$ at the beginning of Period t, define $\pi_{i,t+1} = z(\pi_{i,t})$ for his *information* at the beginning of Period $t+1$, where

$$z(\pi_{i,t}) = \pi_{i,t} p_0 + (1-\pi_{i,t}) p_1 \tag{10}$$

Therefore equations (3a) and (3b) can be written as:

$$u_t^*(\pi_{1,t}, \pi_{2,t}, \dots, \pi_{I,t}) = \sum_{i=1}^{I} \pi_i + \beta \times$$
$$\left\{ \sum_{x_{i,t} \in \{0,1\}, i \in S} \prod_{i \in S} (\pi_{i,t})^{1-x_{i,t}} (1-\pi_{i,t})^{x_{i,t}} \right.$$
$$\left. \Big[u_{t+1}^*\big(\gamma_0(1-x_{1,t}) + \gamma_1(x_{1,t}) \ , \ \dots \ , \ \gamma_0(1-x_{C,t}) + \gamma_1(x_{C,t}), z(\pi_{C+1,t}) \ , \ \dots \ , \ z(\pi_{I,t})\big) \Big] \right\} \tag{11a}$$
$$u_T^*(\pi_{1,T}, \pi_{2,T}, \dots, \pi_{I,T}) = \sum_{i=1}^{I} \pi_{i,T} \tag{11b}$$

where summation $\sum_{x_{i,t} \in \{0,1\}, i \in S}$ includes all possible realizations of $(x_{1,t}, x_{2,t}, \dots, x_{C,t})$.

Proof for Proposition 1

The proofs for both monotonicity and convexity use induction. For both of the proofs we consider two systems in which only one patient's state differs. Then we separate the discussion by two cases: whether this patient is scheduled or not under one of the systems. We only show the part that proves convexity here to save space.

(ii) We prove that $u_t^*(\pi_{1,t}, \pi_{2,t}, \dots, \pi_{I,t})$ is convex in each $\pi_{j,t}, \forall j \in \{1,2,\dots,I\}$ and all $t \in \{1,2\dots,T\}$. By definition we need to show that $\forall t \in \{1,2,\dots,T\}$, $\forall \alpha \in [0,1]$ and $\pi_{j,t}, \pi'_{j,t} \in [0,1]$, the following inequality holds:

$$u_t^*(\pi_{1,t}, \dots, \alpha\pi_{j,t} + (1-\alpha)\pi'_{j,t}, \dots, \pi_{I,t}) \leq$$
$$\alpha u_t^*(\pi_{1,t}, \dots, \pi_{j,t}, \dots, \pi_{I,t}) + (1-\alpha) u_t^*(\pi_{1,t}, \dots, \pi'_{j,t}, \dots, \pi_{I,t}) \tag{12}$$

We use induction to prove (12).

We initialize the induction in Period T. From (11b), it is clear that $u_T^*(\pi_{1,T}, \pi_{2,T}, \dots, \pi_{I,T}) = \sum_{i=1}^{I} \pi_{i,T}$ is linear, and is therefore convex in each $\pi_{i,T}$.

Now assume that $u_{t+1}^*(\pi_{1,t+1}, \pi_{2,t+1}, \dots, \pi_{I,t+1})$ is convex in each of its element. Consider the system where *information* of all patients in the system at the beginning of Period t is $(\pi_{1,t}, \dots, \alpha\pi_{j,t} + (1-\alpha)\pi'_{j,t}, \dots, \pi_{I,t})$. Suppose that $S = \{1,2,\dots,C\}$. Then consider two cases depending on whether Patient j is scheduled in this system or not.

Case(a): Patient j with information $\alpha\pi_j + (1-\alpha)\pi'_j$ is scheduled in Period t.

