MIT Open Access Articles

A network flow approach for tactical resource planning in outpatient clinics

Massachusetts Institute of Technology

A Network Flow Approach for Tactical Resource Planning in Outpatient Clinics

Thi Thu Ba **NGUYEN**[1], Appa Iyer **SIVAKUMAR**[1,2], Stephen C. **GRAVES**[1,3]

[1]Manufacturing Systems and Technology Program, Singapore-MIT Alliance, Nanyang Technological University

65 Nanyang Drive, Singapore 637460

E-mail:✉ ng0003ba@e.ntu.edu.sg; ngthuba@mit.edu

[2]School of Mechanical and Aerospace Engineering, Nanyang Technological University

65 Nanyang Drive, Singapore 637460

E-mail: MSIVA@ntu.edu.sg

[3]Department of Mechanical Engineering and Engineering Systems, Massachusetts Institute of Technology

77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA

E-mail: sgraves@MIT.EDU

A Network Flow Approach for Tactical Resource Planning in Outpatient Clinics

Abstract

This paper introduces a deterministic model to plan the physician requirements for outpatient clinics to achieve service targets for the appointment lead-times of patients. The Ministry of Health of Singapore has established targets for the median, 95th percentile, and 100th percentile of appointment lead-times for patients, since long appointment postponements are regarded as being unacceptable for health care services. The study is to match the capacity of the healthcare providers to the patient demand for a re-entry system, subject to restrictions on the appointment lead-times for patients. We propose a mixed-integer programming model for planning capacity with the minimization of the maximum required capacity as its objective. In the model we assume a finite planning horizon, deterministic arrivals, multiple types of patients, identical physicians, and dependent demand between types of patients. We solve this model with a Branch and Cut algorithm. We test the model with numerical experiments using real data from the chosen specialty at the outpatient clinic of the studied hospital. The results show the value of the proposed model via a systematic push-pull mechanism in scheduling patients' requests to minimize the objective. The clinic should use one of the appointment lead-time targets to determine the patients' appointment dates. Finally, from the sensitivity analyses we demonstrate that the objective is negatively correlated with first-visit patients' appointment lead-time targets, the discharge rates, and the re-visit patients' mean appointment lead-time; we find a positive correlation between the first-visit patients' mean appointment lead-time and the appointment lead-time targets.

Keywords   Appointment Lead-time, Capacity Planning, Mixed Integer Programming, Outpatient Clinics.

1. Introduction

The increase in the patients' appointment lead-time (Figure 1) at the outpatient clinics of hospitals has been a crucial concern of the Ministry of Health (MOH) in Singapore. Currently, the MOH is focused on the performance of healthcare services, particularly the performance of outpatient services, in order to improve the quality of health care service. The MOH has set targets for the median, 95th percentile, and 100th percentile of appointment lead-times as a guideline for hospitals in Singapore. This poses challenges to the hospitals to provide quick access or short appointment lead-times to healthcare services. The hospitals have to determine

2

how to continue to provide quality healthcare while reducing their appointment lead-times with their existing resources.
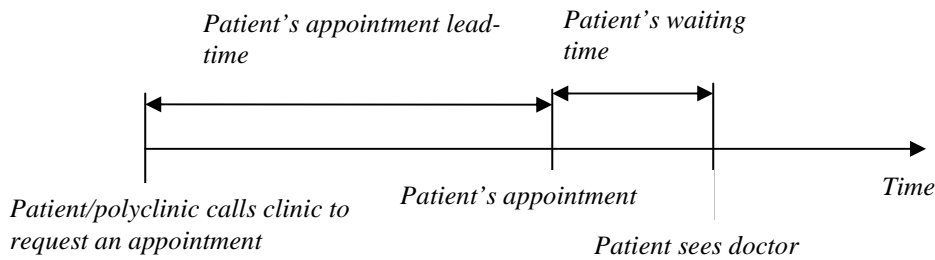

Fig. 1 *Appointment lead-time definition*

This research is based on the study of an outpatient clinic at Tan Tock Seng Hospital (TTSH) in Singapore. A new patient, termed as a first-visit (FV) patient, calls a contact center to book an appointment. Then, the patient has to wait until his/her appointment-date (a reminder of the appointment is sent, for reducing the rate of no-show [14]). After the first visit, a patient typically continues to return to the clinic at regular intervals as part of a treatment plan to assure continuity of care [24]. After the first visit, the patient then becomes a re-visit (RV) patient. There is a restricted range for the appointment lead-time for the next visit by a RV patient. In addition, the patient may need to return as a RV patient several times, depending on treatment protocol. This re-entry system (Figure 2) challenges the service providers in planning their capacity for future demand, because any new arrival (or any FV patient) leads to a number of additional RV arrivals.
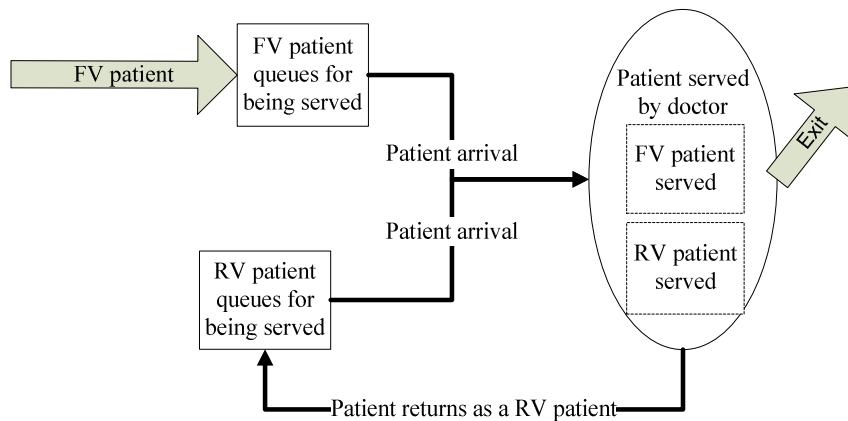
Fig. 2 *Patient flow through the re-entry system*

Currently the clinic plans capacity by allocating the available physician time into blocks dedicated to see FV patients and blocks dedicated to see RV patients. These allocations are based largely on the preferences of each physician. The appointment booking for FV patients is usually on a first-come-first-served basis, where the patient is offered the earliest block with available capacity. Appointments of RV patients are scheduled by the physicians and depend on

3

the prescribed return interval (e.g., come back in three to four weeks), where again the patient is offered the earliest available block in the given range. The current practice results in long appointment lead-times because of the mismatch between the volume and mix of patient demands and the capacity of the healthcare providers. Therefore, we expect there is opportunity for improvement from a more systematic way to plan the required capacities.

