Health Care Management Science manuscript No.
(will be inserted by the editor)

Online Stochastic Optimization of Radiotherapy Patient Scheduling

Antoine Legrain · Marie-Andrée Fortin · Nadia Lahrichi · Louis-Martin Rousseau

Received: date / Accepted: date

Abstract The effective management of a cancer treatment facility for radiation therapy depends mainly on optimizing the use of the linear accelerators. In this project, we schedule patients on these machines taking into account their priority for treatment, the maximum waiting time before the first treatment, and the treatment duration. We collaborate with the Centre Intégré de Cancérologie de Laval to determine the best scheduling policy. Furthermore, we integrate the uncertainty related to the arrival of patients at the center. We develop a hybrid method combining stochastic optimization and online optimization to better meet the needs of central planning. We use information on the future arrivals of patients to provide an accurate picture of the expected utilization of resources. Results based on real data show that our method outperforms the policies typically used in treatment centers.

Keywords Patient scheduling · radiotherapy booking · stochastic optimization · online optimization

Mathematics Subject Classification (2000) MSC 90B36 · MSC 90C15 · MSC 90C39

Research funded in part by the Natural Sciences and Engineering Research Council of Canada and Elekta

Antoine Legrain · Nadia Lahrichi · Louis-Martin Rousseau
CIRRELT and Polytechnique Montréal, Department of Mathematical and Industrial Engineering, C.P. 6079, Succursale Centre-ville, Montréal, QC, Canada, H3C 3A7
E-mail: antoine.legrain@polymtl.ca, nadia.lahrichi@polymtl.ca, louis-martin.rousseau@polymtl.ca

Marie-Andrée Fortin
Head of the Radio-Oncology Department, CICL and Department of Radiology, Radio-Oncology and Nuclear Medicine, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, QC, H3C 3J7
E-mail: mfortin.csssl@ssss.gouv.qc.ca

1 Introduction

The Canadian Cancer Society [7] reports that "on average, one person learns that he or she has cancer every 3 minutes in Canada. Every 7 minutes, one person in Canada dies from cancer." Therefore, it is crucial to ensure that all Canadians have good access to cancer care. In 2011, Quebec authorities inaugurated two new cancer treatment facilities in the Montreal region. We collaborate with one of them: the Centre Intégré de Cancérologie de Laval (CICL).
The services provided by the center include chemotherapy and radiotherapy; the latter uses radiation to destroy malignant cells to prevent them from multiplying. Radiation treatments are administered with specialized equipment such as linear accelerators (linacs). Each patient who arrives at the CICL for radiotherapy must leave the center with an appointment for the linacs. Furthermore, the patients have different specificities that are not known in advance: their day of arrival, their priority, and their treatment duration. They also share the same resources at the CICL. This complicates the scheduling of the linacs and the choice of the first day of treatment.

The objective of this paper is to solve the radiotherapy patient scheduling problem in an online fashion. We aim to ensure that the patients receive an initial treatment within a reasonable time and that there is an optimal use of the resources, which are mainly linacs.
We present an innovative algorithm to solve the online scheduling problem. We implement a hybrid method combining stochastic optimization and online optimization. Our work builds upon the online stochastic algorithm presented by Legrain and Jaillet [21]. Our main contributions are the adaptation of this algorithm and its application to real instances. The modifications al-

low us to model more complex and realistic problems. After various computational tests to set the parameters of the online stochastic algorithm, we solve two real instances. The results outperform the solutions used by the CICL. The improvements are at least 300% as measured by our objective function.

In the next section, we present the radiotherapy scheduling problem and a literature review that places our contribution in the context of previous work on radiotherapy scheduling and online problems. In Section 3, we build an offline model and then an online procedure based on this model. In Section 4, we test our algorithm in a theoretical context and on two real instances. We conclude with some final remarks and ideas for future developments.

2 Problem statement and related work

We address operational management issues in this paper. The CICL administrators face several performance-related challenges. The provincial management closely monitors the access to specialized services, following up on individual patients on the waiting list. In Quebec, the waiting-time target for radiation oncology is set to four weeks (for 90% of the patients) and it is calculated as the time elapsed between the day when the patient is ready for treatment and the first day of treatment [24]. Many studies have shown that a delay in starting radiotherapy has a negative effect on the patient's clinical condition[9]. The impact of such a delay varies greatly depending on the nature and the severity of the cancer and the progression of the disease. The Ministry of Health and Social Services (MHSS) advises meeting certain deadlines depending on the patient's condition.

A radiation oncology department serves both curative and palliative patients, and they have different priorities. Seventy percent of the patients have a curative profile. After a complex treatment-planning process involving several healthcare professionals, the treatments are delivered 5 days a week, for a total of 5 to 44 treatments. The remaining 30% of the patients are palliative and need rapid access to care (in less than 3 days) to be relieved of serious pain.

Before beginning treatment, the patient must undergo a CT scan and a specialized team (radiation oncologist, dosimetrist, physicist, and technologist) must perform a sequence of complex tasks. The efficiency of CICL operations depends on the managers' ability to use their resources to their full potential to ensure that the patients are treated on time. Since the linacs are in great demand in a cancer treatment facility, we focus on scheduling appointments and booking slots on the accelerator. The CICL has decided to divide each linac schedule into time slots of twenty minutes, and assign an appointment to each slot. The center has a policy of treating curative patients at the same time every day. However, palliative patients can be treated at different times. Therefore, the problem involves assigning a first day of treatment and a slot on the accelerator for each patient.
The treatment requires several time slots, so the clerk must make several appointments for the patient while ensuring that the clinical and operational constraints are satisfied. He must also leave free slots in the schedule to allow for the possible arrival of more urgent patients.

The daily management of unforeseen events is important given the cost of cancer treatment. In Quebec this cost was estimated in 2008–2009 at $4400, i.e., $454 million for 103,000 patients [5]. This includes only the operational costs: workforce costs, supplies, and equipment maintenance. The random patient inflow must consequently be managed to minimize the impact on operations. To control the impact of unforeseen events, the appointment scheduling must include a stochastic aspect. Some patients undergo surgery or chemotherapy before starting radiotherapy treatment. The CICL can access information for about 80% of these patients four weeks in advance, and the stochastic optimization can take this information into account.

The appointment optimization literature is wide; [8, 14, 1] are three good reviews of this field. Researchers have drawn to many applications. Begen and Queyranne [2] tackle the single-server appointment scheduling problem where a doctor for example must give a time of appointment to a sequence of patients. Santibáñez et al [28] present the appointment scheduling of operating rooms. Patrick et al [25] study the appointment scheduling of patients with different priorities on a single server like a scanner. However, the optimization of appointments for radiation therapy is a relatively young field with a limited literature. Furthermore, CICL policy requires that patients leave the center with their appointment, so an online algorithm is necessary. This review aims to show how the radiotherapy patient scheduling problem can take advantage of online optimization techniques which have a mature literature.

2.1 Radiotherapy booking

Several simple strategies for the scheduling of a patient's radiotherapy are described in [27]. In particular, the "as soon as possible" (ASAP) strategy gives satisfactory results for this problem. The patients are sorted in ascending order of deadline and given the earliest possible appointments. Other approaches are proposed in [26]. Several ideas are tested, e.g., postponing the assignment of appointments or imposing limits on the use of resources. These ideas can easily be used by a cancer treatment facility.