Define random vector $\vec{\boldsymbol{\mu}}_{j,t+1}$ for the *information* of all patients except for Patient j. Then we obtain:

$$\begin{aligned}
&u_t^*(\pi_{1,t},\ldots,\alpha\pi_{j,t}+(1-\alpha)\pi'_{j,t},\ldots,\pi_{I,t}) \\
&= \alpha\pi_{j,t} + \alpha\sum_{i=1,i\neq j}^{I}\pi_{i,t}\beta\alpha\mathbb{E}_{\vec{\boldsymbol{\mu}}_{j,t+1}}\Big[\pi_{j,t}u_{t+1}^*(\gamma_0,\vec{\boldsymbol{\mu}}_{j,t+1}) + (1-\pi_{j,t})u_{t+1}^*(\gamma_1,\vec{\boldsymbol{\mu}}_{j,t+1})\Big] \\
&+ (1-\alpha)\pi'_{j,t} + (1-\alpha)\sum_{i=i,i\neq j}^{I}\pi_{i,t} + \beta(1-\alpha)\mathbb{E}_{\vec{\boldsymbol{\mu}}_{j,t+1}}\Big[\pi'_{j,t}u_{t+1}^*(\gamma_0,\vec{\boldsymbol{\mu}}_{j,t+1}) + (1-\pi'_{j,t})u_{t+1}^*(\gamma_1,\vec{\boldsymbol{\mu}}_{j,t+1})\Big] \qquad (13)
\end{aligned}$$

Moreover, notice that,

$$u_t^*(\pi_{1,t},\ldots,\pi_{j,t},\ldots,\pi_{I,t}) \geq \pi_{j,t} + \sum_{i=1,i\neq j}^{I}\pi_{i,t} + \beta\mathbb{E}_{\vec{\boldsymbol{\mu}}_{j,t+1}}\Big[\pi_{j,t}u_{t+1}^*(\gamma_0,\vec{\boldsymbol{\mu}}_{j,t+1}) + (1-\pi_{j,t})u_{t+1}^*(\gamma_1,\vec{\boldsymbol{\mu}}_{j,t+1})\Big] \qquad (14)$$

where the left hand side is the value of system when system state is $(\pi_{1,t},\ldots,\pi_{j,t},\ldots,\pi_{I,t})$ at Period t; and the right hand side is the discounted present value if the action in Period t is to schedule Patient $1,2,\ldots,C$ and from Period $t+1$ on is governed by optimal policy.

Similarly,

$$u_t^*(\pi_{1,t},\ldots,\pi'_{j,t},\ldots,\pi_{I,t}) \geq \pi'_{j,t} + \sum_{i=1,i\neq j}^{I}\pi_{i,t} + \beta\mathbb{E}_{\vec{\boldsymbol{\mu}}_{j,t+1}}\Big[\pi'_{j,t}u_{t+1}^*(\gamma_0,\vec{\boldsymbol{\mu}}_{j,t+1}) + (1-\pi'_{j,t})u_{t+1}^*(\gamma_1,\vec{\boldsymbol{\mu}}_{j,t+1})\Big] \qquad (15)$$

Multiply equation (14) by α and (15) by $(1-\alpha)$ and sum them up, the resulting right hand side is equivalent to the right hand side of equation (13). Then we are done.

Case (b): Patient j with information $\alpha\pi_{j,t} + (1-\alpha)\pi'_{j,t}$ is not scheduled in Period t.

Similarly define random vector $\vec{\boldsymbol{\mu}}'_{j,t+1}$. The compact optimality equation gives

$$\begin{aligned}
&u_t^*(\pi_{1,t},\ldots,\alpha\pi_{j,t}+(1-\alpha)\pi'_{j,t},\ldots,\pi_{I,t}) \\
&\quad\leq \alpha\pi_{j,t} + \alpha\sum_{i=1,i\neq j}^{I}\pi_{i,t} + \beta\alpha\mathbb{E}_{\vec{\boldsymbol{\mu}}'_{j,t+1}}\Big[u_{t+1}^*\Big(z(\pi_{j,t}),\vec{\boldsymbol{\mu}}'_{j,t+1}\Big)\Big] \\
&\quad+ (1-\alpha)\pi'_{j,t} + (1-\alpha)\sum_{i=1,i\neq j}^{I}\pi_{i,t} + \beta(1-\alpha)\mathbb{E}_{\vec{\boldsymbol{\mu}}'_{j,t+1}}\Big[u_{t+1}^*\Big(z(\pi'_{j,t}),\vec{\boldsymbol{\mu}}'_{j,t+1}\Big)\Big] \qquad (16)
\end{aligned}$$

Inequality (16) holds because $u_{t+1}^*(\cdot,\ldots,\cdot)$ is componentwise convex by induction assumption, and because expectation is a linear operator.

