



HAL
open science

Daily outpatient chemotherapy appointment scheduling with random deferrals

Thierry Garaix, Salim Rostami, Xiaolan Xie

► **To cite this version:**

Thierry Garaix, Salim Rostami, Xiaolan Xie. Daily outpatient chemotherapy appointment scheduling with random deferrals. Flexible Services and Manufacturing Journal, In press, 10.1007/s10696-018-9326-x . hal-02056838

HAL Id: hal-02056838

<https://hal.science/hal-02056838v1>

Submitted on 4 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Daily Outpatient Chemotherapy Appointment Scheduling with Random Deferrals

Thierry Garaix · Salim Rostami · Xiaolan Xie

Received: date / Accepted: date

Abstract In this paper we propose a heuristic approach that computes the order in which patients will be treated in an ambulatory chemotherapy center. Each patient follows an individual treatment plan that fixes dates for series of drug injections separated by recovery periods. The daily care process has three steps: consultation with the oncologist, drug preparation in the pharmacy and drug injection in medical beds. The facility closes after the last injection. As drug injection varying considerably in duration – from 15 minutes to 6 hours – bad schedules lead to excessive overtime. In addition, after the consultation the oncologist may decide to cancel the injection because of a weak patient's health condition. In the current setting of the chemotherapy facility we work with, *First Come First Served* policy controls the care process. In this study, we propose to compute a common priority list of patients for consultation and injection phases. A unique list of patients is a simple tool used by nurses to manage the flow of patients and to react to uncertain events. A GRASP algorithm is developed to compute optimized list of patients in few seconds as the operating planning context requires. Two objectives are considered; the closing time and the overworking time of the facility. Numerical experiments show that our GRASP is able to quickly reach near optimal solutions and that list of patients policy performance is comparable to more complex scheduling policies. Benchmark data sets are built based on historical data of the French chemotherapy facility, ICL in Saint-Étienne.

Thierry Garaix
École des Mines de Saint-Étienne / LIMOS UMR CNRS 6158
Tel.: +334-77-426641
E-mail: garaix@emse.com

Salim Rostami
IESEG School of Management, Lille, France / ORSTAT, KU Leuven, Belgium

Xialoan Xie
École des Mines de Saint-Étienne / Shanghai Jiao Tong University, Shanghai, China

1 Introduction

Ambulatory care services offer is growing last years with the objectives to improve patient's experience and reduce costs compared to hospitalization. Several chemotherapy treatments are delivered in ambulatory services for a long time. Actually, the chemotherapy care pathway is designed according to a medical protocol. A protocol is defined by series of consultations, drug injections and recovery periods. As recovery periods may last several weeks, the ambulatory model is suitable for several chemotherapy treatments.

In a single day, the patient consults the oncologist and receives the injection. During the consultation, the oncologist examines the patient and analyses the results of the biological examinations. The oncologist can decide to defer the injection when the health condition of the patient is too weak. Otherwise, the dosage specifications are given to the pharmacy which starts the production of the drug. After the consultation – if the injection is prescribed – the patient has to wait drug and bed availability before getting the injection. The completion time of the last injection of the day gives the closing time of the facility. This completion time has to be minimized in order to minimize overtime.

This research is motivated by ICL¹ chemotherapy facility in Saint-Étienne, France. In this cancer center, each patient gets their next appointment before leaving the facility. The appointment day is determined by the protocol. The oncologists' consultation time slots are filled by hand from earliest to latest. Because of the adherence constraint to the protocol, the set of patients to schedule is known in advance. Note that the first injection of new patients is prescribed after a delay of several days; at least the time to get biological examination results. Durations of each step follow, with low variability, standards related to the type of treatment delivered. During the day, patients with available drugs are assigned to beds following a *first come first served* rule. The consultation and injection scheduling strategies are usually suboptimal and may generate avoidable overtime. The pharmacy is sized to cover the demand from different services and the ambulatory one is prioritized compared to hospitalization. The drug preparation duration therefore suffers low uncertainty.

Two decisions are supported in this work: the next consultation appointment schedule and the patient to bed assignment. In the first case, the proposed solution can be implemented into the planning software used by the secretary. But the running time has to be short, since the patients have to get their appointment before they leave the facility. Contrary to consultations, injections are uncertainly performed because of deferrals. Until 10% of the daily demand may be deferred. Therefore, an optimized complete schedule of injections can generate idle times on beds and suboptimal plans. A dynamic scheduling policy is then required. In the absence of real time information systems that tracks the activity in the cancer center, automatic rescheduling tools are not suitable in ICL as in many other cancer centers. The scheduling policy has to be easy to apply by the medical staff. Considering these constraints, we propose to compute a global sequence of patients for each day. This sequence is satisfied at each step. A patient is simply removed from the sequence if their injection is deferred. In this paper we investigate how to

¹ Institut de Cancérologie Lucien Neuwirth

build such global sequence and how global sequence management performs against other decision support approaches.

In Section 2, the proposed approach is compared against the literature on ambulatory chemotherapy and appointment scheduling with no-shows. The optimization problem of computing a global-sequence of patients is described in Section 3 with its mathematical formulation in Section 4. How to evaluate numerically the performance of a global sequence and the GRASP heuristic algorithm developed to compute near-optimal global sequences are detailed in Section 5. Two lower bounds and some easy to compute global sequences, named target sequences, are presented in Section 6. The numerical experiments of Section 7 have been conducted on a benchmark data set built from historical data provided by ICL. The results show the good performance of global sequences compared to several dynamic schedule policies applied to the bed assignment decisions and also the dominance of GRASP over the target sequences.

2 Literature Review

Appointment scheduling in highly constrained health care services has been identified as research challenges in 2008 [10]. Chemotherapy appointments make one of these challenges, as a multistep scheduling in a dynamic and uncertain context. The literature on the topic is scarce but some recent papers try to propose optimization methods and algorithms to manage appointments for chemotherapy or similar health care applications. We give an overview of the most recent papers closely related to the problem we are interested in.

This paper is most closely related to the series of work started with [23] that proposed a linear programming model and a three-stage heuristic for optimizing the medical planning, i.e. the work schedule of oncologists. As the patients lists of oncologists had very different size, an optimized assignment of oncologists consultation days other the week may smooth the activity in the facility.

Several Monte Carlo optimization approaches were proposed in [21] for determining the chemotherapy day of each new patient. Because of the protocol adherence constraints, the day of the first injection was kept by default for all the following injections. The deterministic daily schedule of chemotherapy was solved by a Lagrangian relaxation in [22]. Given a set of patients, their appointment were scheduled in order to minimize the closing time of the facility. This series of works showed significant reduction of the maximum daily bed capacity requirement. The constraints considered in that paper were similar to ours and our paper can be seen as an extension of that previous one by taking into account the random deferrals of chemotherapy sessions.