Furthermore, urgent requests are very rare at the studied clinics. For most of the urgent requests, the patients will be sent to an emergency department that operates separately from the studied clinics. Occasionally, a physician will agree to see an urgent request within the clinic, but this is quite rare. Hence, we can ignore the urgent cases when examining the appointment scheduling and capacity planning at the outpatient clinics.

This study develops a network flow model to plan future required capacity at the tactical level for the re-entry system. We propose a model to minimize the maximum level of physician-hours that are required to meet the MOH's service targets for median, 95th percentile, and 100 percentile of appointment lead-times for patients. We formulate the model in Section 3, which follows the literature review in Section 2. In Section 4, we report on numerical experiments on real case studies; we also present our analysis of the characteristics of the objective and the appointment lead-time targets. We then discuss our findings in Section 5, and we provide a conclusion in Section 6.

2. Literature Review

The appointment systems of outpatient clinics have been studied for many years. One of the earliest studies was in the early of 1950s by Bailey [1] who adopted a queuing perspective to investigate the relationship between patients' waiting times and physicians' idle times. The author analyzed the impact of patients arriving at a clinic (e.g. such as an appointment interval, a number of patients per session, and patients' punctuality) on patients' waiting times and physicians' idle times. Many subsequent studies have investigated how to improve service performance in outpatient appointment systems. The investigations mainly consider patient appointment scheduling and capacity design problems.

We comment here on three approaches to the appointment scheduling problem: queuing theory [1, 4, 10, 17, 19], simulation [6, 9, 13, 15, 23, 25], and heuristic search [5, 12, 18, 26]. Queuing theory has been employed extensively to model and analyze different health care systems (see Lakshmi and Sivakumar [16] for an overview). Applications of queuing theory in scheduling appointment systems for outpatient clinics aim to find the optimal system performance taking into consideration the tradeoff between patients' waiting time versus physicians' idle time as a measurement of service's performance. The studies investigate the impact of various factors such as no-show [10, 17], punctuality of doctors [17], punctuality of patients [10], length of appointment interval [1, 10], and the number of patients per session [1].

Bailey [1], Brahimi and Worthington [4], Hassin and Mendel [10], and Liu and Liu [17] consider a deterministic appointment interval [1, 4, 10, 17] while Pegden and Rosenshine [19] try to determine the optimal appointment interval. In addition, these papers make various assumptions on service time distribution: stochastic service systems are investigated with an exponential distribution [10, 19], a general distribution [4, 17], and a Pearson Type III curve [1].

Secondly, simulation has been used for the appointment scheduling problem to determine the bottle-neck station [6, 9, 25], and to schedule physicians to achieve a given service performance [13, 15, 23]. The first set of papers explored the effect of patient flow and that of the appointment schedule on the stations' performance in clinics. Harper and Gamlin [9], and Wijewickrama and Takakuwa [25] considered different priorities of appointment schedule amongst categories of patients, while Cote [6] gave the same priority to all patients. Cote [6] measured utilization, queue length, probability of full occupied stations, and patient flow time. Harper and Gamlin [9] considered the objectives of patient waiting duration from his/her appointment time until his/her first time served, the percentage of patients that wait more than 30 minutes, and the mean time that a patient spends in the clinic. Wijewickrama and Takakuwa [25] examined the tradeoff between the patient's waiting time and the doctor's idle time. The second set of papers focused on physician scheduling and appointment scheduling policies in order to minimize patients' waiting time and physicians' idle time. Rising et al. [23] provided a policy without differentiating between advanced booking and walk-in patients. Klassen and Rohleder [13] accounted for this, by prioritizing based on types of patients. An overbooking policy was proposed by LaGanga and Lawrence [15] to reduce the impact of no-shows

Finally, many studies propose heuristic searches to optimize the appointment scheduling problem. This approach has been conducted for various problem instances, e.g., the exponential service time distribution [12, 18, 26], general service time distribution [5], single no-show probability [12], different no-show probabilities [5, 18, 26], multi-modularity [12, 26] and unimodal objective [5, 18]. Kaandorp and Koole [12] derived a local search in which a local optimal schedule converges to the global optimal. Zeng et al. [26] proposed a local search to find a local optimal schedule for heterogeneous patients. Muthuraman and Lawley [18] provided proofs of necessary and sufficient conditions for their scheduling policy. Chakraborty et al. [5] developed a sequential scheduling policy which is able to provide an appointment to patients before call is ended. Zeng et al. [26], Muthuraman and Lawlay [18], and Chakraborty et al. [5] addressed the objective of profit maximization involving a trade-off between patient revenue, and cost associated with patients' waiting time and physicians' over-time. Kaandorp and Koole [12] considered an objective that is a function of a weighted mean of expected patients' waiting times, doctors' idle time, and doctors' overtime.

In contrast to the appointment scheduling problem, there are fewer studies that investigate how to plan the capacity for an appointment system. There are two versions of the

problem. The resource allocation version optimizes the usage of given and known resources. The re-design capacity version aims to determine the minimal amount of resources required.

Resource allocation optimization was investigated by Qu et al. [20, 22]. In both articles, the authors provided analytic approaches to determine the optimal percentage of advanced (or pre-booked) access versus open access appointments. The studies aim to match the provider's capacity with patient demand, assuming that a long wait causes a high no-show rate. The authors started with single period [20], and later expanded the study to two time periods [22].

The re-design capacity problem was studied in the 1980s by Ittig [11]. The study determined the capacity level to maximize profit. The profit is a function of patient revenue, additional capacity, and patient cost. Capacity level continued to be investigated by Bowers et al. [3], and Qu and Shi [21]. These articles derived functions to determine required capacity. Bowers et al. [3] provided a simple function to identify the number of clinics required in each specialty from information of new attendances, follow-up rates, no-show rates, and utilization of a clinic. Green and Savin [8] developed a recursion model that could help determine capacity requirement or patient panel size for a same-day appointment system. Qu and Shi [21] proposed a function of demands and no-show rates of two different types of patients (e.g., fixed appointments and open appointments) to determine the expected number of patients to be scheduled, which helps to evaluate the required capacity. According to our best knowledge, there is only one paper which considers appointment lead-time [7]. The study mentions the 95th percentile of appointment lead-time of new patients as a performance target and uses an M/D/1 queuing model. However, the authors did not provide a systematic way to plan the capacity for the given target.

In general, there does not seem to be any literature that examines how to meet multiple service targets, such as for the median, 95th percentile, and 100th percentile of appointment lead-times. Almost all the studies optimize the performance for a given level of capacity, and allow for requests of patients for treatment to be rejected. The MOH of Singapore regards the rejection of the requests for treatment as being unacceptable for healthcare services. Although Bowers et al. [3], Ittig [11], and Qu and Shi [21] designed capacity levels which assumed that all requests of patients are accepted, patient appointment lead-times were not addressed. All of the papers ignore the re-entry characteristic in their studies except Bowers et al. [3]. The authors developed a model including RV patients' demand, but assume that patients return only once into the system. This assumption underestimates the required capacity for re-entry systems in which any new arrival may lead to multiple returns to the system; this is the case for the clinic we study. Thus, our work focuses on designing capacity to match the demand for the re-entry systems while considering appointment lead-times targets. Our performance measure is the maximum required capacity.