By definition of the optimality equation and an argument similar to equations (14) and (15),

$$u_t^*(\pi_{1,t},\ldots,\pi_{j,t},\ldots,\pi_{I,t}) \geq \pi_{j,t} + \sum_{i=1,i\neq j}^{I}\pi_{i,t} + \beta\mathbb{E}_{\vec{\boldsymbol{\mu}}'_{j,t+1}}\Big[u_{t+1}^*(z(\pi_{j,t}),\vec{\boldsymbol{\mu}}'_{j,t+1})\Big] \qquad (17)$$

$$u_t^*(\pi_{1,t},\ldots,\pi'_{j,t},\ldots,\pi_{I,t}) \geq \pi'_{j,t} + \sum_{i=1,i\neq j}^{I}\pi_{i,t} + \beta\mathbb{E}_{\vec{\boldsymbol{\mu}}'_{j,t+1}}\Big[u_{t+1}^*(z(\pi'_{j,t}),\vec{\boldsymbol{\mu}}'_{j,t+1})\Big] \qquad (18)$$

Therefore, multiply equation (17) by α and (18) by $(1-\alpha)$ and sum them up, we complete the proof. ∎

Proof for Theorem 1

Recall the definition of set S and $\overline{S}$ from before.Theorem 1 claims that for any pair of indices $j \in S$ and $j' \in \overline{S}$,it must be that $\pi_{j,t} \leq \pi_{j',t}$.

We prove the theorem by contradiction. If the theorem is not true, then there exists an optimal policy Ω under which $j \in S$ and $j' \in \overline{S}$ but $\pi_{j,t} > \pi_{j',t}$. Suppose without loss of generality, $S = \{1, 3, \ldots, C+1\}$ and $\overline{S} = \{2, C+2, \ldots, I\}$. Moreover, let $j = 1$ and $j' = 2$. We construct an alternate policy Ω', which schedules Patient 2 instead of Patient 1 in Period t and follows the optimal policy from Period $t+1$ to the end of horizon. We show that policy Ω cannot be optimal when $\pi_{1,t} > \pi_{2,t}$, by showing that following policy Ω' yields more benefit than Ω.

Consider a random vector $\vec{\boldsymbol{\mu}}_{12,t+1}$ which has realizations $\vec{\mu}_{12,t+1}$.

$$\vec{\mu}_{12,t+1} = \Big(\gamma_0(1-x_{3,t}) + \gamma_1(x_{3,t})\ ,\ \ldots\ ,\ \gamma_0(1-x_{C+1,t}) + \gamma_1(x_{C+1,t}), z(\pi_{C+2,t})\ ,\ \ldots\ ,\ z(\pi_{I,t})\Big) \tag{19}$$

of probability $\Pi_{i \in S\setminus\{1\}}(\pi_{i,t})^{1-x_{i,t}}(1-\pi_{i,t})^{x_{i,t}}$.

If $1 \in S$ and $2 \in \overline{S}$ in Period t and optimal policy is followed from Period $t+1$ till the end of horizon, we define the value of following policy Ω as:

$$\begin{aligned} u_t(\pi_{1,t}, \ldots, \pi_{I,t}; 1 \in S, 2 \in \overline{S}) = &\sum_{i=1}^{I} \pi_{i,t} \\ &+ \beta \mathbb{E}_{\vec{\boldsymbol{\mu}}_{12,t+1}} \Big[\pi_{1,t} u^*_{t+1}(\gamma_0, z(\pi_{2,t}), \vec{\boldsymbol{\mu}}_{12,t+1}) + (1-\pi_{1,t}) u^*_{t+1}(\gamma_1, z(\pi_{2,t}), \vec{\boldsymbol{\mu}}_{12,t+1})\Big] \end{aligned} \tag{20}$$