[5] proposed a rather qualitative study of next-days chemotherapy scheduling to book chemotherapy session in the following day of consultation session. The results support that using this regulation improves efficiency of pharmacy and nursing and decreases waiting times of patients in receiving their chemotherapy. Later, a survey ([16]) showed that the majority of studied patients prefer chemotherapy session to be on the same day as consultation session. Convenience, long distances and difficulty of transportation were reported as the main reasons.

With uncertainties playing a less important role or being neglected, a simulation based evaluation method was used in [1] to investigate different scheduling

templates which could be used to schedule patients' arrival to the clinic. The goal is to minimize patients' waiting time and to increase the utilization rate of resources. A mathematical formulation was proposed in [2] to optimally schedule patients per day while balancing the nursing workload throughout that day. Discrete event simulation was also used to investigate impacts of changing the nurse-to-patient ratios.

A two-phase meta-heuristic approach was proposed in [17] to schedule patients over a several days planning horizon. In the first step an initial solution was created with a Greedy algorithm considering various strategies, while in the second stage a Tabu search was used to improve the best strategy of the first step. Results on a case study showed the efficiency of the proposed algorithm in scheduling more patients than the practical approach without imposing any overtime and with a better balanced workload between nurses. A discrete event simulation model of daily chemotherapy outpatient operations was proposed in [26] in another attempt to reduce patients' waiting time and increase utilization ratio of beds for a case study. Bed utilization was held to a tolerance level by controlling the excess waiting time.

Outpatient appointment scheduling with individual day-dependent no-show predictions was addressed in [24]. The daily schedule was single task (like one consultation). Target sequences including show and no-show patients that minimize delays and makespan were identified. This fast evaluation was embedded in a stochastic optimization model where the appointment days were scheduled for one real patient and a set of future fictitious patients. The finite set of scenarios of fictitious patients was derived from historical data. The selection of the day of the appointment also is the focus of an algorithm based on Markovian Decision Process in [9], in the context of chemotherapy. A linear-programming-based Approximate Dynamic Programming was employed to obtain near optimal solutions. Then the heuristic results were compared to the ones obtained by practical and easy-to-use heuristic decision rules under diverse scenarios using simulation.

In the context of nuclear medicine, a similar appointment scheduling problem was solved in [20]. A stochastic program modeled a complex care pathway but without no-shows and with unknown patients. The scenarios represented possible demand realizations and the solution approach was based on the Monte Carlo method.

In [3], the authors study of optimal schedules for patient appointment booking, sequencing, and scheduling decisions for outpatient procedure centers. They conduct several numerical experiments with a two-stage stochastic mixed integer model. They provide some insights to practitioners on overbooking strategies and good sequences. They conclude that sequencing patients with high no-show rates or high procedure duration variance later in the day results in lower costs of waiting, idling, and overtime.

In order to address the dynamic optimization of chemotherapy outpatient scheduling with uncertainty, a two-phase approach was proposed in [12]. A template for the schedule with known and forecast demands via a constraint programming algorithm, was detailed in [11]. That template was updated when new demands were revealed by a shuffling algorithm.

In [6], the authors present a stochastic integer programming model for dynamic sequencing and scheduling of appointments to a single stochastic server. They

provide several interesting properties of best schedules. But their considerations are mainly related to waiting times, that are not relevant in our study.

A single-server scheduling problem given a fixed sequence of appointment arrivals with random no-shows and service durations was optimized in [14]. Appointment waiting, server idleness and overtime were penalized in a robust optimization model. The analysis of their optimal schedules may be used to evaluate good sequences when no-shows occurs at the consultation step. But these results can not be directly applied to our study case with parallel machines (beds).

This study differs from the cited references as we are not directly interested in solving the dynamic problem of appointment scheduling where patients to schedule are revealed online. We want to characterize good schedules when uncertainty relates to the availability of patients to receive their injection.

The results of this research aim to be used in three ways. Firstly, as guidelines to build good scheduling strategies for the dynamic appointment scheduling of consultations. Secondly, as a decision tool to give the next appointment to a patient before they leaves the facility. Thirdly, as a subroutine which provides daily appointment schedules in solution algorithms that schedule appointments for several patients over single or multiple days planning horizon.

3 Problem description

A brief explanation on chemotherapy treatments (Section 3.1) along with the key assumptions considered in this research (Section 3.2) are provided in this section.

3.1 Outpatient Chemotherapy Care Pathway

Disease-specific appointment scheduling problems are highly influenced by the type and the characteristics of the disease. As a consequence, we first describe the basics of the chemotherapy procedures before the mathematical formulation of the planning optimization problem we address.

When a patient, who is going to receive chemotherapy treatment, arrives to the medical facility for the first time, they will be assigned to an oncologist who will be the referee physician of the patient for the rest of treatments. A protocol (treatment plan) is decided for the patient based on different factors including cancer type, health condition, age, gender, etc. A protocol identifies the number of required injections, time intervals between injections, and drug types. An example is given in Table 3.1. Some protocols are often repeated, including some recovering periods. In practice, the protocol can not be perfectly satisfied because of the deterioration of the patient's health condition or closing days the facility during national holidays.

Each appointment is consisted of different parts. In the chemotherapy facility that this research is motivated by, the procedure is as follows:

1. *Bio samples*: One day before their appointment time, patients must take a blood test to assess their health condition and eligibility to receive further injections. This test may reveal their inability of receiving injection at the appointment of the next day. In this case the chemotherapy appointment will be postponed to another time in future.

Table 1 R-CHOP protocol

drug	administration	unit doses	time frame
Rituximab	injection	$375mg/m^2$	day 1
Cyclophosphamide	injection	$750mg/m^2$	day 1
Doxorubicine	injection	$50mg/m^2$	days 1 and 21
Vincristine	injection	$1.4mg/m^2$	days 1 and 42
Predisolone	oral	$40mg/m^2$	days 1 to 7

2. *Consultation:* Patients are visited by their referee oncologist. The oncologist decides whether to authorize the injection and if so they decides on the dosage of drugs that should be injected and send the prescription to the pharmacy. Note that at this point, the oncologist has an updated and more reliable estimation of patient's health condition in next appointment. This estimation is vital in the process of decision making about the time of next appointment as it reveals two major pieces of information, (1) the dosage of drugs to be injected in next appointment which implies required injection time, (2) probability of the patient not being able to receive any injection in their next appointment time due to critical health condition – the no-show probability.
3. *Drug preparation:* After the consultation session, the patient has to wait for prescribed drugs to be prepared and sent by the pharmacy. This preparation may take different amounts of time based on types of requested drugs.
4. *Injection:* By arrival of requested drugs and a bed is available, a nurse prepares the patient and starts the chemotherapy injection.
5. *Appointment set up:* Before leaving the facility, the patient gets their next appointment schedule from the registry office.

In this research we aim to show that a precomputed global sequence, common to consultation and injection steps, is an insightful alternative to classical First-in-First-Out our Longest-processing-time-first policies applied in practice. In the context of ICL, the global sequence is computed at the appointment set up step in order to give the date of the next consultation to each patient. The global sequences can be used in different contexts, like scheduling appointments for several patients in a single run. In the next section, we define and comment the main assumptions on the system addressed in this paper.