We propose a network flow model for planning capacity of the re-entry outpatient appointment system at the tactical level. This model allows for restrictions on the length of

patients' appointment lead-times. We provide a detailed description of the model in the next section.

3. Model Development

We introduce a network flow model (Figure 3) with deterministic arrivals of FV patients, arrivals of RV patients that depend on the FV arrivals, and a finite horizon for the re-entry system. One actual patient can create more than one patient visit. Therefore, within the model we will specify the "patient" as being defined as a "patient visit". The following model minimizes the maximum required capacity (physician-hours) when all requests of new appointments must be accepted.

The required capacity must satisfy the achievement of FV appointment lead-time targets for median, p^{th} percentile, and 100th percentile, as well as restrictions on the RV appointment lead-time range and on the RV's mean appointment lead-time. We let u, v, and w correspondingly be a median, p^{th} percentile, and 100th percentile of FV appointment lead-time targets; for instance, we might have targets of two weeks for the median, six weeks for the 95th percentile and ten weeks for the 100th percentile. Furthermore, we let [a, b] and $\bar{a}$ denote the restrictions on the RV appointment lead-time range and on the RV's mean appointment lead-time, respectively; for instance, we might require that a RV patient needs to be scheduled between two and ten weeks, with a mean of seven weeks.
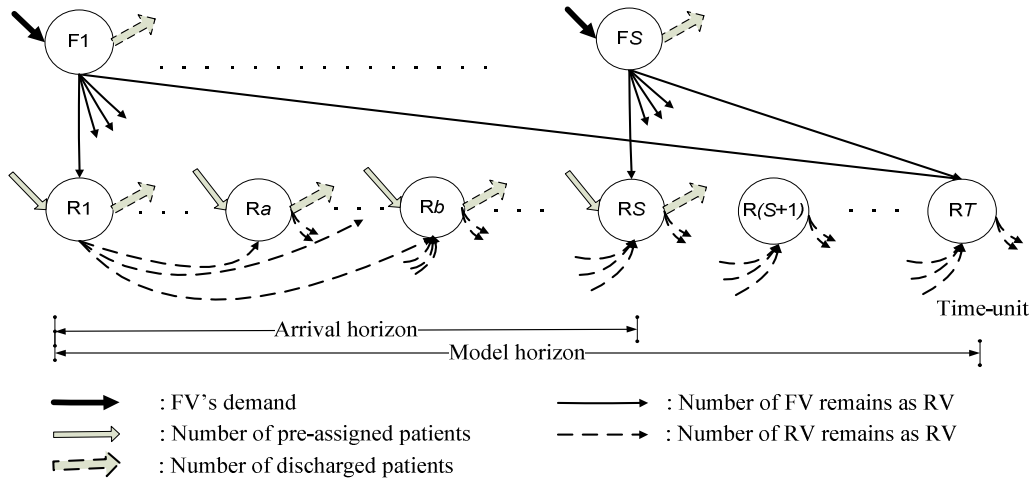
Fig. 3 *A network flow model for planning capacity of the re-entry appointment system*

We term the first S time-units to be the arrival horizon. We assume that FV demands are known for S time-units into the future. However, due to the patient dynamics and the inherent appointment lead-times, we extend the model horizon to T time-units; we need to set T to be significantly long to cover the maximum appointment lead-time allowed for the last arriving FV

7

patient. Moreover, the objective of the problem is the minimization of the maximum required capacity. Therefore, we need to also consider the maximum appointment lead-time for RV patients when setting T. In other words, $T = Max[S + w; S + b]$ so as to cover the maximum possible appointment lead times for both FV and RV patients that arrive within the arrival horizon. We list assumptions and notation in Tables 1 and 2, respectively. The above description leads to design of S and T nodes that represent FV and RV patient nodes (or FV-nodes and RV-nodes), respectively. FV- nodes are numbered F1, F2, up to FS, and RV-nodes R1, R2, up to RT.

The arrival of the FV patient, f_i, at time-unit i^{th} is signified in Figure 3 by the black thick solid arrows that go into the FV-nodes. The grey thick solid arrows that go into the RV-nodes represent the total number of pre-assigned FV patients, r_j^f, and RV patients, r_j^r. The FV pre-assigned patients are patients who made their requests before the start of the arrival horizon and their appointment dates are within the arrival horizon. The RV pre-assigned patients are patients whose last appointment was before the start of the arrival horizon but their following appointment is within the arrival horizon.

Table 1 *List of assumptions*

No.	Assumptions
1	Patient requests cannot be rejected.
2	The designed capacity needs to be sufficient for both FV and RV patients, who are already in the system at the start.
3	The model considers a single type of FV patients and a single type of RV patients in terms of consultation times, appointment lead-times, and discharge rates.
4	Physicians are identical in terms of consultation times
5	A patient's appointment can be scheduled in the same time-unit as the patient's request.
6	There is no uncertainty in the rate at which patients arrive for their appointment requests.
7	Patients continue to be RV patients until they are discharged from the system.
8	An FV patient is a patient who calls to request an initial appointment. After the first appointment the FV patient might be discharged; if not, the FV patient becomes a RV patient and requests a re-visit appointment.
9	The model does not consider discharges after the arrival horizon.
10	There are no new requests after the arrival horizon.
11	There are constant discharge rates α, β for FV and RV patients respectively.

We consider the impact of discharge rates on designing the clinic capacity. The discharge rates may be different following each visit, but for the sake of model simplification, we assume constant discharge rates α, β for FV and RV patients respectively. Since the discharge rates are different, it is necessary for the model to track whether a patient visit is a first visit or a return visit. The thick solid arrows that go out from FV-nodes represent the total number of FV discharges, d_i^f, namely the number of FV patients that are discharged after their first

appointment. For example, the thick solid arrow that goes out from node F1 represents the total number of discharges for FV patients, who make requests for appointments in the first time-unit. However, the schedule for their appointments may be at another time. Some patients' appointment dates could be on the first time-unit, or others' appointment dates could be on the second or up to the Tth time-unit. The thick solid arrows that go out from RV-nodes represent total number of RV discharges, d_j^r, for patients that were discharged as RV patients. For example, the thick solid arrow that goes out from node R1 represents the total number of patients that will be discharged after their next appointment. The time until the next appointment will be between a to b time-units from the 1st time-unit.