If following policy Ω', i.e., scheduling Patient 2 instead of Patient 1 in Period t and following optimal policy from Period $t+1$ on, then the value will be:

$$\begin{aligned} u_t(\pi_{1,t}, \ldots, \pi_{I,t}; 2 \in S, 1 \in \overline{S}) = &\sum_{i=1}^{I} \pi_{i,t} \\ &+ \beta \mathbb{E}_{\vec{\boldsymbol{\mu}}_{12,t+1}} \Big[\pi_{2,t} u^*_{t+1}(\gamma_0, z(\pi_{1,t}), \vec{\boldsymbol{\mu}}_{12,t+1}) + (1-\pi_{2,t}) u^*_{t+1}(\gamma_1, z(\pi_{1,t}), \vec{\boldsymbol{\mu}}_{12,t+1})\Big] \end{aligned} \tag{21}$$

Now we show that $u_{t+1}(\pi_{1,t}, \ldots, \pi_{I,t} | 1 \in S, 2 \in \overline{S}) \le u_{t+1}(\pi_{1,t}, \ldots, \pi_{I,t} | 2 \in S, 1 \in \overline{S})$ by showing that for any given realization $\vec{\mu}_{j,t+1}$, the following inequality holds,

$$\begin{aligned} &\pi_{1,t} u_{t+1}\big(\gamma_0, z(\pi_{2,t}), \vec{\mu}_{12,t+1}\big) + (1-\pi_{1,t}) u_{t+1}\big(\gamma_1, z(\pi_{2,t}), \vec{\mu}_{12,t+1}\big) \le \\ &\qquad \pi_{2,t} u_{t+1}\big(\gamma_0, z(\pi_{1,t}), \vec{\mu}_{12,t+1}\big) + (1-\pi_{2,t}) u_{t+1}\big(\gamma_1, z(\pi_{1,t}), \vec{\mu}_{12,t+1}\big) \end{aligned} \tag{22}$$

Define $f((\pi_{2,t}, \pi_{1,t}); \vec{\mu}_{12,t+1})$, $f((\pi_{1,t}, \pi_{2,t}); \vec{\mu}_{j,t+1})$ and a univariate function $g(\pi_{2,t}; \pi_{1,t}, \vec{\mu}_{12,t+1})$ as:

$$\begin{aligned} f((\pi_{2,t}, \pi_{1,t}); \vec{\mu}_{12,t+1}) &= \pi_{2,t} u^*_{t+1}(\gamma_0, z(\pi_{1,t}), \vec{\mu}_{12,t+1}) + (1-\pi_{2,t}) u^*_{t+1}(\gamma_1, z(\pi_{1,t}), \vec{\mu}_{12,t+1}) \\ f((\pi_{1,t}, \pi_{2,t}); \vec{\mu}_{12,t+1}) &= \pi_{1,t} u^*_{t+1}(\gamma_0, z(\pi_{2,t}), \vec{\mu}_{12,t+1}) + (1-\pi_{1,t}) u^*_{t+1}(\gamma_1, z(\pi_{2,t}), \vec{\mu}_{12,t+1}) \\ g(\pi_{2,t}; \pi_{1,t}, \vec{\mu}_{12,t+1}) &= f((\pi_{2,t}, \pi_{1,t}); \vec{\mu}_{12,t+1}) - f((\pi_{1,t}, \pi_{2,t}); \vec{\mu}_{12,t+1}) \end{aligned} \tag{23}$$

Then to prove inequality (22) is equivalent to prove that $g(\pi_{2,t}; \pi_{1,t}, \vec{\mu}_{12,t+1}) \ge 0$ if and only if $\pi_{2,t} \le \pi_{1,t}$.