Kapamara et al [18] provide a good review of exact methods and metaheuristics for the radiotherapy scheduling problem. Conforti et al [11, 12] and Jacquemin et al [15] present techniques based on linear programming. However, these methods do not take into account the arrival distribution of the patients or future events. One contribution of our paper is a novel method that takes these two features into account.

To the best of our knowledge, stochastic optimization has been used only once in radiation oncology. Sauré et al [29] introduce the radiotherapy patient scheduling problem and solve it using techniques presented in Patrick et al [25]. They take into account the different patient priorities and use a dynamic approach to deal with uncertain arrivals. They develop a method based on a Markov decision process and achieve very good results on randomly generated instances based on statistics from the British Columbia Cancer Agency. However, the Markov decision process computes an optimal policy (105 min of computing time) in advance and becomes effective when the process reaches a steady state. The CICL has advance information for some patients, and it is difficult to integrate such information into this method. Indeed, it is necessary to recalculate the optimal strategy with this information and to wait for the process to return to a steady state: information about the future has already been revealed at this point, and the recalculation remains computationally demanding. Erdogan and Denton [13] present a model for booking daily appointments in an online fashion; the instances are small (20 patients) but complex. They use stochastic optimization to deal with the uncertainties. Since the patient arrival distribution of the CICL changes dynamically for each patient arrival, stochastic and online optimization techniques are more appropriate.

2.2 Online optimization

Online optimization techniques have been developed to make a decision quickly when an event occurs while maintaining a good solution. For example, advertisement placement by search engines (the Adwords problem) is a well studied problem. For each user search, the engine displays advertisements related to the keywords of the search; the goal is to maximize income. The decisions must be made extremely quickly and should be close to optimal. The state of the art of online methods is presented in [17]; the solution time is the most important criterion. The problems studied can be modeled as allocation problems or vehicle routing problems, which are complex combinatorial problems. Online algorithms are designed to provide robust solutions regardless of the future events. Often the goal is to prove that it is impossible to obtain a solution worse than a certain bound.

There is an abundant literature on assignment problems, ranging from coupling problems [20] in a bipartite graph to more complex problems such as the Adwords problem [23]. Mehta et al [23] analyzed a greedy algorithm and more sophisticated methods for the Adwords problem. The primal-dual algorithm is applied in [6] where fifteen different problems (including Adwords) are explored. One of the advantages of this algorithm is that it can update the selection parameters during the procedure to take past decisions into account. Recently, these algorithms have incorporated probabilistic information. To make better future decisions, Manshadi et al [22] use an offline strategy based on the available information before the beginning of the algorithm. Other methods using offline statistics are presented in [19, 16]. Several papers propose algorithms that use the available information when making decisions. Ciocan and Farias [10] present a very general model that uses re-optimization techniques to incorporate this information. We use similar techniques to solve our problem of scheduling radiation therapy; these techniques are outlined in [21].

3 Methodology

The goal of the optimization is to determine the first day of treatment on a linac and then the slot. A linear program is used to schedule appointments on linacs. It must ensure the availability of sufficient resources to meet the patients' deadlines. Pretreatment (scanner, dosimetry, etc.) is performed before the radiotherapy treatment in a fixed amount of time depending on the priority of the patient.

Let $\mathcal{P}$ be the set of patients, $\mathcal{P}_c$ the set of curative patients, and $\mathcal{P}_p$ the set of palliative patients. We must schedule a set of treatments on linacs that have the following parameters for patient $j \in \mathcal{P}$:

- r_j: day when patient can start treatment (pretreatment is finished);
- d_j: deadline for first treatment;
- p_j: treatment duration in days.

We first present an offline version of the problem for two reasons. The solution is optimal and therefore a lower bound for the online problem. Furthermore, the proposed model is used for the online procedure.

3.1 The Offline Model

We solve the problem for a known set of patients. The model is a set partitioning problem: we allocate each patient to a schedule while respecting the capacity of the linacs. A schedule consists of the treatment days for a patient. The cost of a schedule for a patient takes into account the number of waiting days and the number of days beyond its deadline. In our case, there are three deadlines: 3 days for palliative patients, 14 days for a specific type of curative patients, and 28 days for all other patients. Finally, the constraint on the ready date is a hard constraint.

Let H be the index set of the working days given the planning horizon, $\mathcal{B}$ the index set of Mondays, and M the set of available linacs. Define S_j to be the index set of feasible schedules for patient j, a^m_{ijk} the plan $i \in S_j$ ($= 1$ if the patient is treated on machine m on day k, and 0 otherwise), and c_{ij} the cost of the schedule $i \in S_j$. Let F^m_k be the number of available slots on linac m on day k, O_{day} the maximum number of overtime slots on linac m on day k, O_{week} the maximum number of overtime slots on linac m for a week, and c^o the cost of one overtime slot.

The variable x_{ij} defines the allocation of the plan $i \in S_j$ to patient j (=1 if allocated, and 0 otherwise). z_{mk} is the number of overtime slots on linac m on day k. This gives the following formulation for a known set of patients:

$$\min \sum_{j\in\mathcal{P}} \sum_{i\in S_j} c_{ij}x_{ij} + \sum_{k\in H} \sum_{m\in M} c^o z_{mk} \tag{1a}$$

subject to:

$$\sum_{i\in S_j} x_{ij} = 1, \qquad \forall j \in \mathcal{P} \tag{1b}$$

$$\sum_{j\in\mathcal{P}} \sum_{i\in S_j} a^m_{ijk} x_{ij} \le F^m_k + z_{mk}, \qquad \forall m \in M, \forall k \in H \tag{1c}$$

$$\sum_{j\in\mathcal{P}_p} \sum_{i\in S_j} a^m_{ijk} x_{ij} \ge z_{mk}, \qquad \forall m \in M, \forall k \in H \tag{1d}$$

$$\sum_{k=b}^{b+4} z_{mk} \le O_{week}, \qquad \forall m \in M, \forall b \in \mathcal{B} \tag{1e}$$

$$z_{mk} \in [0, O_{day}], \qquad \forall m \in M, \forall k \in H \tag{1f}$$

$$x_{ij} \in \{0, 1\}, \qquad \forall j \in \mathcal{P}, \forall i \in S_j \tag{1g}$$

Constraint (1b) ensures that each patient j is scheduled. Constraint (1c) ensures that the capacity (including the overtime) of each linac is not exceeded. Constraint (1d) prevents curative patients from being scheduled in overtime slots. Constraints (1e) and (1f) bound respectively the weekly and daily overtime on each linac. Constraint (1g) is an integrality constraint. Finally, the objective (1a) is divided into two parts: the cost of a plan and the cost of overtime.

If all the plans are considered at the beginning, the computational time is too high. Therefore, the sets S_j of schedules are dynamically updated by column generation according to Algorithm 1.

Algorithm 1 Offline Algorithm - Step 1

while there is a schedule with a negative reduced cost **do**
  solve linear relaxation of formulation (1)
  enumerate all feasible schedules for each patient j (start of treatment must be after day r_j and before day $d_j + \delta$)
  compute reduced costs of schedules from the dual variables
  insert schedules of patient j with a negative reduced cost into S_j, $\forall j \in \mathcal{P}$
end while
Solve formulation (1)

The parameter δ is a constant and allows us to miss the deadline by a few days. This value ensures the feasibility of formulation (1) and should be small (e.g., less than a week). Algorithm 1 gives a nearly optimal solution: some schedules not generated by the procedure could improve the integer solutions. The plan i chosen for each patient j determines only the first treatment day k and the linac m where the patient is treated. A slot must then be booked for each patient. The following theorem shows that a feasible solution always exists.