Table 2 *List of notation*

	Notation	Descriptions
Definition of sets	N	Index set for the arrival horizon.
	Z	Set of all $z_{ij}: j - i \geq 0$.
	L^m, L^p	Set of all $z_{ij} \in Z$ that have $j - i \leq u$ and set of all $z_{ij} \in Z$ that have $j - i \leq v$, respectively.
Decision variables	z_{ij}	Number of FV patients who make a request in the ith time-unit and have their appointment scheduled in the jth time-unit.
	x_{ij}	The number of FV patients who request in the ith time-unit and have appointment in the jth time-unit, and still remain in the system as RV patients after their appointment in the jth time-unit.
	y_{ij}	The number of patients who have an appointment in the ith time-unit and have their next appointment in the jth time-unit, and still remain as RV patients after the jth time-unit.
	d_i^f	The number of FV patients who make a request in the ith time-unit, and are discharged after their first appointment.
	d_j^r	The number of RV patients who are discharged after their appointment in time-unit j.
	c_i^f	Capacity for FV patients in the ith time-unit, minutes.
	c_j^r	Capacity for RV patients in the jth time-unit, minutes.
	c_i	Total capacity in the ith time-unit, minutes.
	q	The maximum required capacity per time-unit, minutes.
Inputs	u, v, w	Appointment lead-time targets for median, p^{th} percentile, and 100th percentile respectively, time-units.
	p	A pth percentile of a given appointment lead-time target v ($0 < p < 1$)
	τ^f, τ^r	Consultation times for FV and RV patients, in minutes. These times are specified by the doctors.
	α, β	Discharge rates for FV and RV patients, respectively ($0 < \alpha, \beta < 1$)
	f_i	Number of FV patients' requests that arrive in the ith time-unit.
	r_j^f, r_j^r	The number of pre-scheduled FV and RV patients with appointments in time-unit j, who still remain as RV patients after their appointments.

	[a, b]	Restricted appointment lead-time of RV patients or range of RV appointment lead-time, time-units.
	$\bar{a}$	RV's mean appointment lead-time, time-units.
Other notation	S	The number of time units in the arrival horizon
	T	The number of time-units in the planning horizon, $T = Max(S + b, \quad S + w)$.
	i, j	Indices for time-units.

The thin solid lines that connect FV-nodes and RV-nodes show the number of FV patients who make their request in time-unit i and are assigned appointments in time-unit j, and that will remain in the system after their consultation in time-unit j. We denote this by x_{ij}. We allow the appointment date to be the same as the request date. As mentioned above, the constant α is assumed. Therefore, the thin solid arrows that connect FV-nodes and RV-nodes can also be used to relate to the assignments of appointment requests for FV patients.

The dashed arrows that connect two RV-nodes signify the number of RV patients with appointments in time-unit i that have their next appointment in time-unit j, and that will not be discharged at time-unit j. We denote this by y_{ij}. We do not model randomness in the no-show rate. Effectively, we assume that all patients will arrive; alternatively we could assume that there is a fixed fraction that does not arrive. We also assume that physicians are identical, both in the eyes of the patients and in their consultation times. Hence, patients exercise no choice as to which physician to see, or the timing of their appointment. These assumptions simplify the model, and we believe that it is acceptable given the intent to use the model for tactical planning. Finally, we assume that a patient remains as a RV patient until he/she is discharged.

Min q $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (1)

Subject to:

$$q \geq c_j, \forall j \in N, \qquad\qquad\qquad (2)$$

$$\sum_{j=i}^{T} z_{ij} = f_i, \forall i \in N, \qquad\qquad\qquad (3)$$

$$x_{ij} - (1 - \alpha)z_{ij} = 0, \ \forall i, j \in N, \qquad\qquad\qquad (4)$$

$$d_i^f = \alpha f_i, \ \forall i \in N, \qquad\qquad\qquad (5)$$

$$\left(r_j^f + r_j^r + \sum_{i=1}^{j} x_{ij} + \sum_{i=1}^{j} y_{ij}\right) - \left(d_j^r + \sum_{i=j}^{T+1} y_{ji}\right) = 0, \ \forall j \in N, \qquad (6)$$

$$d_j^r - \beta\left(r_j^f + r_j^r + \sum_{i=1}^{j} y_{ij} + \sum_{i=1}^{j} x_{ij}\right) = 0, \forall j \in N, \qquad (7)$$

$$y_{ij} = 0, \forall j - i < a, \forall i, j \in N, \qquad\qquad\qquad (8)$$

$$y_{ij} = 0, \forall j - i > b, \forall i, j \in N, \qquad\qquad\qquad (9)$$

$$y_{ij} = 0, \quad \forall j \geq T + 1, \forall i \leq S, \qquad\qquad\qquad (10)$$

$$\sum_{i=S+1}^{T} d_i^r = 0 , \tag{11}$$

$$\sum_{z_{ij} \in L^m} z_{ij} \geq \left(\tfrac{1}{2}\sum_{i \in N} f_i\right) + 1, \tag{12}$$

$$\sum_{z_{ij} \in L^p} z_{ij} \geq p \sum_{i \in N} f_i, \tag{13}$$

$$z_{ij} = 0, \forall j - i > w, \forall j - i < 0, \tag{14}$$

$$\sum_{j=1}^{T}\sum_{i=1}^{j}(j-i)\, y_{ij} - \bar{a} \sum_{j=1}^{T}\sum_{i=1}^{j} y_{ij} \leq 0, \tag{15}$$

$$c_j^f - \left(\tfrac{1}{1-\alpha}\tau^f\, r_j^f + \tau^f \sum_{i=1}^{j} z_{ij}\right) = 0, \ \forall j \in N, \tag{16}$$

$$c_j^r - \left(\tfrac{1}{1-\beta}\tau^r\, r_j^r + \tfrac{1}{1-\beta}\tau^r \sum_{i=1}^{j} y_{ij}\right) = 0, \ \forall j \in N, \tag{17}$$

$$c_j - \left(c_j^f + c_j^r\right) = 0 , \ \forall j \in N, \tag{18}$$

$$x_{ij}, y_{ij}, z_{ij} , c_i^f, \ c_j^r, d_i^f, d_j^r \geq 0, \ \forall i,j \in N, \tag{19}$$