Since $u^*_{t+1}(\cdot, \ldots, \cdot)$ is componentwise convex by Proposition 1, it is also continuous. Therefore, $g(\pi_{2,t}; \pi_{1,t}, \vec{\mu}_{12,t+1})$ is continuous in $\pi_{2,t}$. We identify some characteristics of function $g(\pi_{2,t}; \pi_{1,t}, \vec{\mu}_{12,t+1})$ by Lemma 2, 3 and 4 in the following sections, which are useful in the proof of the claim that $g(\pi_{2,t}; \pi_{1,t}, \vec{\mu}_{12,t+1}) \ge 0$ if and only if $\pi_{2,t} \le \pi_{1,t}$.

LEMMA 2. *If $u^*_{t+1}(\cdot, \ldots, \cdot)$ is a componentwise convex function, then $g(\cdot; \pi_{1,t}, \vec{\mu}_{12,t+1})$ is a concave function for all $\pi_{1,t} \in [0, 1]$ and $\vec{\mu}_{12,t+1} \in \prod_{i \in \{3,4,\ldots,I\}} [0, 1]$.*

LEMMA 3. *Given any* $0 \le \pi_{1,t} \le 1$ *and* $\vec{\mu}_{12,t+1} \in \prod_{i\in\{3,4,\ldots,I\}}[0,1]$*, there are two possible cases for* $g(1;\pi_{1,t},\vec{\mu}_{12,t+1})$*:*

(i) When $\pi_{1,t}=1$*,* $g(1;\pi_{1,t},\vec{\mu}_{12,t+1})=0$

(ii) When $\pi_{1,t}<1$*,* $g(1;\pi_{1,t},\vec{\mu}_{12,t+1})<0$

LEMMA 4. *Given any* $0\le \pi_{1,t}\le 1$ *and* $\vec{\mu}_{12,t+1}\in\prod_{i\in\{3,4,\ldots,I\}}[0,1]$*,* $g(0;\pi_{1,t},\vec{\mu}_{12,t+1})\ge 0$*.*

Proof. Proofs for Lemma 2, Lemma 4 and Lemma 3 are omitted to save space. ■

B. Proofs of theoretical results in Section 5.2

Establish optimality equation

To facilitate the proof of Proposition 2 and Theorem 2, in the following section we establish a more detailed version of optimality equation. The result from this section only applies to the proof of Proposition 2 and Theorem 2. The detailed derivation is omitted. Under the optimal policy that schedules Patients $1,2,\ldots,C$ in Period t, equations (3a) and (3b) can be written as:

$$u_t^*(\vec{\pi}_{1,t}\ ,\ \ldots\ ,\ \vec{\pi}_{I,t})=\sum_{i=1}^{I}\phi(\vec{\pi}_{i,t}) + \beta\sum_{x_{i,t},i\in S}\left[\prod_{i\in S}(\vec{\pi}_{i,t})_{x_{i,t}}u_{t+1}^*\Big(\vec{\gamma}_{x_{1,t}}\ ,\ \ldots\ ,\ \vec{\gamma}_{x_{C,t}},z(\vec{\pi}_{C+1,t}),\ldots,z(\vec{\pi}_{I,t})\Big)\right] \tag{24}$$

$$u_t^*(\vec{\pi}_{1,T}\ ,\ \ldots\ ,\ \vec{\pi}_{I,T})=\sum_{i=1}^{I}\phi(\vec{\pi}_{i,T}) \tag{25}$$

where $\sum_{x_{i,t},i\in\{1,\ldots,C\}}$ counts for all the possible realizations $(x_{1,t},x_{2,t},\ldots,x_{C,t})$.

Proof for Proposition 2 To prove Proposition 2, without loss of generality, we prove that at the beginning of a given Period t, if $\vec{\pi}_{i,t}=\vec{\pi}'_{i,t}$ for all i, except for the information vector of Patient j, i.e., $\vec{\pi}_{j,t}\succ_{s.t.}\vec{\pi}'_{j,t}$, then $u_t^*(\vec{\pi}_{1,t},\ldots,\vec{\pi}_{j,t},\ldots,\vec{\pi}_{I,t})\le u_t^*(\vec{\pi}_{1,t},\ldots,\vec{\pi}'_{j,t},\ldots,\vec{\pi}_{I,t})$. We prove it using induction.