3.2 Assumptions

The assumptions of this research are as follows:

- (A1) All the patients to schedule are known.
- (A2) All beds are identical and suit every patient during injection time.
- (A3) The durations of the consultation sessions are equal for all patients.
- (A4) Durations of drug preparation and drug injection are variable among patients and yet known for each patient. Moreover, the pharmacy is not a bottleneck and drugs arrive in time.
- (A5) Each patient could only be visited by their referee oncologist.
- (A6) Oncologists are available throughout the scheduling day and no idle-time is allowed in their schedules.
- (A7) No-show probability for the consultation session is zero.

- (A8) All patients arrive on-time for the consultation sessions.
- (A9) An adequate number of nurses is assumed to be always available to prepare the beds and the patients for the injection as soon as the drugs are ready. The presence of a nurse is not required during the injection.

All assumptions except (A6) et (A7) are similar to those of our previous works [23,21,22] and we do not repeat justifications presented in our previous relevant publications. Assumption (A6) on the availability of oncologists is only introduced for simplicity of the presentation and may be relaxed later. Assumption (A7) is motivated by our field observation that, for continuity of care and for better monitoring, the consultation appointment is often kept even if the bio sample reviews the inability of the patient for chemotherapy the next day. The consultation appointment is canceled only for patients with really bad health conditions. Note that when an injection is deferred, it is usually postponed to the next session scheduled by the protocol. The start of a new protocol is scheduled in advance, at least one week, in order to conduct preliminary medical exams. For these two reasons, Assumption (A1) that all patients to schedule are known holds even under the dynamic context where appointments are assigned to each patient before they leave the cancer center.

4 Problem Formulation

Under the on-going assumptions, the problem under consideration is characterized by a set P of chemotherapy patients, a set O of oncologists and a set B of identical beds. Each patient is identified by a referee oncologist, a common consultation time d_1 with their referee oncologist, a known yet variable drug preparation time d_{2p} , an injection time in a bed d_{3p} and a deferral probability λ_p that the patient will leave after the consultation (without any injection or drug preparation). The set P is partitioned into PO_o (set of patients of oncologist o). Each oncologist consults all their patients without interruption. After the consultation, a patient who is not deferred is ready for injection d_{2p} time units later, the injection starts if a bed is available, the bed is held by the patient during the injection duration d_{3p} . The goal is to minimize the expected makespan or the expected overtime, i.e. the opening the outpatient facility beyond some regular closing time.

A feasible solution considered in this paper is a global sequence π of all patients which gives oncologist-specific sub-sequences for the consultation stage as well as a global priority list for bed assignment. The set PO_o of each oncologist o is ordered according to the global sequence π and are given appointment times $0, d_1, 2d_1, \dots$. A patient is assigned to a bed if all preceding patients in sequence π either have started their injection or have left without injection.

Imposing the same global sequence for consultation and bed allocation leads to simple problem formulation and reduced solution space. A potential drawback is the degraded bed utilization due to the idle time of beds waiting for high priority patients. Numerical experiments will be conducted to show that this potential loss is not significant for realistic cases.

A discrete time model will be used in this paper with the time horizon divided into T identical time slots of length l . In our case study, the consultation time of each patient takes 15 minutes and d_{2p} and d_{3p} could be quantified as integer multiples of this duration. Thus we consider time slots of 15 minutes.

Scenario-based formulation will be proposed in this paper. Each scenario k is a realization of deferral or permission status of injection of patients. The total number of scenarios for set P would be equal to $2^{|P|}$.

4.1 Notations

We consider the following notations:

Given sets:

P : Set of all patients who based on their treatment plans are expected to need an appointment on the scheduling day

O : Set of oncologists

PO_o : Set of patients of oncologist $o \in O$ (PO_o sets are a partition of P)

PS_k : Set of patients able to receive their injection in scenario k ($PS_k \subset P$)

Parameters:

K : Number of scenarios

T : Number of available identical time slots on the scheduling day

d_1 : Number of time slots required for consultation of a patient

d_{2p} : Number of time slots required for drug preparation of patient p

d_{3p} : Number of time slots required for injection of patient p

λ_p : Estimated injection-deferral probability of patient p

B : Number of available beds

ξ_k : Vector $(\xi_{k1}, \dots, \xi_{k|P|})$ of scenario k where ξ_{kp} is equal to 1, if in scenario k , patient p is able to receive injection, or is equal to 0 otherwise

Decision Variables:

$x_{pp'}$: Binary variable equal to 1 if and only if patient p has higher priority than patient p' .

Auxiliary Variables:

z_{ptk} : Binary variable equal to 1 if and only if the injection of patient p starts on time slot t in scenario k

y_p : Drug preparation start-time of patient p

s_{pk} : Injection start-time of patient p in scenario k

u_k : Total makespan in scenario k

4.2 Mathematical Model

$$\min \sum_{k=1}^K q_k \cdot u_k \quad (1)$$

s.t.

$$x_{pp'} + x_{p'p} = 1 \quad \forall (p, p') \in \{P \times P | p \neq p'\} \quad (2)$$

$$y_p \geq d_1 \cdot \left(\sum_{p' \in PO_o} x_{p'p} + 1 \right) + 1 \quad \forall p \in PO_o, \quad o \in O \quad (3)$$

$$\sum_{t=1}^T z_{ptk} = 1 \quad \forall p \in P, \quad k \in \{1, \dots, K\} \quad (4)$$

$$s_{pk} = \sum_{t=1}^T t \cdot z_{ptk} \quad \forall p \in P, \quad k \in \{1, \dots, K\} \quad (5)$$

$$s_{pk} \geq y_p + d_{2p} \cdot \xi_{kp} \quad \forall p \in P, \quad k \in \{1, \dots, K\} \quad (6)$$

$$s_{pk} \geq s_{p'k} - (1 - x_{p'p}) \cdot M \quad \forall (p, p') \in \{P \times P | p \neq p'\}, \quad k \in \{1, \dots, K\} \quad (7)$$

$$\sum_{p \in PS_k} \sum_{t'=t-(d_{3p}-1)}^t z_{pt'k} \leq B \quad \forall t \in \{1, \dots, T\}, \quad k \in \{1, \dots, K\} \quad (8)$$

$$u_k \geq s_{pk} + d_{3p} \cdot \xi_{kp} \quad \forall p \in P, \quad k \in \{1, \dots, K\} \quad (9)$$

$$x_{pp'}, z_{ptk} \in \{0, 1\} \quad \forall p, p' \in P, \quad t \in \{1, \dots, T\}, \quad k \in \{1, \dots, K\} \quad (10)$$

where q_k is the probability of scenario k and it is calculated as follows:

$$q_k = \prod_{p \in PS_k} (1 - \lambda_p) \cdot \prod_{p \in P \setminus PS_k} \lambda_p = \prod_{p \in P} (\xi_{kp} + \lambda_p - 2\xi_{kp}\lambda_p) \quad \forall k \in \{1, \dots, K\} \quad (11)$$

Vector $X = \{x_{pp'} | p, p' \in P\}$, represents the selected global sequence. Vector $Y = \{y_p | p \in P\}$ is equivalent to time when the patient's consultation session is finished. Since all patients are present for their consultation sessions, vector Y is independent from the scenarios. It is worth mentioning that variables z_{ptk} are defined as integer binary variables (and not as continue variables) as it simplifies checking bed capacity in Equation (8). Note that vectors Y and S are included into the model for the sake of simplicity of understanding the model, and they could be easily replaced by their defining equations (3) and (5).