Theorem: For each feasible solution of formulation (1), there exists at least one feasible solution of the original problem. This solution can be built by Algorithm 2.

Algorithm 2 Offline Algorithm - Step 2

Sort all curative patients ($\mathcal{P}_c$) in ascending order of first day of treatment: let $\mathcal{R}$ be the resulting set;
while $\mathcal{R}$ is not empty **do**
  consider the first patient $j \in \mathcal{R}$
  choose the first free slot p on the first treatment day k on the chosen linac m
  choose slot p on linac m for all treatment days
  remove patient j from $\mathcal{R}$
end while
Book palliative patients in the remaining free slots

This theorem also justifies the two-step approach: it shows that solving the overall problem is equivalent to solving the two problems sequentially.

Proof: First, we will prove by induction on $\mathcal{R}$ that the plan for the curative patients is feasible.

$\#\mathcal{R} = 0$: There are no patients to schedule: the plan is feasible.
$\#\mathcal{R} = n + 1$: Suppose that the plan built with the constructive procedure is feasible for the first n patients. Let j be the last patient of $\mathcal{R}$. There exists a free slot p on the first treatment day k on linac m for patient j, because $\sum_{l\in\mathcal{P}_c}\sum_{i\in S_l} a^m_{ilk}x_{il} \le F^m_k$, $\forall m \in M, \forall k \in H$. This inequality holds because $\sum_{l\in\mathcal{P}}\sum_{i\in S_l} a^m_{ilk}x_{il} \le F^m_k + z_{mk}$ and $\sum_{l\in\mathcal{P}_p}\sum_{i\in S_l} a^m_{ilk}x_{il} \ge z_{mk}$. Slot p is free for all the subsequent days, because if a slot is not booked on day k, it will not be booked subsequently. A treatment occurs in the same slot on every day, and no patient currently booked starts treatment after day k. We therefore assign slot p on linac m to patient j. This plan is feasible.
By induction, the plan for all the curative patients is feasible.
Since there are enough free slots for all the palliative patients, the overall plan is also feasible. Consequently, if formulation (1) is solved to optimality, the overall solution will also be optimal.

3.2 The Online Procedure

Now we solve the model in an online fashion, i.e., the patients arrive sequentially and future patients are not known in advance. We present different algorithms for this problem.

3.2.1 Online algorithms

Two algorithms have often been used for online assignment problems: the greedy algorithm and the primal-dual algorithm. The first chooses the feasible plan with the lowest cost. We will use a slightly different algorithm: a minimum of b slots per linac per day are reserved for palliative patients. This number is set by the CICL. This policy ensures that sufficient slots are reserved for palliative patients while the remaining slots are allocated to curative patients. Indeed, at the CICL curative patients are scheduled at least one week before the beginning of their treatment while palliative patients are planned only a few days before. Algorithm 3 presents this modified greedy procedure. It is also the algorithm used by the CICL.

Algorithm 3 finds the cheapest feasible plan in a greedy fashion, subject to a constraint on the minimal number of slots reserved for palliative patients. This procedure is also close to the ASAP procedure. The longer the patient must wait, the more expensive the plan. Therefore, this algorithm will generally assign the first day of treatment to occur as soon as possible.

Algorithm 3 CICL Algorithm

```
for all patient arrivals j do
  if patient j is curative then
    choose first treament day k on linac m such that:
    1) the cost of the associated plan is minimal
    2) there are fewer than F_h^m − b slots booked on all treatment days
  else
    book palliative patient with the cheapest feasible plan
  end if
end for
```

We use a primal-dual algorithm that draws inspiration from the examples of [6]. The algorithm makes an irrevocable decision at each patient arrival. The objective function is a simple maximize (or minimize) objective of the form $\sum_i (c_{ij} - r_i)x_{ij}$, where c_{ij} is a cost, r_i a constant dual variable, and x_{ij} the decision variable. Once the algorithm has made a choice, the dual variables are updated. Different primal-dual algorithms use different approaches to update the dual variables. In our algorithm, these variables are estimated at each patient arrival and are calculated using stochastic optimization tools.

3.2.2 Stochastic primal-dual algorithm

The stochastic optimization model is based on the offline formulation (1) and a sample set Ω_j of future scenarios. After the arrival of patient j, we consider the scenarios for future patients who arrive before day $d_j + p_j$. Let $\mathcal{P}^\omega$ be the set of these future patients for a scenario $\omega \in \Omega_j$. F^m_k is from now on the remaining number of free slots (patients who arrived before j are already scheduled). O_{week} and O_{day} have been modified to take into account the palliative patients already booked in overtime. This gives the following formulation derived from formulation (1) to schedule patient j:

$$\min \sum_{i \in S_j} c_{ij} x_{ij} + \mathbb{E}_{\omega \in \Omega_j} [\sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} c_{il} y^\omega_{il} + \sum_{k \in H} \sum_{m \in M} c^o z^\omega_{mk}]$$

subject to:

$$\begin{aligned}
&\sum_{i \in S_j} x_{ij} = 1 \\
&\sum_{i \in S_l} y^\omega_{il} = 1, \qquad \forall \omega \in \Omega_j, \forall l \in \mathcal{P}^\omega \\
&\sum_{i \in S_j} a^m_{ijk} x_{ij} + \sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} a^m_{ilk} y^\omega_{il} \le F^m_k + z^\omega_{mk}, \\
&\qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j \\
&\mathbb{1}_{\mathcal{P}_p}(j) \sum_{i \in S_j} a^m_{ijk} x_{ij} + \sum_{l \in \mathcal{P}^\omega_p} \sum_{i \in S_l} a^m_{ilk} y^\omega_{il} \ge z^\omega_{mk}, \\
&\qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j \\
&\sum_{k=b}^{b+4} z^\omega_{mk} \le O_{week}, \qquad \forall m \in M, \forall b \in \mathcal{B}, \forall \omega \in \Omega_j \\
&z^\omega_{mk} \in [0, O_{day}], \qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j \\
&x_{ij} \in \{0,1\}, \qquad \forall i \in S_j \\
&y^\omega_{il} \in \{0,1\}, \qquad \forall l \in \mathcal{P}^\omega, \forall i \in S_l, \forall \omega \in \Omega_j
\end{aligned} \tag{2}$$

For the details of the formula of the variables u_{ij}, see Appendix A. To summarize, our current method can be stated as follows:

Algorithm 4 Stochastic Primal-Dual Algorithm

for all patient arrivals j **do**
  for all scenarios ω **do**
    solve dual slave problem (5) with value $x = 0$ for event ω
  end for
  compute the variables u_{ij} (as in Appendix A)
  select a feasible schedule i that minimizes the cost $(c_{ij} + u_{ij})$
  update remaining resources F^m_k
end for

The reduced cost has two parts: c_{ij} is the current cost of the pattern, and u_{ij} is the expected cost of this pattern in the future. Algorithm 4 enumerates every feasible schedule to find the one that minimizes this reduced cost.