$$z_{ij} \ \text{integer}, \ \forall i,j \in N. \tag{20}$$

The objective (1) minimizes the maximum required capacity, per time-unit, to achieve the appointment lead-time targets. The maximum required capacity for the clinic is shown in constraint (2). The above objective leads to the optimal allocated capacity between FV and RV patients as well, in each time-unit. The conservation flow at FV-nodes is represented in constraints (3), (4), and (5). Constraint (3) schedules an appointment for each request. Constraint (4) determines the RV patients who remain after the first visit. The number of FV patients who are discharged after their first visit is specified in constraint (5). The conservation of flow at RV-nodes is modeled in constraints (6) and (7). The first bracket of constraint (6) is the total number of pre-assigned FV and RV patients plus the total number of FV patients, plus the total number of RV patients, all with appointments in time-unit j. The second bracket of constraint (6) is the number of RV patients discharged after their visit in time-unit j plus the number of RV patients remaining with a subsequent appointment (the flows that go out from the same time-unit j). Constraint (7) specifies the number of patients discharged after their visit in time-unit j. Constraints (8) and (9) assure that no RV patients will have appointment lead-times falling outside the restricted range [a, b]. We accomplish this by defining y_{ij} to be zero if the time interval between i and j falls outside the restricted range. Constraint (10) prohibits RV patients' appointments to be made after last date of the model horizon T. Constraint (11) specifies the number of RV patients discharged after their visit in time-unit i. Constraints (12) and (13) assure that the targets of the appointment lead-time for the median and pth percentile are satisfied, respectively. Constraint (14) assures a realistic appointment date and does not allow appointments that go beyond the 100th percentile of limit. The restriction on the RV mean lead-time is represented in constraint (15). The required capacities of FV and RV patients in time-unit j are determined in constraints (16) and (17), respectively. The total required capacity in time-unit j is presented in constraint (18). The required capacity for either FV or RV patients

includes the required capacity of pre-assigned patients and that of the new arrivals. Constraint (19) requires non-negative variables. Finally, constraint (20) requires the integer assignment of FV patients.

4. Numerical experiments

In this section, we report on numerical experiments to investigate how the proposed mathematical model sets the appointments for patients. We report on four performance metrics: maximum required capacity, mean required capacity, FV's mean appointment lead-time, and RV's mean appointment lead-time.

4.1. Experimental design

Data from years 2008, 2009, and 2010 were obtained from the Urology specialty. We use the data from year 2008 to determine the number of pre-assigned FV and RV patients in the model. We use the data from years 2009 and 2010 to determine the inputs on the arrival of FV patients, on the discharge rates of FV and RV, and on the restricted appointment lead-times of RV patients. The design of data is due to the fact that the FV and RV pre-assigned patients of year 2009 came from requests in 2008, and the FV demands of year 2009 were scheduled with appointments in years 2009 and 2010. The length of the RV mean appointment lead-time is currently not restricted. Hence, the optimization model, as formulated, would tend to set the appointment lead-time for RV patients to be unrealistically long. To correct for this, we set the range for the RV appointment lead-times based on expert judgment; we also imposed a constraint on the mean RV appointment lead-time. The detailed inputs are listed in Table 3.

Table 3 The inputs included in the model for planning capacity of re-entry appointment system

No.	Descriptions		Urology
1		Median (u)	2 weeks
2	Appointment lead-time target	95th percentile(v)	6 weeks
3		100 percentile(w)	9 weeks
4		FV (τ^f)	15 minutes
5	Consultation time slots	RV (τ^r)	10 minutes
6		FV (α)	38 (%)
7	Discharge rates	RV (β)	32 (%)
8		Min (a)	2 weeks
9	Appointment lead-time of RV	Max (b)	30 weeks
10		Mean ($\bar{a}$)	30 weeks

The numerical experiments were implemented with the year 2009 as the arrival horizon. The planning horizon extends into 2010. The total number of pre-assigned FV and RV patients are 1,277 and 10,970 respectively. As the arrival horizon is 2009, pre-assigned patients are any patient who made an appointment request before Jan 1st 2009 and had their appointment scheduled on or after Jan 1st 2009. In addition, the number of total new FV requests in 2009 is 9,760 patients. We set up the model to plan the required capacity on a weekly basis. The Branch and Cut algorithm [2], a variant of the Branch and Bound algorithm, is chosen to solve the problem.

The computation time to find optimal solutions to all numerical examples is less than four minutes, implemented in IBM ILOG CPLEX Optimization Studio V12.4 on HP Pavilion G6 Notebook 1.9GHz PC with the Windows 7 operating system.

4.2. Experimental results

Given the inputs described in section 4.1, we report the results in Figures 4, 5 and 6. These figures illustrate the comparison between historical and model's performances in terms of the cumulative required capacity, the maximum required capacity per week, and the mean required capacity per week, respectively. The actual cumulative capacity (Figure 4) for FV patients was always less than what should have been provided according to the model. However, in terms of RV patients, the actual cumulative capacity was above what is obtained from the model. The implication of the result is that the hospital could achieve the lead-time targets by providing more capacity to FV patients and less capacity to RV patients, while maintaining the same total capacity. In other words, the proposed model provides the most efficient and effective planned capacity by systematically pushing and pulling the RV and FV patients' appointments so that the clinic can achieve the FV's appointment lead-time targets with the least maximum required capacity.
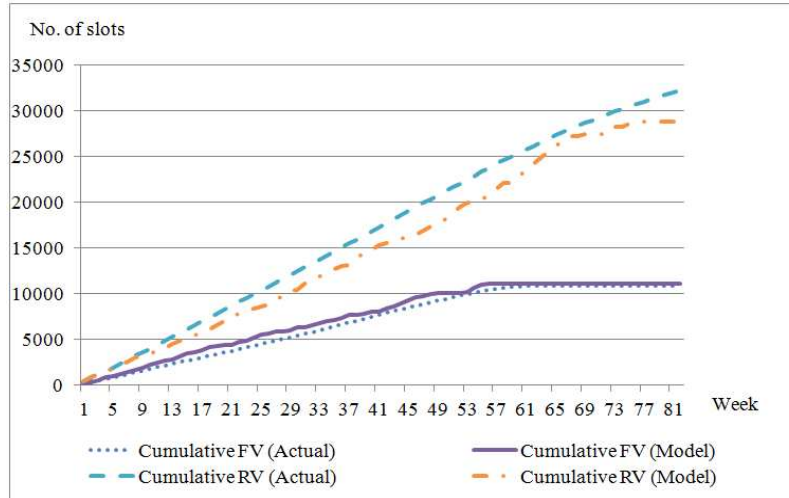
Fig. 4 The cumulative capacity of actual vs. model