For the expected overworking minimization objective function, equation (1) should be changed to:

$$\min \sum_{k=1}^K q_k \cdot \max(u_k - T, 0) \quad (12)$$

4.3 Complexity

The makespan minimization of our problem is strongly NP-hard (NP-hard for any given number of beds) as any parallel machine scheduling problem (see [8] for its complexity) can be polynomially reduced to a special case of our problem with one patient per oncologist with equal drug preparation time and without deferral. Our problem becomes a standard parallel machine scheduling problem if we neglect the consultation, drug preparation and the deferrals. These results can be easily extended to the minimization of overworking cost which depends only on the makespan.

The mathematical model of the previous section approximately includes $|P| \cdot T \cdot K$ variables and $K \cdot [|P| \cdot (|P| + T + 2) + T]$ constraints. Consequently, for a realistic problem instance which could include 40 patients and 40 time slots per day, the model would include $K = 2^{40} \approx 1.1 \times 10^{12}$ scenarios, 1.76×10^{15} variables and

3.65×10^{15} constraints. Consequently, this model is astronomically large to be solved directly.

The complexity comes from two different sources: (i) the combinatorial solution space, *i.e.* the set of all possible global sequences, and (ii) the exponential complexity of the solution evaluation, *i.e.* the evaluation of the $2^{|P|}$ potential scenarios.

In the following sections we propose methods to overcome these two problems in order to find (near-) optimal solutions.

5 Solution Algorithm

In this section we propose a GRASP (for Greedy Randomized Adaptive Search Procedure) heuristic [7] to obtain sequences of high quality for instances with more than 40 patients. Before we explain how to numerically evaluate the performance of a solution with the objective function (1).

5.1 Solution Evaluation

This subsection addresses the evaluation of a given solution, *i.e.* the computation of the criterion value $Z(X)$ (makespan or overworking cost) for a given global sequence X . Two approaches are considered.

The first straightforward approach is to enumerate the set Ω of all possible scenarios, evaluate the makespan and overworking for each scenario, and determine the expected criteria values using equations (1), (12) and (13).

The second approach, termed Random Scenario Sampling (RSS), consists in (i) uniformly sampling a given number K of scenarios, *i.e.* each patient i is sampled with probability λ_p in each scenario; (ii) evaluate the criterion value u_k of each scenario, and (iii) averaging the results of (1), *i.e.*

$$Z(X) \approx \frac{1}{K} \sum_{k=1}^K u_k(X) \quad (13)$$

Due to the huge number of possible scenarios, the first approach is possible only for small size problems. The estimation of the RSS approach is statistically unbiased and converges with probability 1 to $Z(X)$ as the number of scenarios increases.

5.2 GRASP Heuristic Algorithm

GRASP typically consists of iterations made up from successive constructions of a greedy randomized solution and a subsequent local search. In our algorithm, the local search step is replaced by an improvement procedure – such variant can be called semigreedy heuristic [13].

The core of our algorithm is the constructive procedure described in Algorithm 1. A list X of patients is incrementally built in order to give a new solution. Patients are iteratively added to X according to a restricted candidate list (RCL)

of best solutions that have already been visited. At each iteration, one solution X' of RCL is randomly (uniformly) chosen, the number np of remaining patients of X' that do not belong to X are then added to X in the same relative order.

In the first GRASP iterations when few solutions are available as $|RCL|$ is lower than *WorkSize* parameter, with a low frequency (controlled by *pRand* and *pBiasedRand*), two other procedures are used to build X . The first procedure randomly selects one new patient as the next patient. In the second approach, the selection of the next patient is done through a *Regret based biased random sampling*. The probability ρ_j of selection of patient $j \in P \setminus X$ is determined by means of the following function:

$$\rho_j = \frac{d_{3j} + 1}{\sum_{i \in P \setminus X} (d_{3i} + 1)}.$$

Once a complete X is at hand, we apply *double justification* technique ([18,19,25]) which is shown to be an effective method to improve the quality of scheduling solutions without substantially more computational efforts. First, a schedule \mathbf{S} is built by applying X to a single scenario with all the patients receiving their injection. A double justification consists of shifting injection timeslots to right as far as possible in non-increasing order of their finish time without altering the total makespan of the schedule and then re-shifting them to left in non-decreasing order of their finish time. The justified \mathbf{S} is then reconverted into a list by ordering patients in non-decreasing order of their injection start time.

In order to maintain the restricted candidate list of best solutions and the best known solution, the list obtained from the constructive procedure is evaluated through simulation. In our application evaluation is time consuming. We observe experimentally, that a low number of replications that allows building more solutions, gives better results. The best value found for this parameter is 10 replications for a complete runs of 10,000 iterations on 40-patients instances.

Algorithm 1 List construction

```

1:  $RCL$  is the current list of solutions
2:  $X = \emptyset$ 
3: while  $|X| < |P|$  do
4:    $\tau = UNIF([0, 1])$ 
5:   if  $\tau > pRand + pBiasedRand$  and  $|RCL| = WorkSize$  then
6:      $X' = UNIF(RCL)$ 
7:      $np = UNIF([0, |P \setminus X|])$ 
8:      $X = X +$  the  $np$  first patients of  $X' \setminus X$ 
9:   else if  $\tau < pRand$  then
10:     $p = UNIF(P \setminus X)$ 
11:   else
12:     $p = RegretBasedSample(P \setminus X)$ 
13:   end if
14: end while
15:  $\mathbf{S}$  = Schedule obtained by applying  $X$  to a scenario with no deferrals
16: Double justification applied to  $\mathbf{S}$ 
17:  $X$  = List obtained from improved schedule  $\mathbf{S}$ 
18: Return  $X$ 

```

It is worth mentioning that other reference lists such as “Inverse solution” were studied to extend the diversity of produced solutions but the results were not satisfactory.

6 Lower Bounds and Target Sequences

In order to evaluate the performance of the GRASP algorithm on instances intractable for our linear program solver, two lower bounds and some reference global sequences, called target sequences, are proposed in this section.