3.2.3 Online stochastic algorithm

Algorithm 4 chooses only the first day of treatment. We use another procedure to choose the slot, as in the offline case. When palliative patients arrive at the center, some curative patients have already been booked for the next seven days. Palliative patients simply fill the remaining slots of the plan and no optimization is necessary. Algorithm 4 is used only for curative patients and so $\mathbb{1}_{\mathcal{P}_p}(j) = 0$.
For a curative patient starting treatment on linac m on day k, we introduce the measure $G_{mk}(p)$ for a slot p. $G_{mk}(p)$ is equal to the number of consecutive free slots p just before day k on linac m. We choose the slot p with the minimum $G_{mk}(p)$. If a tie occurs, we choose the earliest slot in the day. We obtain Algorithm 5.

Algorithm 5 Online Stochastic Algorithm

for all patient arrivals j **do**
  if patient j is curative **then**
    choose first treatment day k on linac m using Algorithm 4
    choose treatment slot p such that:
    $G_{mk}(p)$ is minimal and (if a tie occurs) p is minimal
  else
    book patient ASAP
  end if
end for

Algorithm 5 has three main benefits:

1. The remaining free slots for palliative patients are not constant but optimized according to the current and future workload;
2. The ASAP procedure allows us to give palliative patients the earliest possible appointments;
3. The procedure that chooses slot p tries to keep the largest slots for future patients.

4 Results

We will first present the context of the results. Then, we will study the algorithms in a theoretical context to set some parameters. Finally, we will solve two real instances from the CICL.

4.1 Context of the results

We consider three types (t) of patients, each defined by its urgency, i.e., the deadline d_j for the start of treatment. Most urgent patients are palliative and must be treated within 3 days; the lower-priority patients are curative and must be treated within 14 or 28 days. The center's main objective is to ensure that the patients meet their treatment deadlines. Our objective function has a penalty if a deadline is missed. Each penalty is modeled by a piecewise function: this is the most important part of the objective. Furthermore, the patient waiting time is slightly penalized to break the symmetry of the solution. Finally, the overtime is penalized to avoid unnecessary overtime. This penalty is difficult to set, because it depends on CICL policy, which tries to balance overtime and delayed treatments. We have set this penalty in such a way that overtime slots are booked only if a palliative patient misses his deadline by more than one day.

We compare three decision rules for scheduling patients: the CICL method (Algorithm 3), our online stochastic strategy (Algorithm 5) and the offline strategy (Algorithms 1 and 2). The CICL method reserves two free slots on each linac to accommodate possible future palliative patients (three-day deadline). The online stochastic algorithm uses scenarios that take into account all the unforeseen events that we want to model. The uncertainties that we consider are the number of patients, their arrival dates, their priorities, and their treatment durations. The offline strategy is almost optimal and represents the situation where all information is known in advance. Indeed, if every pattern is generated at the beginning, the optimality gaps are around 0.1%.

4.2 Sensitivity analysis

The scenarios are generated from a modified Poisson distribution. The number of each patient type arriving in a day follows a Poisson distribution truncated at a ceiling of twice the mean. For each patient type, the cancer type follows a Bernoulli distribution. The Poisson distributions are truncated to reduce the variability in the total number of patients. The parameters used were inferred from CICL statistics. The Poisson distribution is described by a total rate (λ) and the proportion (ρ_t) given in Table 1. Thus, for each patient type (t), the mean of the Poisson distribution is equal to $\rho_t\lambda$. The parameter ρ_t was inferred from CICL statistics. However, the arrival rate was not inferred, because it depends of the number of linacs. As the CICL is a new cancer treatment center, its number of linacs grows slowly from 2 up to 6. To obtain reasonable results it is important to use a moderate arrival rate λ that is representative of the number of linacs. We will infer this arrival rate by machine learning techniques in the future.

We allow 29 slots per linac per day and 3 overtime slots per linac per day. We allow a maximum of 5 overtime slots per linac per week. The online stochastic algorithm has been tested in two situations. The first uses the probability distribution: we refer to this as the online stochastic algorithm. The second uses both the probability distribution and all available information on low-priority patients, i.e., the patients with a 28-day deadline are known in advance: we refer to this as the online clairvoyant algorithm. We have assumed that all low-priority patients are known although, in reality, only 80% of them are registered at the CICL in advance.

4.2.1 Number of scenarios

The set Ω_j is normally very large, so we represent it by a few generated scenarios. We have to find a trade-off between the number of scenarios, the solution quality, and the computational time. We have tested different numbers of scenarios ($|\Omega_j|$) to study the behavior of the algorithms. We compare the value of the offline objective, the competitive ratio of the CICL algorithm, the competitive ratio of the online stochastic algorithm, and the ratio of the objectives of the CICL algorithm and the online stochastic algorithm. These are respectively:
$Z^* = Obj_{offline}$, $c_{cicl} = \dfrac{Obj_{cicl}}{Obj_{offline}}$, $c_{online} = \dfrac{Obj_{online\ stochastic}}{Obj_{offline}}$, $c_{relative} = \dfrac{Obj_{cicl}}{Obj_{online\ stochastic}}$.
The same ratios are defined for the online clairvoyant algorithm. We ran the tests 200 times and found the averages of these different measures. The planning horizon is 100 days, and the average number of patients is 168.5. We considered one linac with a moderate arrival rate λ of 1.7. We allowed a maximum of 5 s of solution time for each curative patient; the CICL administrators consider this reasonable. All the tests were conducted with CPLEX 12.5 on an Intel(R) Xeon(R) X5675 (3.07 GHz) CPU.

The results presented in Table 2 are for the online stochastic algorithm and the online clairvoyant algorithm. For the former procedure, the use of available information improves the performance. As the number of scenarios increases, c_{online} decreases, stabilizing around 2.7. At the same time, $c_{relative}$ increases, stabilizing around 1.17. Note that these two ratios have apparently opposite behavior.

The online clairvoyant algorithm gives better results than the online stochastic procedure. The ratios follow the same trend as the number of scenarios increases: c_{online} decreases to 2.10 and $c_{relative}$ increases to 1.53. The use of the information greatly decreases the variability of the scenarios and has an immediate impact on the results.

Furthermore, the more scenarios the algorithms have, the better the ratios. However, more than 15 sce-

Patient type t	palliative	curative 1 (14-day deadline)	curative 2 (28-day deadline)
Proportion ρ_t (%)	31	19	50

Table 1 Proportions of patient types inferred from CICL data

$\|\Omega_j\| =$	1	2	3	5	10	15	20	25	30
Z^*	807776								
c_{cicl}	3.16								
Online Stochastic Algorithm									
Average c_{online}	3.17	2.92	2.80	2.79	2.70	2.73	2.72	2.78	2.71
Average $c_{relative}$	1.05	1.11	1.15	1.15	1.18	1.17	1.17	1.16	1.17
Average number of scenarios	1.0	2.0	3.0	5.0	10.0	14.8	19.1	22.8	25.9
Online Clairvoyant Algorithm									
Average c_{online}	2.32	2.24	2.17	2.16	2.13	2.12	2.10	2.10	2.10
Average $c_{relative}$	1.36	1.41	1.48	1.48	1.51	1.52	1.52	1.52	1.53
Average number of scenarios	1.0	2.0	3.0	5.0	10.0	15.0	19.8	24.4	28.8

Table 2 Comparison of measures for different values of $|\Omega_j|$

narios does not noticeably improve the results; 10 scenarios seem to suffice.