Similar to the proof for Proposition 1, we separate the analysis by whether Patient j with information $\vec{\pi}_{j,t}$ is scheduled or not in Period t. The detailed proof is omitted here to save space. ■

Proof for Theorem 2

Similar to the proof of Theorem 1, given any I patients with information vector $\vec{\pi}_{1,t},\vec{\pi}_{2,t},\ldots,\vec{\pi}_{I,t}$ at the beginning of a given Period $t\in\{1,2,\ldots,T-1\}$, recall that set S includes the indices of all patients that are scheduled in Period t and $\overline{S}$ includes the indices of all patients who are not scheduled in this period. The theorem claims that $\forall j\in S$ and $\forall j'\in\overline{S}$, it must be that $\vec{\pi}_{j,t}\succ_{s.t.}\vec{\pi}_{j',t}$.

We prove the theorem using contradiction. If the theorem is not true, then there exits an optimal policy under which Patient $j\in S$ and Patient $j'\in\overline{S}$, but $\vec{\pi}_{j,t}\prec_{s.t.}\vec{\pi}_{j',t}$ and $\vec{\pi}_{j,t}\ne\vec{\pi}_{j',t}$. Suppose without loss of generality, that under the optimal policy Ω, $S=\{1,3,\ldots,C+1\}$ and $\overline{S}=\{2,C+2,\ldots,I\}$. Moreover, we assume $j=1$ and $j'=2$, i.e., $\vec{\pi}_{1,t}\prec_{s.t.}\vec{\pi}_{2,t}$ and $\vec{\pi}_{1,t}\ne\vec{\pi}_{2,t}$. We show that policy Ω cannot be optimal, by constructing policy Ω' that schedules Patient 2 instead of Patient 1 in Period t but keep everything else the same as policy Ω, and showing that following policy Ω' yields more benefit than Ω.

Consider random vector $\vec{\boldsymbol{\mu}}_{12,t+1}$ which consists of $(I-2)$ random vectors and has realization $\vec{\mu}_{12,t+1}$ where

$$\vec{\mu}_{12,t+1}=\Big(\vec{\gamma}_{x_{3,t}}\ldots,\vec{\gamma}_{x_{C+1,t}},z(\vec{\pi}_{C+2,t})\ ,\ \ldots\ ,\ z(\vec{\pi}_{I,t})\Big) \tag{26}$$

where $\vec{\mu}_{12,t+1}$ represents the information vector for all patients except for Patient 1 and Patient 2, and its realization (26) occurs with probability $\prod_{i\in S\setminus\{1,2\}}(\vec{\pi}_i)_{x_{i,t}}$.

Using random vector $\vec{\boldsymbol{\mu}}_{12,t+1}$, the value function when scheduling Patient 1 instead of Patient 2 can be rewritten as:

$$u_t(\vec{\pi}_{1,t},\dots,\vec{\pi}_{I,t};1\in S,2\in\overline{S})=\sum_{i=1}^{I}\phi(\vec{\pi}_{i,t})+\beta\mathbb{E}_{\vec{\boldsymbol{\mu}}_{12,t+1}}\Big[\sum_{k=0}^{K}(\vec{\pi}_{1,t})_k u^*_{t+1}(\vec{\gamma}_k,z(\vec{\pi}_{2,t}),\vec{\boldsymbol{\mu}}_{12,t+1})\Big] \tag{27}$$

Similarly, the value function when scheduling Patient 2 instead of Patient 1 would be:

$$u_t(\vec{\pi}_{1,t},\dots,\vec{\pi}_{I,t};2\in S,1\in\overline{S})=\sum_{i=1}^{I}\phi(\vec{\pi}_{i,t})+\beta\mathbb{E}_{\vec{\boldsymbol{\mu}}_{12,t+1}}\Big[\sum_{k=0}^{K}(\vec{\pi}_{2,t})_k u^*_{t+1}(z(\vec{\pi}_{1,t}),\vec{\gamma}_k,\vec{\boldsymbol{\mu}}_{12,t+1})\Big] \tag{28}$$