6.1 Lower Bounds

Two different lower bounds are considered in this paper. The first is the “Linear relaxation” lower bound obtained by relaxing the integrity constraints of the binary variables $x_{pp'}$, z_{ptk} in our mathematical model. The second is the so-called “Perfect information” lower bound by replacing the single global sequence with scenario-dependent global sequences. As a result, our problem decomposes into independent sub-problems, one for each scenario to determine the scenario-dependent global sequence to minimize the overall criterion for this scenario. More precisely, in order to obtain the perfect information lower bound in our problem we iteratively solve the problem considering only one scenario. For each scenario the optimal sequence is obtained. Then the average of the makespan values of the optimal sequence of each scenario is known as perfect information lower bound.

Although the computational complexity of the lower bounds is significantly lower than that of the original problem, they still suffer from the complexity related to the exponentially increasing number of scenarios to consider. For this reason, these lower bounds will only be used for problems of small sizes.

6.2 Target Sequence

In addition to our GRASP algorithm, we consider four other strategies to build patient sequences, “Longest processing time first”, “Longest expected processing time first”, “Highest injection probability ($1 - \lambda_p$) first” and the inverse of the LEPT sequence, respectively denoted LPT, LEPT, HIP, LEPTinv.

The identification of good and simple target sequences has important practical impact, since decision makers may apply a target sequence without any optimization software.

Among all possible target sequences “Longest Processing Time First (LPT)” and “Longest Expected Processing Time First (LEPT)” are of particular interest. Due to the significant longer injection duration, the injection duration is used in LPT and LEPT uses the injection duration multiplied by the injection probability ($1 - \lambda_p$).

We now provide justifications for the selection of these two rules. The second stage of our problem (injections) has the most impact in the makespan and overworking in the sense that the average of injection durations is significantly longer

than consultation and drug preparation durations. Further the second stage problem is similar to a parallel machine scheduling problem for which LPT is known to perform well. [4] prove that LPT has a performance guarantee of 1.5 in an on-line parallel machine scheduling problem and no on-line algorithm could have a performance guarantee less than 1.3473. From Johnson's Rule ([15]), for a two-stage flow shop $F2||C_{max}$ with identical and shorter first-stage processing time, the optimal schedule is a global sequence sorted by LPT in the second stage. In the special (extreme) case with unlimited number of beds, identical drug preparation time and no deferral, LPT provides the optimal schedule.

LEPT is a natural extension of LPT to account for the deferrals. The deferral probability can be modeled as an uncertain injection duration that can take only two values: 0 if the injection is deferred and its deterministic duration otherwise. In the case with one oncologist, one bed and equal drug preparation and injection times, least deferral probability first is the optimal schedule.

HIP and LEPT_{inv} are proposed as a priori bad global sequences. HIP focuses on deferral probabilities and does not consider injection durations. LEPT_{inv} should perform worse than LEPT as its reverse sequence.

The comparison results regarding target sequences and the best sequences obtained by GRASP algorithm are provided in Section 7.2.

7 Computational Results

Several numerical experiments have been conducted in order to show the interest of global sequences and the performance of our GRASP algorithm to find out high quality global sequences.

The evaluation of the quality of global sequences is based on the the expected makespan (or overworking time) which is obtained by means of a simulation with sufficiently large number (10^7) of replications for the final solutions.

In the experiments presented in this section, the GRASP algorithm has been limited to 10,000 iterations. With this setting, the maximum running time over all instances is 600 seconds.

The datasets used in our experiments are described in Section 7.1. In Section 7.2, the results concerning the effectiveness of different global sequences in terms of expected makespan and overtime minimization are compared. In Section 7.3, the global sequence approach is compared against different scheduling policies that do not impose the same sequence on both consultation and injection steps.

All the experiments of this research have been performed on a personal computer with Intel®Core™ i7-3770 CPU with 3.40 GHz clock speed and 8.00 GB RAM. The codes were written and run in C++ using Microsoft Visual Studio Express 2012 for Windows desktop programming environment under the Windows 7 Professional x64 operating system.

7.1 Test Instances

Seven datasets with 100 instances, named *Basic Datasets* have been generated with a parameter setting which corresponds to the ICL chemotherapy facility.

Using similar parameters, the 300 small size instances of the *OptSize Dataset* have been built in order to solve to optimality the mathematical model presented in Section 4.2. Two other group of six datasets of 100 instances, the *Short-Art Datasets* and the *Long-Art Datasets*, have been generated in order to consider different settings on injection duration length and variability.

We investigated the historical data of ICL corresponding to a two-month period of the facility from December 2013 to February 2014.

All durations are given in time-slot units; a time-slot is equal to 15 minutes.

Four parameters have a fixed value. The daily *number of patients* is set to 40. There are 6 *available beds* for injections. There are 6 *oncologists* present at each time-slot to perform consultation sessions. The *consultation duration* is 1 time-slot for all patients.

The *referee oncologist* of each patient is chosen randomly and uniformly.

The *drug preparation duration* is randomly sampled to be either 1 or 2 time-slots.

The *injection duration* is selected based on the probability distribution of Table 2.

Table 2 Probability distribution of injection durations in time-slots of 15 minutes

d_{3p}	Probability	d_{3p}	Prob.	d_{3p}	Prob.	d_{3p}	Prob.
20	0.0090	14	0.1084	6	0.1067	2	0.2133
18	0.0098	12	0.1244	4	0.2631	1	0.0400
16	0.0231	8	0.0240	3	0.0782		

Each *Basic Dataset* corresponds to one of the following *average deferral probability* (overall patients) $\gamma \in \{0.1 ; 0.15 ; 0.2 ; 0.3 ; 0.4 ; 0.5 ; 0.6\}$. Historical data meet low average deferral probabilities ($\gamma \leq 0.2$). Instances with higher values of γ are introduced to cover some contexts with more deferrals.

The possible *individual deferral probabilities* are limited to five key values. These key values could be different over all instances. When $\gamma \leq 0.2$, key values are lower than 0.5. This constraint is consistent with the operation of ICL where appointments of patients with high deferral probability would be postponed to another time based on their treatment plan.

Unless otherwise mentioned, the *Basic Datasets* are used in all sections of this paper. In order to study some extreme or special cases, three additional datasets are generated with the following specifications.

- The *OptSize Dataset* is generated following the same procedure but with reduced number of patients and resources. For each of the three combinations of the number of patients and the number of beds ($(|P|=5, B=5)$, $(5, 3)$ and $(8, 3)$), 100 instances with 1 oncologist are generated.
- For the *Short-Art Dataset* and the *Long-Art Dataset*, injection durations are not determined by Table 2 but they are randomly chosen from uniform distributions $U([2, 4])$ and $U([10, 12])$, respectively.

7.2 Performance Evaluation of GRASP and Target Sequences

This subsection compares the four target sequences and the GRASP solution for both objectives, makespan and overtime minimization. Overall, LPT and LEPT outperform others target sequences and perform close to GRASP (GRASP reaches best solutions for most part of the runs) for both objective.