Finally, when there are more than 20 scenarios, on average the online stochastic algorithm does not have enough time to solve all the scenarios. This is not the case for the clairvoyant algorithm. The problems of the online stochastic algorithm are harder to solve because the scenarios for a particular patient are more varied.

To summarize, the online stochastic and clairvoyant algorithms should generate 15 scenarios for the curative patients to achieve very good results. If we have good information about the future, the online clairvoyant algorithm can take advantage of it and perform much better than the CICL algorithm.
Since the CICL knows about 80% of the future low-priority curative patients, we will use the online clairvoyant algorithm in the following sections. Finally, if there is a tight constraint on the solution time, the algorithms will explore 15 scenarios; otherwise they will use their 5 seconds for each curative patient.

4.2.2 Evolution of the algorithms

We study in this section the evolution of the CICL algorithm and the online clairvoyant procedure. We calculate the instantaneous rate of utilization by computing the average rate of utilization for a period of thirty days from the current day. This allows us to see the future workload. We want to measure the impact of a high or low rate on the behavior of the algorithms. We also study the evolution of the cumulative number of delayed patients per day. This cumulative function increases by one on a given day if a patient who has arrived on that day is scheduled after his deadline. The number of delayed patients increases quickly when the instantaneous rate is high. We also want to explore whether the online clairvoyant algorithm stabilizes the number of delayed patients and their waiting time compared to the CICL method. This could explain why the online clairvoyant algorithm outperforms the CICL procedure.
The planning horizon is 300 days, and there are 408 patients. The instance has been solved only once for one linac with an arrival rate λ of 1.5, and 15 scenarios have been used to find a schedule for each curative patient.

Until day 70 in Figure 1, there are no delayed patients because there are enough free slots on the linac to schedule all the patients before their deadlines. The instantaneous rate of utilization then starts increasing and goes up to 0.7. The number of delayed patients increases slightly for the CICL algorithm. The instantaneous rate decreases slightly and there are no delayed patients for either algorithm. The instantaneous rate then increases again. After this second rise, the number of delayed patients increases until the end of the time horizon: there are not enough slots to book new patients. The online clairvoyant algorithm increases in steps and not as quickly, and its rate of utilization has less variability. This algorithm obtains better results, because it stabilizes this rate and, consequently, it reserves a stable number of free slots for future patients.

There is a final difficulty for palliative patients. If a patient arrives on a Friday and is not treated the same day, the first day of treatment is the subsequent Monday and the patient has already waited three days. Consequently, the patient has missed his deadline while waiting only one working day. Table 3 presents the distribution of delayed palliative patients by arrival day.

The CICL algorithm has scheduled more palliative patients after their deadlines. This is normal, because there are many delayed patients for this instance. However, these palliative patients are uniformly distributed across the week. For the clairvoyant algorithm, most of the delayed palliative patients arrive at the end of the

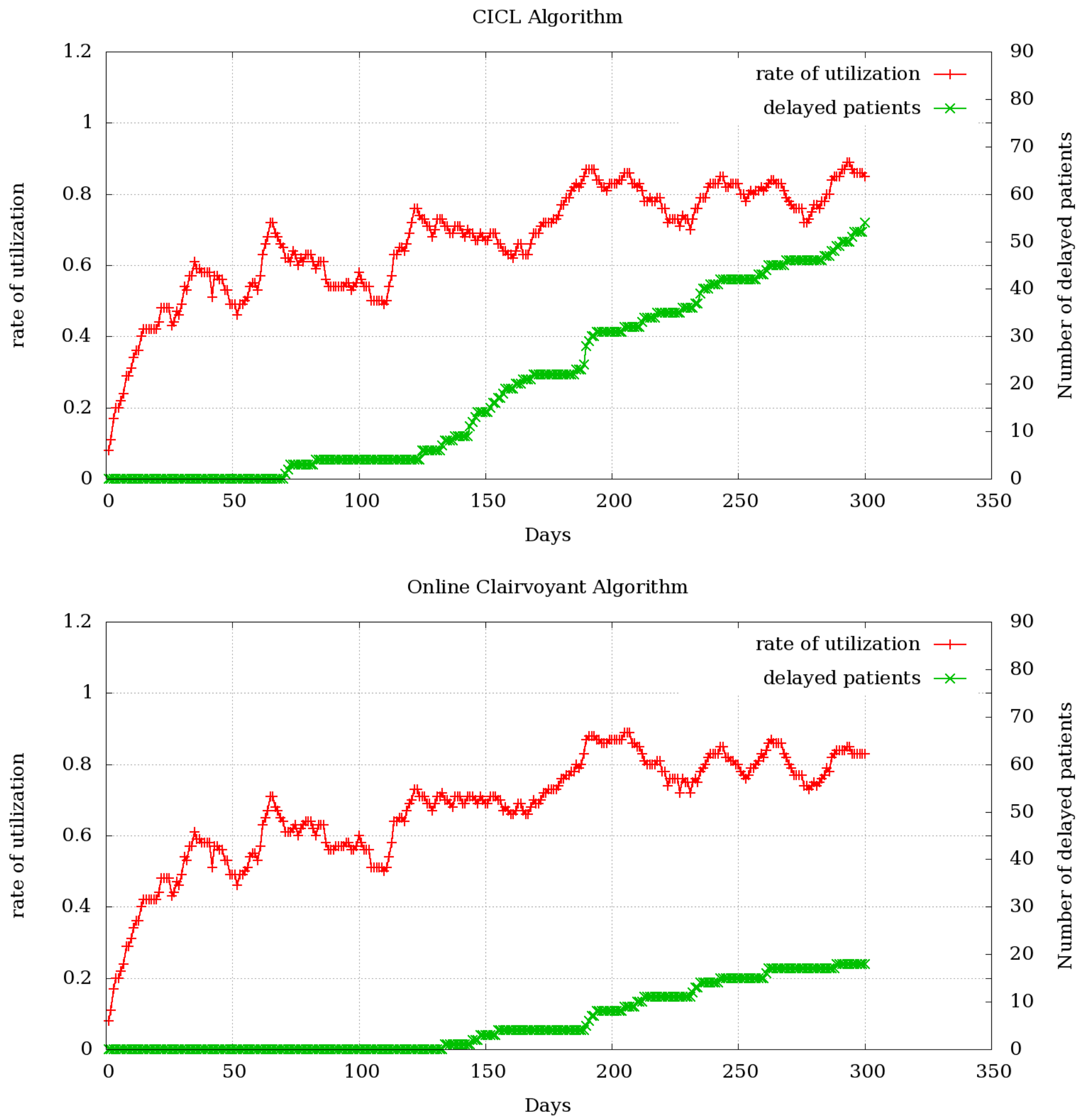

Fig. 1 Evolution of the number of delayed patients and the instantaneous rate of utilization

	Arrival Day				
Algorithm	Monday	Tuesday	Wednesday	Thursday	Friday
CICL	4	6	8	6	8
Online Clairvoyant	0	1	2	2	5

Table 3 Distribution of delayed palliative patients by arrival day

week, and half of them arrive on Friday. If the weekend days were not counted toward the delay, this algorithm would have better results.