For perfect treatment, equation (27) and (28) can be reduced to the following ones because the terms within summation do not depend on k, and $u^*_{t+1}(\cdot,\dots,\cdot)$ is symmetric in its variates.

$$u_t(\vec{\pi}_{1,t},\dots,\vec{\pi}_{K,t};1\in S,2\in\overline{S})=\sum_{i=1}^{I}\phi(\vec{\pi}_{i,t})+\beta\mathbb{E}_{\vec{\boldsymbol{\mu}}_{\mathbf{12},\mathbf{t+1}}}\Big[u^*_{t+1}(\vec{\gamma}_0,z(\vec{\pi}_{2,t}),\vec{\boldsymbol{\mu}}_{12,t+1})\Big] \tag{29}$$

$$u_t(\vec{\pi}_{1,t},\dots,\vec{\pi}_{K,t};2\in S,1\in\overline{S})=\sum_{i=1}^{I}\phi(\vec{\pi}_{i,t})+\beta\mathbb{E}_{\vec{\boldsymbol{\mu}}_{\mathbf{12},\mathbf{t+1}}}\Big[u^*_{t+1}(\vec{\gamma}_0,z(\vec{\pi}_{1,t}),\vec{\boldsymbol{\mu}}_{12,t+1})\Big] \tag{30}$$

Since $\vec{\pi}_{1,t}\prec_{s.t.}\vec{\pi}_{2,t}$ and $z(\cdot)$ preserves stochastic ordering, $z(\vec{\pi}_{1,t})\prec_{s.t.}z(\vec{\pi}_{2,t})$. From Proposition 2, $u^*_{t+1}(\cdot,\dots,\cdot)$ is componentwise decreasing. Therefore for any given realization $\vec{\mu}_{12,t+1}$,

$$u^*_{t+1}(\vec{\gamma}_0,z(\vec{\pi}_{2,t}),\vec{\mu}_{12,t+1})\le u^*_{t+1}(\vec{\gamma}_0,z(\vec{\pi}_{1,t}),\vec{\mu}_{12,t+1})$$

Therefore from equations (29) and (30)

$$u_t(\vec{\pi}_{1,t},\dots,\vec{\pi}_{K,t};1\in S,2\in\overline{S})\le u_t(\vec{\pi}_{1,t},\dots,\vec{\pi}_{K,t};2\in S,1\in\overline{S})$$

which implies that policy Ω that does not schedule Patient 2 cannot be optimal. This completes the proof of Theorem 2. ∎

C. Proof for theoretical results in Section 6

Proof for Theorem 3

We prove the theorem by contradiction. If the theorem is not true, then there exists a myopic policy Π under which Patient j is scheduled and Patient j' is not scheduled, with $\vec{\pi}_{j,t}\prec_{s.t.}\vec{\pi}_{j',t}$, assuming that $\vec{\pi}_{j,t}\neq\vec{\pi}_{j',t}$. We show that this policy cannot be optimal for the maximization problem in (7), by demonstrating that policy Π' which schedules Patient j' instead of Patient j yields a higher reward in the next period than policy Π; hence Π cannot be a myopic policy.

The objective function in (7) can be written as the sum of three QALY values in Period t: (i) for patients other than j and j', (ii) for Patient j and (iii) for Patient j'.

$$\begin{aligned}&\mathbb{E}\Big[\sum_{i=1}^{I}\phi(\vec{\boldsymbol{\pi}}_{i,t+1})|\vec{\pi}_{1,t},\vec{\pi}_{2,t},\dots,\vec{\pi}_{I,t}\Big]\\&=\mathbb{E}\Big[\sum_{i=1,i\neq j,j'}^{I}\phi(\vec{\boldsymbol{\pi}}_{i,t+1})|\vec{\pi}_{1,t},\vec{\pi}_{2,t},\dots,\vec{\pi}_{I,t}\Big]+\mathbb{E}\Big[\phi(\vec{\boldsymbol{\pi}}_{j,t+1})|\vec{\pi}_{j,t}\Big]+\mathbb{E}\Big[\phi(\vec{\boldsymbol{\pi}}_{j',t+1})|\vec{\pi}_{j',t}\Big]\end{aligned} \tag{31}$$