We carried out four series of numerical experiments. First, we compared on small size instances target sequences with optimal global sequences of Model (1)-(11) and two classical lower bounds directly derived from this model, “Linear relaxation” and “Perfect information” described in Section 6.1. Detailed results on *OptSize Dataset* are not reported here. The main observations on small-size instances include: (i) LPT and LEPT are very close to optimal global sequences with less than 2% and 4%, respectively; (ii) lower bounds are not very tight with a gap greater than 3%; and (iii) as expected HIP and LEPTinv provide the worst solutions with optimality gap greater than 7%.

In the second series of experiments on the *Basic Datasets*, we minimize the expected makespan. These results reported in Section 7.2.1 allow verifying the dominance of GRASP over the target sequences. In the two last series of experiments with *Short-Art* and *Long-Art Datasets* of Section 7.2.2, we minimize the overworking time objective. The results obtained reinforce the statement on the dominance of GRASP over other global sequence strategies.

7.2.1 Makespan minimization

Consider first the results of makespan minimization for the *Basic Datasets* given by the Figure 1. For low average deferral probability ($\gamma \leq 0.2$) corresponding to the real-world setting, LEPT and LPT perform the same. As deferral probabilities decrease, expected injection time of each patient gets close to its real injection time. Consequently, LEPT and its performance gets closer to LPT. On the other hand, with high deferral probability ($\gamma \geq 0.3$) LEPT performs worse than LPT.

In these instances, patients with long injection time and low deferral probability have low expected processing time and consequently low priority in LEPT. Such a sequence would schedule the injection session of such patients at the end. This decision is reasonable in scenarios without such patients. But in scenario where these patients are present this decision would lead to significant increase in the makespan.

Finally, as expected, by relying solely on the deferral probabilities and disregarding processing times, HIP leads to undesirable increase in expected makespan and only performs slightly better than LEPTinv.

It is also worth mentioning that each dataset is consisted of multiple instances (here 100) and the values presented in the figure is the average of the expected makespan over all instances of the dataset. In fact, GRASP performs worse than the best target sequences (LEPT/LPT) in a few single instances of the datasets (approx. 3% – 5%).

We provide a further sensitivity analysis to see whether the dominance of GRASP and LPT over LEPT is dependent on the instances and parameters.

Theoretically we expect to see different results when the variation of injection durations decreases, meaning that generally, patients have more similar injection durations. In this case the explained weakness of LEPT is expected to be resolved.

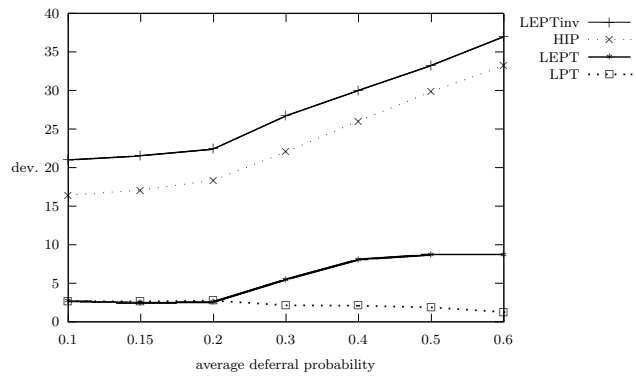
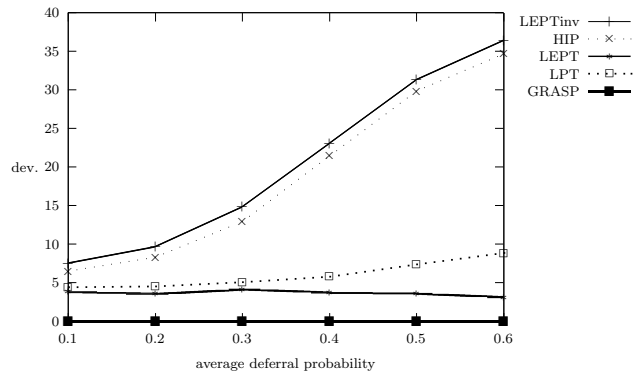
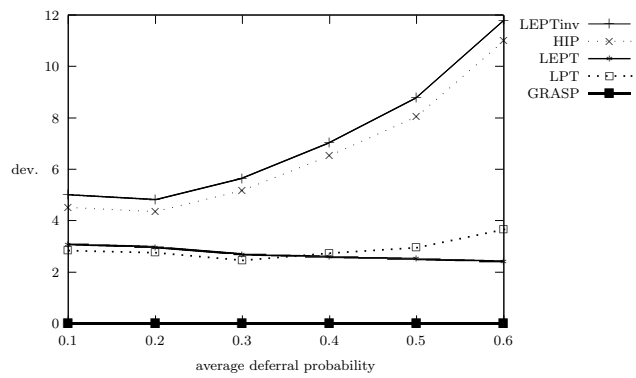
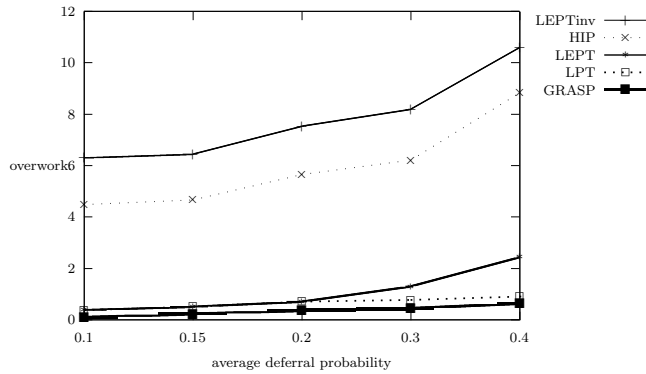
Fig. 1 Percentage deviation from GRASP of different target sequences**Fig. 2** Deviation percentage from the best makespan for short injection durations and small variation**Fig. 3** Deviation percentage from the best makespan for long injection durations and small variation

Fig. 4 Expected overworking time vs deferral probability

In order to test this case we produced two new datasets called *Long-Art* and *Short-Art* defined in Section 7.1 instead of the *Basic Datasets* used in Figure 1. The results are depicted in Figure 2 and Figure 3.

For low average deferral probability ($\gamma \leq 0.2$), all target sequences perform similarly, as there are little difference between the patients regarding the deferral probability and injection duration.

In the instances with high deferral probability ($\gamma \geq 0.4$) LEPT performs noticeably better than LPT.

7.2.2 Overworking minimization

The results of the different global sequences in overworking time minimization on the *Basic Datasets* are presented in Figure 4.

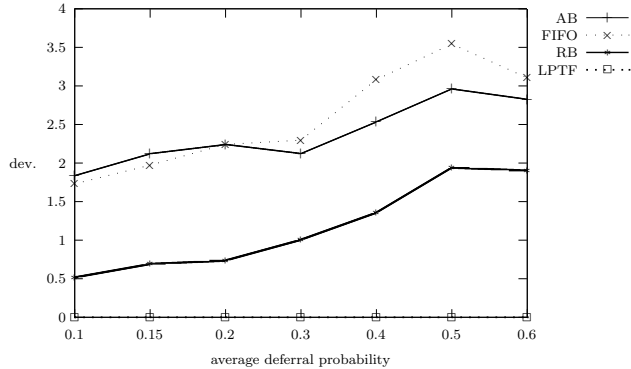
As it is shown in the picture, the results for overworking time is similar to the ones for makespan minimization. GRASP outperforms the other sequences. Among the target sequences, LPT and LEPT perform the same as each other and slightly worse than GRASP in the real-world setting (the instances with low deferral probability ($\gamma \leq 0.2$)). On the other hand, in the instances with high deferral probability ($\gamma \geq 0.3$), LPT outperforms LEPT with the same explanation we provided in the case of makespan minimization.