4.2.3 Objective function criteria

We present one instance with the value of the objective function as well as the number of delayed patients, the average delay, and the number of overtime slots. The aim is to check whether or not the objective function correctly measures the quality of the solution in a perfect situation, where the probability distribution is perfectly known. The planning horizon is 100 days and only one linac is used, as in Section 4.2.1.

Table 4 shows the results for the three strategies. The first set of columns gives for each strategy the number of patients who miss their deadlines, i.e., the waiting time is more than 3, 14, or 28 days. The next set of columns specifies the average waiting time for treatment, once again for each patient type. The final column shows the number of overtime slots. This instance has 140 patients, and the arrival rate λ is still 1.7. The online clairvoyant algorithm uses 15 scenarios for each curative patient. We obtain the following results: $Z^* = 494.5$, $c_{cicl} = 3.47$, $c_{online} = 1.33$, and $c_{relative} = 2.60$.

In terms of solution quality, the results presented in Table 4 are satisfactory. They show that the online clairvoyant procedure meets the needs of booking policies better than the strategy typically used in cancer centers: 5 palliative patients (3.6%) and 7 curative patients (5.0%) with a 14-day deadline are better served.

In terms of waiting times, patients with a 14-day deadline have an average delay of 8.72 days, compared to 11.31 days with the CICL strategy. However, those with a 28-day deadline have an average delay of 13.63 days. The distribution of these delays is better aligned with the original deadlines, i.e., those with a later deadline tend to wait more. The average deadlines obtained by the online clairvoyant and the offline algorithms compare very well: the two solutions have roughly the same average waiting time. It is slightly higher for the online clairvoyant procedure; it is because this algorithm tries to build in some flexibility for future patients whereas the offline algorithm knows the patients in advance.

To summarize, the offline solution is almost optimal and all the deadlines are met in this example. The online clairvoyant approach compares very well to the offline solution: no patients miss their deadlines, and the average waiting time is comparable to that of the offline solution. The CICL approach is clearly the simplest since it involves no computation, but 12 patients miss their deadlines, and the average waiting time is no better than that of the online clairvoyant algorithm.

All the measures also indicate that the offline algorithm provides the best solution: the ratios c_{cicl} and c_{online} are greater than 1. This solution is optimal and can only be achieved in a perfect world where the future patients are completely known. Furthermore, the ratio c_{online} is close to 1, indicating that the online solution is close to the offline solution, despite its imperfect knowledge of the future. Finally, we note that the ratios are better than their averages (presented in Table 2), but the differences between the CICL method and the online clairvoyant algorithm are large. To conclude, the ratios c_{cicl} and $c_{relative}$ confirm that the CICL solution is the worst in this case.

4.3 Real instances

This section presents two real CICL instances with 74 and 77 working days. We ran the tests with 2 linacs, 23 slots per day for the first instance and 29 for the second, and 3 overtime slots per linac per day for both. We allowed a maximum of 5 overtime slots per linac per week. The first instance books 159 patients and the second 181. We generated future patients with an arrival rate λ of 2.75 for the first instance and 3.5 for the second. The tests were run 200 times for each instance with 15 scenarios and a maximum of 5 s of solution time, but only one run is presented in Tables 5 and 6.

The solution presented in Table 5 has the following values: $c_{online} = 3.07$ (for a mean of 2.82 and a standard deviation of 0.17), $c_{relative} = 11.91$ (for a mean of 13.33 and a standard deviation of 0.98), and 14.3 scenarios on average. This performance is slightly worse than the average. However, the solution obtained for the online clairvoyant algorithm outperforms the CICL solution. In particular, the CICL solution delays 16 curative patients including 11 high-priority curative patients who have waited more than 21 days. However, the low-priority curative patients wait in average two more days for the online clairvoyant solution. The improvement of the global solution penalizes these patients, but they are still treated before their deadline. Finally, the offline solution delays no patients and the high-priority palliative and curative patients are nearly served at their ready date.

The solution presented in Table 6 has the following values: $c_{online} = 2.42$ (for a mean of 1.74 and a standard deviation of 0.13), $c_{relative} = 4.08$ (for a mean of 5.76 and a standard deviation of 0.95), and 55.4 scenarios on average. These ratios are still worse than the average, and the online clairvoyant algorithm continues to perform better than the CICL procedure. The online clairvoyant algorithm keeps delaying low-priority curative patients while high-priority patients have a better access to the center. Finally, the online clairvoyant and offline results are similar for this instance.

Conclusion

We have presented the problem of scheduling patient appointments in an online fashion. The goal is to offer patients a reasonable waiting time for their first treatment, while maximizing the resource utilization, and the ultimate goal is to provide better access to treatment. Curative patients are offered regular appointments, i.e., at the same time every treatment day, to accommodate their other commitments. Palliative patients are offered varying appointments. Booking cura-

	Deadline			Average delay			Overtime slots
Algorithm	>3 days	>14 days	>28 days	3 days	14 days	28 days	
CICL	5	7	0	0.62	11.31	10.66	1
Online Clairvoyant	0	0	0	0.11	8.72	13.63	0
Offline	0	0	0	0.0	7.31	10.08	0

Table 4 Results for a randomly generated instance

	Deadline			Average delay			Overtime slots
Algorithm	>3 days	>14 days	>28 days	3 days	14 days	28 days	
CICL	6	16 (11)	0	1.11	16.32	14.23	1
Online Clairvoyant	4	6	0	0.57	10.24	16.03	2
Offline	0	0	0	0.0	7.12	14.42	0

Table 5 159 patients in 74 working days

	Deadline			Average delay			Overtime slots
Algorithm	>3 days	>14 days	>28 days	3 days	14 days	28 days	
CICL	15	4	0	1.78	9.19	8.65	13
Online Clairvoyant	5	0	0	0.78	8.54	9.98	0
Offline	0	0	0	0.0	7.13	10.14	0

Table 6 181 patients in 77 working days

tive patients has a clear impact on the utilization of the machines since they might not be available for future palliative patients. An offline two-step algorithm provides an optimal solution when all patients are known.

To reflect the uncertainty related to the arrival of patients and the treatment duration, we developed an approach combining stochastic optimization and online optimization. This approach is innovative and our contributions are in both the methodology and the application. The results show that this approach works well on real instances and outperforms the current strategy.

In future work we will schedule the upstream flow. Our current formulation assumes that the pretreatment takes a constant numbers of days whatever the workload. We want to estimate this number of days more accurately to avoid delays. If the pretreatment is not completed on time, the patient cannot start radiotherapy and the corresponding slots should be canceled. Instead of enumerating all feasible plans, we will solve a more complex problem that takes into account the pretreatment and continues to minimize the reduced costs. This research will lead to a way to solve more complex stochastic column generation problems.

Furthermore, as the CICL and its head of the radio-oncology department pilot the project, the algorithm is suited to the CICL operations. The software provider and the CICL have also a collaborative agreement for the integration of the algorithm to the software.

Acknowledgements We would like to thanks Dr Bruno Carrozzo and Ms Julie Heon for their great help to understand CICL operations and to obtain their data.