Note that with perfect treatment, when Patient j is scheduled, $\mathbb{E}\Big[\phi(\vec{\boldsymbol{\pi}}_{j,t+1})|\vec{\pi}_{j,t}\Big]$ is independent of $\vec{\pi}_{j,t}$ since $\vec{\pi}_{j,t}\mathcal{Q} = \vec{e}_0$. Thus, equation (31) differs between policies Π and Π' only in the expected reward from the unscheduled patient (j or j'). Policy Π which does not schedule Patient j' outperforms policy Π' which does not schedule Patient j if and only if the following inequality holds: $\phi(\vec{\pi}_{j',t}\mathcal{P}) \geq \phi(\vec{\pi}_{j,t}\mathcal{P})$ By assumption $\vec{\pi}_{j,t} \prec_{s.t.} \vec{\pi}_{j',t}$ and $\mathcal{P}$ preserves stochastic ordering, then $\vec{\pi}_{j,t}\mathcal{P} \prec_{s.t.} \vec{\pi}_{j',t}\mathcal{P}$. Since $\phi(\vec{\pi}) = \sum_{k=0}^{K} b_k(\vec{\pi})_k$ and $\{b_k\}_{k=0}^{K}$ a decreasing sequence; therefore, by Shaked and Shanthikumar (2007), given that $\vec{\pi}_{j,t} \neq \vec{\pi}_{j',t}$, it follows that $\phi(\vec{\pi}_{j',t}\mathcal{P}) < \phi(\vec{\pi}_{j,t}\mathcal{P})$ which is contradictory to the previous inequality. This completes the proof of theorem 3. ∎

Proof for Theorem 4

We prove the theorem by contradiction. If the theorem is not true, then there exists a myopic policy Π under which Patient j is scheduled and Patient j' is not scheduled, with $\pi_{j,t} > \pi_{j',t}$. We show that this policy cannot be optimal for the maximization problem in (7), by demonstrating that policy Π' which schedules Patient j' instead of Patient j yields a higher reward in the next period than policy Π; hence Π cannot be a myopic policy.

We write the objective function in (7) as the sum of QALY values in Period t for (i) patients other than j and j', (ii) for Patient j and (iii) for Patient j'.

$$\begin{aligned} &\mathbb{E}\Big[\sum_{i=1}^{I}\phi(\boldsymbol{\pi}_{i,t+1})|\pi_{1,t},\pi_{2,t},\ldots,\pi_{I,t}\Big] \\ &= \mathbb{E}\Big[\sum_{i=1,i\neq j,j'}^{I}\phi(\boldsymbol{\pi}_{i,t+1})|\pi_{1,t},\pi_{2,t},\ldots,\pi_{I,t}\Big] + \mathbb{E}\Big[\phi(\boldsymbol{\pi}_{j,t+1})|\pi_{j,t}\Big] + \mathbb{E}\Big[\phi(\boldsymbol{\pi}_{j',t+1})|\pi_{j',t}\Big] \end{aligned} \tag{32}$$

Policy Π which schedules Patient j and not Patient j' outperforms policy Π' which schedules Patient j' and not Patient j if and only if $(1-\pi_{j,t})\phi(qp_0+(1-q)p_1)+(1-\pi_{j',t})\phi(p_1) \geq (1-\pi_{j',t})\phi(qp_0+(1-q)p_1)+(1-\pi_{j,t})\phi(p_1)$. Rearranging terms, the above inequality is equivalent to $\phi(p_1 + q(p_0 - p_1)) \leq \phi(p_1)$. Given that $p_0 > p_1$, $q > 0$ and the definition of ϕ (i.e., $\phi(\pi_{j,t}) = \pi_{j,t}$), the above inequality cannot hold: Policy Π cannot outperform policy Π'. This completes the proof for Theorem 4. ∎