7.3 Relevance of the Global Sequence Approach

The goal of this subsection is to investigate the relevance of the global sequence approach by checking the insensitivity of the optimal global sequence-based solution and the performance ranking of the target sequences with respect to bed assignment strategies.

Using a global sequence may impose some deliberate idle-time on the beds if the consultation session of the next patient with the highest priority is not finished yet, even if there are some other patients with lower priority available. In this section, we compare global sequence approach with three other approaches

Fig. 5 Percentage deviation from the best solution found with four bed assignment policies in GRASP



that relax these constraints. These approaches are defined by their bed assignment policy. We propose four policies “Activity Based (AB)”, “Resource Based (RB)”, “Longest Processing Time First (LPTF)” and “First-In-first-Out (FIFO)”. The AB policy, is to use the global sequence as a given priority list which should be strictly respected in both stages. In this strategy there are start-to-start precedence constraints between patients, meaning that injection session of patient i could not be started earlier than that of patient j , if patient j has higher priority than patient i according to the global sequence.

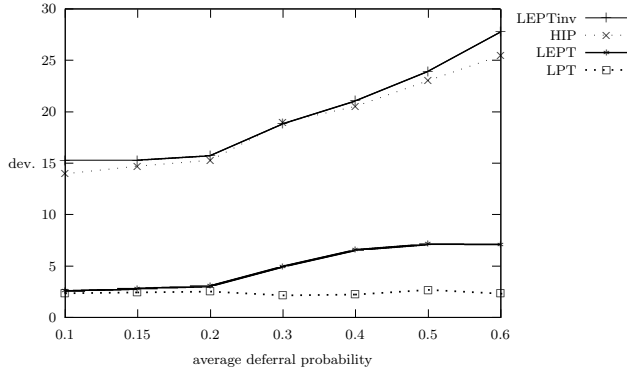
The second policy removes the start-to-start precedence constraints of the patients. Hence at each decision point the patient with highest priority among the available patients would be chosen. We refer to this strategy as RB strategy.

In the third and fourth policies, the global sequence only applies for the consultation stage. In the injection section, however, the chosen patient is the available one with highest processing time, or the first arrived first served rule applies. We call these two LPTF and FIFO policy, respectively. FIFO is the policy that is used by default in the most chemotherapy facilities.

7.3.1 GRASP Sequence Combined with four Bed Assignment Policies

We first ran our GRASP algorithm combined with four different bed assignment policies in order to evaluate the performance loss of enforcing the global sequence at the bed assignment stage.

Figure 5 shows that AB policy with the same global sequence priority at both consultation and injection stages is dominated by RB and LPTF but the deviation is rather small by less than 3% (each point gives the average deviation percentage over 100 instances) on the seven sets of instances of the *Basic Datasets*. The dominance of LPTF increases with deferral probability, as LPTF is a good policy in a pure dynamic context. Surprisingly, the AB policy dominates the FIFO policy when the deferral probability is high and has similar performance when the deferral probability is low.

Fig. 6 Percentage deviation from GRASP of different target sequences combined with FIFO policy

7.3.2 Target Sequences Combined with four Bed Assignment Policies

We now evaluate the ranking of the four target sequences with respect to the bed assignment strategies. All of them are combined with each of the four bed assignment policies, in order to verify the insensitivity of the consultation appointment strategies. In Figures 6, 7 and 8, LPT, LEPT, HIP, LEPTinv are compared against GRASP for FIFO, RB and LPTF bed assignment policies, respectively. A similar comparison is presented for AB in Figure 1.

GRASP clearly outperforms other strategies excepted for some instances with a high deferral rate when “RB” is applied. The target sequence ranking in AB bed assignment (when the global sequence is imposed) is LPT-LEPT- $\{\text{HIP}, \text{LEPTinv}\}$ from the best to the worst. The same target sequence ranking is consistently observed in all other bed assignment strategies as follows: (i) HIP and LEPTinv-based consultation appointments are by far the worst, (ii) LPT-based consultation appointment is consistently the best, (iii) LEPT is as good as LPT when the deferral probability is small but it is clearly dominated by LPT when the deferral probability is large.

8 Conclusion

In this paper we address the problem of Chemotherapy Appointment Scheduling. In this research we focus on the optimization problem at one decision point during a day where all the demands of the day are known. Uncertainties on the ability of patients to accept all treatment steps are not revealed but probabilities on such deferrals are known.

We introduced global sequences as reasonable solutions for the problem and showed that using a global rule for both consultation and injection stages does not deteriorate the quality of scheduling policies. Global sequence strategy even outperforms the current method of chemotherapy facilities where in the injection stage First-in-First-Out rule is applied.

Fig. 7 Percentage deviation from GRASP of different target sequences combined with RB policy

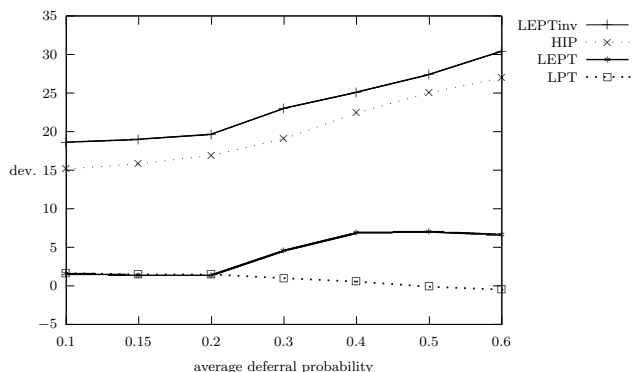
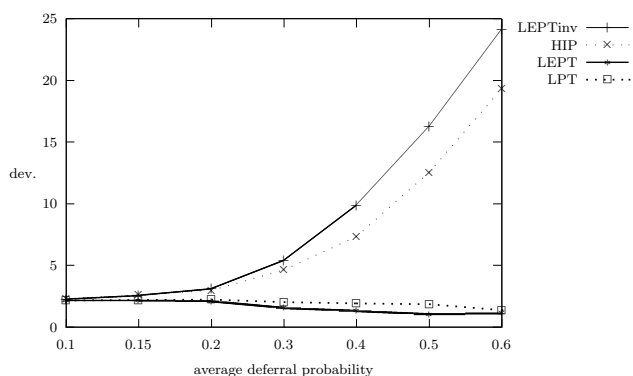


Fig. 8 Percentage deviation from GRASP of different target sequences combined with LPTF policy



We proposed the concept of “Target Sequences” and showed that based on some reasonable characteristics of the high quality solutions, it is possible to find some target sequences that perform sufficiently good for any instance of this problem.