A The online stochastic algorithm

We present here the online stochastic algorithm introduced by Legrain and Jaillet [21]. We rewrite this algorithm for our application. Recall the stochastic optimization formulation (2):

$$\min \sum_{i \in S_j} c_{ij} x_{ij} + \mathbb{E}_{\omega \in \Omega_j} [\sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} c_{il} y^\omega_{il} + \sum_{k \in H} \sum_{m \in M} c^o z^\omega_{mk}]$$

subject to:

$$\sum_{i \in S_j} x_{ij} = 1$$

$$\sum_{i \in S_l} y^\omega_{il} = 1, \qquad \forall \omega \in \Omega_j, \forall l \in \mathcal{P}^\omega$$

$$\sum_{i \in S_j} a^m_{ijk} x_{ij} + \sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} a^m_{ilk} y^\omega_{il} \leq F^m_k + z^\omega_{mk}, \qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j$$

$$\mathbb{1}_{\mathcal{P}_p}(j) \sum_{i \in S_j} a^m_{ijk} x_{ij} + \sum_{l \in \mathcal{P}^\omega_p} \sum_{i \in S_l} a^m_{ilk} y^\omega_{il} \geq z^\omega_{mk}, \qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j$$

$$\sum_{k=b}^{b+4} z^\omega_{mk} \leq O_{week} \qquad \forall m \in M, \forall b \in \mathcal{B}, \forall \omega \in \Omega_j$$

$$z^\omega_{mk} \in [0, O_{day}], \qquad \forall m \in M, \forall k \in H, \forall \omega \in \Omega_j$$

$$x_{ij} \in \{0,1\}, \qquad \forall i \in S_j$$

$$y^\omega_{il} \in \{0,1\}, \qquad \forall l \in \mathcal{P}^\omega, \forall i \in S_l, \forall \omega \in \Omega_j$$

A.1 L-shaped method

The L-Shaped method [4] enables us to solve formulation (2). This technique is based on Benders decomposition [3]. Formulation (2) is first decomposed into a master problem (3) and slave problems (4).

$$\text{(Master problem)} \quad \min \sum_{i \in S_j} c_{ij} x_{ij} + \mathbb{E}_{\omega \in \Omega_j}[Q(x,\omega)]$$

subject to:

$$\begin{aligned}
&\sum_{i \in S_j} x_{ij} = 1 \\
&x_{ij} \in \{0,1\}, \qquad \forall i \in S_j
\end{aligned} \tag{3}$$

The function $Q(x, \omega)$ is called the recourse; it deals with the stochastic part of the objective. The goal of the slave problems (4) is to compute the value of this function for each x and each ω. The variables inside the parentheses (.) are the dual variables associated with the constraint.

(Slave problems)

$$Q(x,\omega) \;=\; \min \sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} c_{il} y^\omega_{il} + \sum_{k \in H} \sum_{m \in M} c^o z^\omega_{mk}$$

subject to:

$$\begin{aligned}
&\sum_{i \in S_l} y^\omega_{il} = 1, \qquad \forall l \in \mathcal{P}^\omega \;(\alpha^\omega_l) \\
&\sum_{l \in \mathcal{P}^\omega} \sum_{i \in S_l} a^m_{ilk} y^\omega_{il} \le F^m_k + z^\omega_{mk} - \sum_{i \in S_j} a^m_{ijk} x_{ij}, \\
&\qquad\qquad\qquad \forall m \in M, \forall k \in H \;(\beta^\omega_{mk}) \\
&\sum_{l \in \mathcal{P}^\omega_p} \sum_{i \in S_l} a^m_{ilk} y^\omega_{il} \ge z^\omega_{mk} - \mathbb{1}_{\mathcal{P}_p}(j) \sum_{i \in S_j} a^m_{ijk} x_{ij}, \\
&\qquad\qquad\qquad \forall m \in M, \forall k \in H \;(\gamma^\omega_{mk}) \\
&\sum_{k=b}^{b+4} z^\omega_{mk} \le O_{week}, \qquad \forall m \in M, \forall b \in \mathcal{B} \;(\pi^{1\omega}_{mb}) \\
&z^\omega_{mk} \in [0, O_{day}], \qquad \forall m \in M, \forall k \in H \;(\pi^{2\omega}_{mk}) \\
&y^\omega_{il} \in \{0,1\}, \qquad \forall l \in \mathcal{P}^\omega, \forall i \in S_l \;(\pi^{3\omega}_{mk})
\end{aligned} \tag{4}$$

Since the recourse function Q has to be computed for each value of its variables, an approximation of Q is built. First, we relax the integrality constraint on the variables y^ω_{il} and then we use the duals (5) of the slave problems.

(Dual slave problems)

$$\begin{aligned}
&\max \sum_{l \in \mathcal{P}^\omega} \alpha^\omega_l - \sum_{m \in M} \{ \sum_{k \in H} [(F^m_k - \sum_{i \in S_j} a^m_{ijk} x_{ij}) \beta^\omega_{mk} + \\
&\mathbb{1}_{\mathcal{P}_p}(j) (\sum_{i \in S_j} a^m_{ijk} x_{ij}) \gamma^\omega_{mk} + O_{day} \pi^{2\omega}_{mk} + \pi^{3\omega}_{mk}] + \sum_{b \in \mathcal{B}} O_{week} \pi^{1\omega}_{mb} \}
\end{aligned}$$

subject to:

$$\begin{aligned}
&\alpha^\omega_l + \sum_{m \in M} \sum_{k \in H} a^m_{ilk} (\gamma^\omega_{mk} - \beta^\omega_{mk}) - \pi^{3\omega}_{mk} \ge c_{il}, \\
&\qquad\qquad\qquad \forall l \in \mathcal{P}^\omega, \forall i \in S_l \\
&\beta^\omega_{mk} - \gamma^\omega_{mk} - \pi^{1\omega}_{mb(k)} - \pi^{2\omega}_{mk} \ge c^o, \qquad \forall m \in M, \forall k \in H \\
&\beta^\omega_{mk},\; \gamma^\omega_{mk},\; \pi^{2\omega}_{mk},\; \pi^{3\omega}_{mk} \ge 0, \qquad \forall m \in M, \forall k \in H \\
&\pi^{1\omega}_{mb} \ge 0, \qquad \forall m \in M, \forall b \in \mathcal{B}
\end{aligned} \tag{5}$$

where $b(k)$ is the index of the Monday of the same week as the working day indexed by k. These problems are always feasible; furthermore they are bounded when there are enough free slots on the linacs for the chosen pattern i ($x_{ij} = 1$). There always exists such a pattern, because the first day of treatment can be sufficiently far from the current day to ensure that there are enough free slots.
Thus, the weak duality theorem gives an approximation (a cut) of the recourse function Q:

$$\begin{aligned}
&\forall i \in S_j, \forall \omega \in \Omega_j, \\
&Q(x,\omega) \ge \sum_{l \in \mathcal{P}^\omega} \alpha^\omega_l - \sum_{m \in M} \{ \sum_{k \in H} [F^m_k \beta^\omega_{mk} + \\
&\qquad \sum_{i \in S_j} a^m_{ijk} x_{ij} (-\beta^\omega_{mk} + \mathbb{1}_{\mathcal{P}_p}(j) \gamma^\omega_{mk}) + O_{day} \pi^{2\omega}_{mk} + \pi^{3\omega}_{mk}] + \\
&\qquad \sum_{b \in \mathcal{B}} O_{week} \pi^{1\omega}_{mb} \}
\end{aligned} \tag{6}$$