As further look at the results we compare the mean required capacity, the maximum required capacity, and mean appointment lead-time of the proposed model and the actual plan. Figure 5 shows that for the FV patients the mean number of the required slots (135 slots per week) from the model is slightly more than that of the actual provided slots (132 slots per week). However, for the RV patients the mean number of the required slots (352 slots per week) from the model is much less than that provided (394 slots per week) from the actual plan. The mean total number of the required slots from the model (487 slots per week) is less than that of the provided slots from the actual plan (525 slots per week) for both FV and RV patients. In general, we can achieve the appointment lead-time targets without requiring the additional capacity.
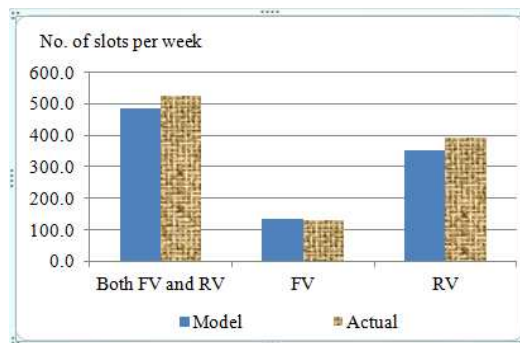

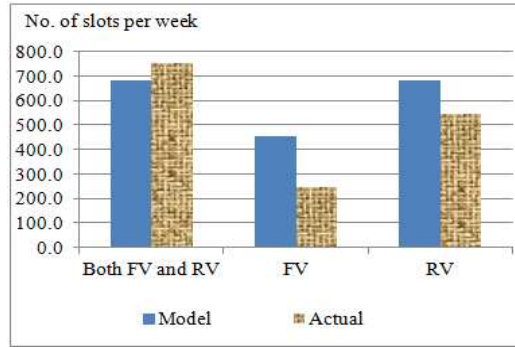**Fig. 5** The mean required capacity of actual vs. model

Fig. 6 The maximum required capacity of actual vs. model

Furthermore, Figure 6 shows that the maximum required capacity obtained from the model is only 680 slots per week (14.5 full-time equivalent doctors) for both FV and RV patients. This is less than values extracted from the actual plan, 751 slots per week (17.8 full-time equivalent doctors). This is evidence of the potential efficacy of the model for capacity planning. Since physicians are a scarce resource, the reduction in required resources is significant.

We investigate the drivers for this improvement with a detailed study on the maximum required capacity for each FV or RV patient type. The results (Figure 6) illustrate that the maximum required capacities for each FV and RV patient type from the proposed model are actually 2 and 1.3 times, respectively, as much as those from the actual plan. This suggests that reasonable assignments or allocations of these resources might vary from week to week to accommodate the appointment schedules and to achieve the appointment targets. The suggested model can help to do this, by proactively adjusting the appointments allocations for RV and FV patients to minimize the maximum required number of physicians.

Finally, we observe that the model produces shorter appointment lead-times for the newly arriving patients while still being able to maintain the continuity of treatments for RV patients. The FV's mean appointment lead-time from the model is 4.2 weeks. This is 35% lower than the FV's mean appointment lead-time of the actual plan (6.4 weeks). A consequence is that the appointment lead-times for the RV's patients are increased. The mean lead-times of RV patients from the model and the actual plan are 23 and 16 weeks respectively. However, this increase in lead-time does not violate the time constraints between treatments, and the continuity of care is maintained.

From the above results, we can see that the proposed model provides a systematic view to optimize the required capacity via the push-pull mechanism, and can produce better capacity plans than the actual practice.

15

4.3. Sensitivity analysis

We now examine the sensitivity of the model's performance measures to changes to the appointment lead-time targets, to the restriction on RV's mean appointment lead-time, and to the discharge rates. We define in Table 4 our terminology conventions for displaying and discussing the findings.

Table 4 *Definition of signs*

No.	Sign	Descriptions
1	-	A negative correlation between the lead-time target, under a specific lead-time target category, and the measurement.
2	+	A positive correlation between the lead-time target, under a specific lead-time target category, and the measurement.
3	0	The measurement's value remains the same while changing the lead-time target that belongs to the same category.
4	1	An indeterminate correlation.

4.3.1 Sensitivity to changes of the appointment lead-times

We investigate the sensitivity of the maximum required capacity, the mean required capacity, FV's mean appointment lead-time, and RV's mean appointment lead-time when changing the median, or the 95th percentile, or the 100th percentile of appointment lead-time target as well as changing the restricted length of RV's mean appointment lead-time. We performed numerical experiments for 5 different sets of discharge rates, for 6 levels of the RV's mean lead-time, and for 15 different sets of appointment lead-time targets (refer to Appendix A1 for the combinations that were considered). Thus, we set up and solved $5 \times 6 \times 15 = 450$ test cases. The details of those inputs and results are presented in Table 5.

We find that the maximum required capacity decreases as we increase an appointment lead-time target for FV patients; this is not surprising, but as expected. This finding is consistent for all considered pairs of discharge rates for the FV and RV patients. Similarly we find that the maximum required capacity decreases as we increase the RV's mean appointment lead-time. However, the maximum required capacity remains unchanged with the change of the appointment lead-time targets when the RV's mean appointment lead-time is large enough.

The mean required capacity is generally negatively correlated or indeterminate with the appointment lead-time targets, but not for all cases considered. When the RV's mean appointment lead-time is large enough, there is not a definitive correlation pattern.

The mean appointment lead-time of the FV patients is positively correlated to median, 95th percentile, and 100th percentile of appointment lead-time targets in most of the cases. This

trend may not exist when the restriction of mean lead-time for RV patients is large, as then the model can have alternative optimal solutions, which have different mean lead-times for FV patients. Besides, increasing the RV's mean appointment lead-time decreases the FV's mean appointment lead-time; this trend is consistent for all settings of the discharge rates.

We fit a regression model to these numerical results to estimate the relationship between the FV's mean appointment lead-time and the appointment lead-time targets, u, v and w. In equation (21) we provide the model with $\bar{l}$ being the estimate of the FV's mean appointment lead-time.

$$\bar{l} = 0.5u + 0.45v + 0.05w \qquad\qquad (21)$$

The mean appointment lead-time of the RV patient tends to be positively correlated with the change of the restricted length for RV's appointment lead-time. However, it remains unchanged when the limit on the RV's mean appointment lead-time is small. This reinforces the push- pull mechanism in assigning the FV and RV appointments.