We evaluated different targets and concluded that sorting the patients in non-increasing order of their processing time (LPT) and expected processing time (LEPT) are good target sequences. In order to evaluate and compare the results of our proposed target sequences we developed a GRASP algorithm for large instances and a linear stochastic model for small instances and showed that the two mentioned target sequences perform almost as good as optimal solutions in small instances and the best found heuristic solutions in large instances in both makespan and overworking time minimization. While both target sequences perform equally well in the instances with low deferral probability, LPT performs better in the instances with high deferral probability and high variability of injection durations and LEPT performs better in the instances with high deferral probability and low variability of injection durations.

We numerically showed that applying global sequences for both consultation and injection stages is a reasonable strategy considering that

1. its application is easier for nurses of chemotherapy facilities,
2. it makes it possible to model the problem as a linear mathematical model,
3. and it is computationally faster and easier to apply in simulation-optimization techniques.

One direct extension to this work is to consider separate morning and afternoon shifts where physicians' availability may be restricted to only one shift (see [23]). The proposed method in this paper based on global sequences and target sequences is fully compatible with this extension as it is possible to offer separate global sequences for each shift.

Many studies of the literature on appointment scheduling have been carried out with chemotherapy centers as partners. Their results show improvements against current practical scheduling strategies. An interesting perspective is to benchmark these approaches on a generic appointment scheduling problem that includes multistep treatments, uncertainty on demands and on treatment durations.

9 References

References

1. Ahmed, Z., ElMekkawy, T., Bates, S.: Developing an efficient scheduling template of a chemotherapy treatment unit: A case study. *The Australasian medical journal* **4**(10), 575 (2011)
2. Athawale, S.: Chemotherapy appointment scheduling and operations planning. Ph.D. thesis, The University of Akron (2015)
3. Berg, B.P., Denton, B.T., Erdogan, S.A., Rohleder, T., Huschka, T.: Optimal booking and scheduling in outpatient procedure centers. *Computers & Operations Research* **50**, 24–37 (2014)
4. Chen, B., Vestjens, A.P.: Scheduling on identical machines: How good is lpt in an on-line setting? *Operations Research Letters* **21**(4), 165–169 (1997)
5. Dobish, R.: Next-day chemotherapy scheduling: a multidisciplinary approach to solving workload issues in a tertiary oncology center. *Journal of Oncology Pharmacy Practice* **9**(1), 37–42 (2003)
6. Erdogan, S.A., Gose, A., Denton, B.T.: Online appointment sequencing and scheduling. *IIE Transactions* **47**(11), 1267–1286 (2015)
7. Feo, T.A., Resende, M.G.: A probabilistic heuristic for a computationally difficult set covering problem. *Operations research letters* **8**(2), 67–71 (1989)
8. Garey, M.R., Johnson, D.S., Sethi, R.: The complexity of flowshop and jobshop scheduling. *Mathematics of operations research* **1**(2), 117–129 (1976)
9. Gocgun, Y., Puterman, M.L.: Dynamic scheduling with due dates and time windows: an application to chemotherapy patient appointment booking. *Health Care Management Science* **17**(1), 60 – 76 (2014)
10. Gupta, D., Denton, B.: Appointment scheduling in health care: Challenges and opportunities. *IIE transactions* **40**(9), 800–819 (2008)
11. Hahn-Goldberg, S., Beck, J.C., Carter, M.W., Trudeau, M., Sousa, P., Beattie, K.: Solving the chemotherapy outpatient scheduling problem with constraint programming. *Journal of Applied Operational Research* **6**(3), 135–144 (2014)
12. Hahn-Goldberg, S., Carter, M.W., Beck, J.C., Trudeau, M., Sousa, P., Beattie, K.: Dynamic optimization of chemotherapy outpatient scheduling with uncertainty. *Health care management science* **17**(4), 379–392 (2014)
13. Hart, J.P., Shogan, A.W.: Semi-greedy heuristics: An empirical study. *Operations Research Letters* **6**(3), 107–114 (1987)

14. Jiang, R., Shen, S., Zhang, Y.: Integer programming approaches for appointment scheduling with random no-shows and service durations. *Operations Research* **65**(6), 1638–1656 (2017)
15. Johnson, S.M.: Optimal two-and three-stage production schedules with setup times included. *Naval research logistics quarterly* **1**(1), 61–68 (1954)
16. Lau, P.K., Watson, M.J., Hasani, A.: Patients prefer chemotherapy on the same day as their medical oncology outpatient appointment. *Journal of Oncology Practice* **10**(6), e380–e384 (2014)
17. Le, M.D., Nguyen, N., Minh, H., Baril, C., Gascon, V., Dinh, T.B.: Heuristics to solve appointment scheduling in chemotherapy. In: *Computing & Communication Technologies-Research, Innovation, and Vision for the Future (RIVF)*, 2015 IEEE RIVF International Conference on, pp. 59–64. IEEE (2015)
18. Li, K., Willis, R.: An iterative scheduling technique for resource-constrained project scheduling. *European Journal of Operational Research* **56**, 370–379 (1992)
19. Özdamar, L., Ulusoy, G.: A note on an iterative forward/backward scheduling technique with reference to a procedure by Li and Willis. *European Journal of Operational Research* **89**, 400–407 (1996)
20. Pérez, E., Ntamo, L., Malavé, C.O., Bailey, C., McCormack, P.: Stochastic online appointment scheduling of multi-step sequential procedures in nuclear medicine. *Health care management science* **16**(4), 281–299 (2013)
21. Sadki, A., Xie, X., Chauvin, F.: Patients assignment for an oncology outpatient unit. In: *Automation Science and Engineering (CASE)*, 2010 IEEE Conference on, pp. 891–896. IEEE (2010)
22. Sadki, A., Xie, X., Chauvin, F.: Appointment scheduling of oncology outpatients. In: *Automation Science and Engineering (CASE)*, 2011 IEEE Conference on, pp. 513–518. IEEE (2011)
23. Sadki, A., Xie, X., Chauvin, F.: Planning oncologists of ambulatory care units. *Decision Support Systems* **55**(2), 640–649 (2013)
24. Samorani, M., LaGanga, L.R.: Outpatient appointment scheduling given individual day-dependent no-show predictions. *European Journal of Operational Research* **240**(1), 245–257 (2015)
25. Valls, V., Ballestín, F., Quintanilla, S.: Justification and RCPSP: a technique that pays. *European Journal of Operational Research* **165**, 375–386 (2005)
26. Yokouchi, M., Aoki, S., Sang, H., Zhao, R., Takakuwa, S.: Operations analysis and appointment scheduling for an outpatient chemotherapy department. In: *Proceedings of the Winter Simulation Conference*, p. 81. Winter Simulation Conference (2012)