Table 5 Summarization of sensitivity analysis to the changes of appointment lead-times

Restriction of RV's mean appointment lead-time (weeks)		Maximum required capacity					Mean required capacity					FV's mean appointment lead-time					RV's mean appointment lead-time				
	FV's discharge rate	0.38		0.38*	0.5	0.7	0.38		0.38*	0.5	0.7	0.38		0.38*	0.5	0.7	0.38		0.38*	0.5	0.7
	RV's discharge rate	0.1	0.2	0.32*	0.32		0.1	0.2	0.32*	0.32		0.1	0.2	0.32*	0.32		0.1	0.2	0.32*	0.32	
	Lead-time target																				
5	Median	-	-	-	-	-	-	-	-	-	1	+	+	+	+	+	0	0	0	0	0
	95th percentile	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	0	0	0	0	0
	100th percentile	-	-	-	-	-	1	1	1	1	1	+	+	+	+	+	0	0	0	0	0
10	Median	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	0	0	0	0	0
	95th percentile	-	-	-	-	-	-	1	-	-	1	+	+	+	+	+	0	0	0	0	0
	100th percentile	-	-	-	-	-	-	1	-	1	1	+	+	+	+	+	0	0	0	0	0
16*	Median	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	0	0	0	0	0
	95th percentile	-	-	-	-	-	-	1	-	-	1	+	+	+	+	+	0	0	0	0	0
	100th percentile	-	-	-	-	-	1	1	-	1	1	+	+	+	+	+	0	0	0	0	0
20	Median	-	-	-	-	0	-	-	-	-	1	+	+	+	+	1	0	0	0	0	0
	95th percentile	-	-	-	-	0	1	1	-	1	1	+	+	+	+	+	0	0	0	0	0
	100th percentile	-	-	-	-	0	1	-	1	1	1	+	+	+	+	1	0	0	0	0	0
25	Median	-	0	0	0	0	1	1	1	1	1	+	+	+	1	+	0	0	-	1	1
	95th percentile	-	-	0	0	0	1	1	1	1	1	+	+	+	+	+	0	0	1	1	1
	100th percentile	-	0	0	0	0	1	1	1	1	1	+	1	1	+	1	0	0	1	1	1
30	Median	0	0	0	0	0	0	-	-	1	1	+	+	+	1	1	1	-	1	1	1
	95th percentile	-	0	0	0	0	1	1	+	1	1	+	+	+	+	+	1	1	1	1	1
	100th percentile	-	0	0	0	0	1	1	1	-	1	+	1	1	-	1	1	1	1	1	1

* These are the historical RV's mean appointment lead-time and discharge rates at TTSH.

4.3.2 Sensitivity to changes of the discharge rates

In this part, we analyze the sensitivity of the maximum required capacity, the mean required capacity, the FV's mean appointment lead-time and the RV's mean appointment lead-time to the changes to the discharge rates. The maximum required capacity decreases when the discharge rate increases. This pattern exists for changes of both FV's and RV's discharge rates. However, this negative correlation is not always true when measuring the mean required capacity. The mean required capacity is not minimized by the model; hence, when the discharge rate increases, in order for the model to decrease the maximum required capacity, it might need to increase the mean required capacity.

In terms of appointment lead-time, the change of either FV's or RV's discharge rates does not impact the general way of providing appointments to patients, specifically FV patients. We find through the numerical results that the FV's and RV's mean appointment lead-times remain unchanged with changes of the discharge rates.

4.3.3 Guidelines for scheduling an appointment

We implemented the numerical experiment with 450 sets of inputs corresponding with 450 runs, which are combined from different levels of the discharge rates, the RV's mean lead-time, and the appointment lead-time targets. From these results we observe that the model solution tends to set the appointment lead-times of the FV patients to one of the appointment lead-time targets. That is, the optimal solution plans appointments for the FV patients to be with lead times of u time-units (the median target), or v time-units (the 95th percentile target) or w time-units (the 100th percentile target). Table 6 reports the percent of the FV patients whose appointments are set to one of the appointment lead-time targets. From 450 cases experimented, a minimum 3%, mean 39%, median 47%, and maximum 50% of the FV demands are assigned a median appointment lead-time. Similarly, a minimum 3%, mean 37.4%, and maximum 45% of the total number of FV demands are assigned a 95th percentile appointment lead-time.

Furthermore, the estimated value of FV's mean appointment lead-time,$\bar{l}$, should correspond to the MOH's targets in any time-unit during the model's period, when providing appointments to FV patients. Hence, the scheduling of patients' requests needs not only to pay attention to maintaining the satisfaction of the MOH's appointment lead-time targets but to preventing the risk of violation of the MOH's appointment lead-time targets in a long term as well.

Table 6 *Percentage of patients whose appointment lead-times are equal to one of appointment lead-time targets*

Targets of patients' appointment lead-time	Percentage of patients (%)			
	Min	Median	Max	Mean

Median	3.0	47.0	50.0	39.0
95th percentile	3.0	45.0	45.0	37.4
100th percentile	0.0	5.0	5.0	4.4

This finding implies that to achieve the appointment lead-time targets, a scheduling policy may focus just on these three appointment lead-time targets and the estimated value of mean appointment lead-time for FV patients. This is an interesting way of scheduling because it not only enables easy control of the appointment lead-time targets to plan capacity, but also gives more fairness to all patients making requests. Hence, this scheduling approach should be implemented in providing appointment dates to patients.

5. Discussions

We have proposed a mixed integer model to determine the maximum required capacity for the re-entry system, with a specified accounting for the appointment lead-time characteristics. The proposed model minimizes the maximum required capacity while achieving service targets on the median, 95th percentile, and 100th percentile for the appointment lead-times for FV patients and satisfying the constraints on the revisit appointment lead-times.

With numerical experiments, we show that the proposed model can provide systematic capacity planning for the re-entry system at the tactical level. The proposed model pushes and pulls the appointments to minimize the maximum required capacity. This push-pull mechanism of the proposed model is to delay the appointment to patients as long as possible without violating the FV's appointment lead-time targets and the RV's appointment lead-time constraints. Regardless of how the system pushes and pulls the appointments to optimize the required capacity, one of the appointment lead-time targets should be the appointment lead-time for FV patients' requests and the estimated value of FV's mean appointment lead-time should be within the MOH targets in any time-unit during the model's period. However, there is no specific trend in scheduling the appointments of RV patients. Therefore, a deeper investigation on the scheduling of RV patients is encouraged.

We conducted a sensitivity analysis to understand how the maximum required capacity, the mean required capacity, the FV's mean appointment lead-time, and the RV's mean appointment lead-time are affected by changes to the appointment lead-time targets, the discharge rates, and the restriction of RV's mean appointment lead-time.

- The maximum required capacity is negatively correlated with the appointment lead-time targets, the discharge rates, and the restriction of RV's mean appointment lead-time. However, this is not always true for the mean required capacity measure.

20

- The FV's mean appointment lead-time is positively correlated with the FV's appointment lead-time targets. However, the RV's mean appointment lead-time generally increases with increases in the restriction on the RV's mean appointment lead-time, until the latter becomes very large.

- Changing the discharge rates does not have much impact on the determination of the FV patient's appointment.

The proposed model does not guarantee that the total amount of required capacity is minimized because it optimizes only the maximum required capacity. This shortcoming might be addressed by adopting multiple objectives in planning the capacity. The study is limited by the consideration of a single category of patients, for which all patients have the same consultation times, the same FV appointment lead-time targets, the same restrictions of RV appointment lead-time, and the same discharge rates factors. For systems with different categories of patients, we could develop an extension to the model. Furthermore, we do not consider any uncertainty due to randomness in the arrival demands, or discharge rates, or no-show rates, in the proposed model. Consequently, the model may prescribe an underestimate to the capacity needed, in light of these uncertainties. Future research should examine how to account for multiple patient categories and the uncertainties in modeling and determining the required capacity.

6. Conclusions

In this study, we develop a mixed integer programme to optimize the maximum required capacity for outpatient clinics which are characterized as a re-entry system. The service targets of median, 95th percentile, and 100th percentile for patient appointment lead-times were investigated. Given the estimated FV arrival demands and the current FV and RV patients, and an appointment lead-time policy, we obtain an optimal planned capacity from the model. Then, the proposed guidelines for scheduling an individual patient' appointment are carried out to manage the service targets. The study found that there is a strong relationship between the service targets and how appointments are planned for the FV patients. In addition, we examine the sensitivity of the maximum required capacity, the mean required capacity, the FV's mean appointment lead-time, and the RV's mean appointment lead-time to changes to the service targets, the discharge rates, and the restricted length for RV's mean lead-time.

The assumptions of unified and deterministic arrival demands for FV patients, and identical discharge and no-show rates are arguably restrictive, and could limit the applicability of the model in a variety of scenarios. We expect that the required capacity obtained from the proposed model may underestimate what is needed, due to possible uncertainties in reality. Hence, to enhance the applicability of the proposed approach, future work should explore how to extend the model and its analysis to allow for multiple patient categories as well as uncertainties in arrivals, and in the patient's consultation times, no-show behavior and discharge rates.

Acknowledgement

Appendix

Appendix A1*: Sets of the appointment lead-time targets*

No	Lead-time category changed	Appointment lead-time targets (weeks)		
		Median	**95th percentile**	**100th percentile**
1	**Median**	*1*	6	9
2*		*2*	6	9
3		*3*	6	9
4		*4*	6	9
5		*5*	6	9
6	**95th percentile**	2	*3*	9
7		2	*4*	9
8		2	*5*	9
9*		2	*6*	9
10		2	*7*	9
11		2	*8*	9
12	**100th percentile**	2	6	*7*
13		2	6	*8*
14*		2	6	*9*
15		2	6	*10*
16		2	6	*11*
17		2	6	*12*

The same cases are presented for the purpose of sensitivity analysis

References

1. Bailey N.T.J. (1952) A study of queues and appointment systems in hospital out-patient departments, with special reference to waiting-times. Journal of the Royal Statistical Society 14(2), 185-199.

2. Bertsimas D., Tsitsiklis J.N. (1997) Introduction to Linear Optimization, page: 272-278. Athena Scientific, Belmont, Massachusetts, USA.

3. Bowers J., Lyons B., Mould G., Symonds T. (2005) Modelling outpatient capacity for a diagnosis and treatment centre. Health Care Management Science 8(3), 205-211.

4. Brahimi M., Worthington D.J. (1991) Queueing models for out-patient appointment systems-a case study. The Journal of the Operational Research Society 42(9), 733-746.

5. Chakraborty S., Muthuraman K., Lawley M. (2010) Sequential clinical scheduling with patient no-shows and general service time distributions. IIE Transactions 42(5), 354-366.

6. Cote M.J. (1999) Patient flow and resource utilization in an outpatient clinic. Socio-Economic Planning Sciences 33(3), 231-245.

7. Elkhuzen S.G., Das S.F., Bakker P.J.M., Hontelez J.A.M. (2007) Using computer simulation to reduce access time for outpatient departments. Quality and Safety in Health Care 16(5), 382-386.

8. Green L.V., Savin S. (2008) Reducing delays for medical appointments: A queueing approach. Operations Research 56(6), 1526-1538.

9. Harper P.R., Gamlin H.M. (2003) Reduced outpatient waiting times with improved appointment scheduling: a simulation modelling approach. OR Spectrum 25(2), 207-222.

10. Hassin R., Mendel S. (2008) Scheduling arrivals to queues: a single-server model with no-shows. Management Science 54(3), 565-572.

11. Ittig P.T. (1985) Capacity planning in service operations: the case of hospital outpatient facilities. Socio-Economic Planning Sciences 19(6), 425-429.

12. Kaandorp G.C., Koole G. (2007) Optimal outpatient appointment scheduling. Health Care Management Science 10(3), 217-229.

13. Klassen K.J., Rohleder T.R. (1996) Scheduling outpatient appointments in dynamic environment. Journal of Operations Management 14(2), 83-101.

14. Kourany R.F.C., Grarber J., Tornusciolo G. (1990) Improving first appointment attendance rates in child psychiatry outpatient clinics. Journal of the American Academy of Child & Adolescent Psychiatry 29(4), 657-660.

15. LaGanga L.R., Lawrence S.R. (2007) Clinic overbooking to improve patient access and increase provider productivity. Decision Sciences 38(2), 251-276.

16. Lakshmi C., Sivakumar A.I. (2013) Application of queueing theory in health care: A literature review. Operations Research for Health Care 2(1-2), 25-39.

17. Liu L., Liu X. (1998) Block appointment systems for outpatient clinics with multiple doctors. The Journal of the Operational Research Society 49(12), 1254-1259.

18. Muthuraman K., Lawley M. (2008) A stochastic overbooking model for outpatient clinical scheduling with no-shows. IIE Transactions 40(9), 820-837.

19. Pegden C.D., Rosenshine M. (1990) Scheduling arrivals to queues. Computers and Operations Research 17(4), 343-348.

20. Qu S., Rardin R.L., Williams J.A.S., Willis D.R. (2007) Matching daily healthcare provider capacity to demand in advanced access scheduling systems. European Journal of Operational Research 183(2), 812-826.

21. Qu X., Shi J. (2009) Effect of two-level provider capacities on the performance of open access clinics. Health Care Management Science 12(1), 99-114.

22. Qu X., Rardin R.L., Williams J.A.S. (2011) Single versus hybrid time horizons for open access scheduling. Computers & Industrial Engineering 60(1), 56-65.

23. Rising E.J., Baron R., Barry B. (1973) A systems analysis of a university-health-service outpatient clinic. Annals of Operations Research 21(5), 1030-1047.

24. Sudhakar-Krishnan V., Rudolf M.C.J. (2007) How important is continuity of care? Archives of Disease in Childhood 92(5), 381-383.

25. Wijewickrama A., Tkakuwa S. (2005) Simulation analysis of appointment scheduling in an outpatient department of internal medicine. Proceedings of the 37th Winter Simulation Conference, Dec. 2005, Orlando, FL, USA.

26. Zeng B., Turkcan A., Lin J., Lawley M. (2010) Clinic scheduling models with overbooking for patients with heterogeneous no-show probabilities. Operations Research 178(1), 1030-1047.