For the qth cut for scenario ω, approximation (6) can be simplified. All the constants of this cut can be gathered into one constant C^ω_q:

$$C^\omega_q = \sum_{l \in \mathcal{P}^\omega} \alpha^\omega_l - \sum_{m \in M} \{ \sum_{k \in H} [F^m_k \beta^\omega_{mk} + O_{day} \pi^{2\omega}_{mk} + \pi^{3\omega}_{mk}] + \sum_{b \in \mathcal{B}} O_{week} \pi^{1\omega}_{mb} \}$$

Each time the slave problems are solved, we obtain a cut for the master problem (3). The master problem is then transformed as follows:

$$\text{(Master problem)} \qquad \min \sum_{i \in S_j} c_{ij} x_{ij} + \mathbb{E}_{\omega \in \Omega_j}[\theta^\omega] \tag{7a}$$

subject to:

$$\sum_{i \in S_j} x_{ij} = 1$$

$$\theta^\omega \geq C_q^\omega + \sum_{i \in S_j} (\sum_{m \in M} \sum_{k \in H} a_{ijk}^m (\beta_{mk}^\omega - \mathbb{1}_{\mathcal{P}_p}(j) \gamma_{mk}^\omega)) x_{ij}, \quad \forall \omega \in \Omega_j, \forall q \tag{7b}$$

$$x_{ij} \in \{0,1\}, \quad \forall i \in S_j$$

Benders Decomposition is also known as row generation. The duals (5) of the slave problems generate cuts that are inserted into the master problem (7). The L-Shaped procedure (Algorithm 6) uses this decomposition.

Algorithm 6 L-Shaped procedure

$q = 0,\ x^q = 0, x = 1$
while $x^q \neq x$ **do**
  $x = x^q$
  $q++$
  for all $\omega \in \Omega_j$ **do**
    solve the dual slave problem with the value x for the event ω
    add the qth cut to the master problem (7)
  end for
  solve the master problem (7) and save the solution in x^q
end while

When the L-Shaped procedure stops, the optimum is reached. Indeed, when the dual slave problems are solved for a value x, the recourse function Q is equal to the cut in this point thanks to the strong duality theorem. Then, if the master problem (7) finds the same solution x, it has to be the optimum, because there is no approximation at this point and all the other points are underestimated.

A.2 Stochastic primal-dual algorithm

Legrain and Jaillet [21] make a simplification: the master problem is solved only once. Then, for each event ω, there is only one cut inserted into the master problem (7). Thus, the constraints (7b) become equalities due to the minimization. The variables θ are therefore replaced in the objective (7a) by their expressions. The objective (7a) becomes: $\min \sum_{i \in S_j} c_{ij} x_{ij} + \mathbb{E}_{\omega \in \Omega_j} [C_q^\omega + \sum_{i \in S_j} (\sum_{m \in M} \sum_{k \in H} a_{ijk}^m (\beta_{mk}^\omega - \mathbb{1}_{\mathcal{P}_p}(j) \gamma_{mk}^\omega)) x_{ij}]$.

We can also remove the constants C_0^ω from the objective to obtain the final master problem (8):

(Master problem)

$$\min \sum_{i \in S_j} (c_{ij} + \sum_{m \in M} \sum_{k \in H} \mathbb{E}_{\omega \in \Omega_j} [\beta_{mk}^\omega - \mathbb{1}_{\mathcal{P}_p}(j) \gamma_{mk}^\omega] a_{ijk}^m) x_{ij}$$

subject to:

$$\sum_{i \in S_j} x_{ij} = 1$$

$$x_{ij} \in \{0,1\}, \quad \forall i \in S_j \tag{8}$$

Finally, the master problem (8) just chooses the pattern with the minimum objective. We can link our procedure with the primal-dual algorithm. The two algorithms are similar; the only difference is the computation of the variables r_i. If we define the variables $u_{ij} = \sum_{m \in M} \sum_{k \in H} \mathbb{E}_{\omega \in \Omega_j} [\beta_{mk}^\omega - \mathbb{1}_{\mathcal{P}_p}(j) \gamma_{mk}^\omega] a_{ijk}^m$, we obtain the stochastic primal-dual algorithm (4).

References

1. Batun S, Begen MA (2013) Optimization in healthcare delivery modeling: Methods and applications. In: Handbook of Healthcare Operations Management, Springer, pp 75–119
2. Begen MA, Queyranne M (2011) Appointment scheduling with discrete random durations. Mathematics of Operations Research 36(2):240–257
3. Benders JF (1962) Partitioning procedures for solving mixed-variables programming problems. Numerische mathematik 4(1):238–252
4. Birge JR, Louveaux F (1997) Introduction to Stochastic Programming. Springer Verlag
5. Boucher P, Beauregard H (2010) Le coût économique du cancer au Québec, en 2008. Tech. rep., Coalition priorité cancer au Québec
6. Buchbinder N (2008) Designing competitive online algorithms via a primal-dual approach. PhD thesis, Technion
7. Canadian Cancer Society (2013) Canadian cancer statistics. Tech. rep.
8. Cayirli T, Veral E (2003) Outpatient scheduling in health care: a review of literature. Production and Operations Management 12(4):519–549
9. Chen Z, King W, Pearcey R, Kerba M, Mackillop WJ (2008) The relationship between waiting time for radiotherapy and clinical outcomes: A systematic review of the literature. Radiotherapy and Oncology 87(1):3–16
10. Ciocan DF, Farias VF (2012) Dynamic allocation problems with volatile demand. Mathematics of Operations Research 37(3):501–525
11. Conforti D, Guerriero F, Guido R (2008) Optimization models for radiotherapy patient scheduling. 4OR: A Quarterly Journal of Operations Research 6(3):263–278
12. Conforti D, Guerriero F, Guido R (2009) Non-block scheduling with priority for radiotherapy treatments. European Journal of Operational Research 201(1):289–296
13. Erdogan SA, Denton B (2013) Dynamic appointment scheduling of a stochastic server with uncertain demand. INFORMS Journal on Computing 25(1):116–132

14. Gupta D, Denton B (2008) Appointment scheduling in health care: Challenges and opportunities. IIE transactions 40(9):800–819
15. Jacquemin Y, Marcon E, Pommier P (2010) Towards an improved resolution of radiotherapy, scheduling. In: 2010 IEEE Workshop on Health Care Management (WHCM), pp 1–6
16. Jaillet P, Lu X (2013) Near-optimal online algorithms for dynamic resource allocations. Working paper
17. Jaillet P, Wagner MR (2010) Online Optimization - An Introduction, INFORMS TutORials in Operations Research, vol 7. INFORMS
18. Kapamara T, Sheibani K, Haas O, Reeves C, Petrovic D (2006) A review of scheduling problems in radiotherapy. In: Proceedings of the 18th International Conference on Systems Engineering (ICSE2006), Coventry University, UK, pp 201–207
19. Karande C, Mehta A, Tripathi P (2011) Online bipartite matching with unknown distributions. In: Proceedings of the 43rd Annual ACM Symposium on Theory of Computing, ACM, pp 587–596
20. Karp RM, Vazirani UV, Vazirani VV (1990) An optimal algorithm for on-line bipartite matching. In: Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, ACM, pp 352–358
21. Legrain A, Jaillet P (2013) Stochastic online bipartite resource allocation problems. CIRRELT
22. Manshadi VH, Gharan SO, Saberi A (2011) Online stochastic matching: Online actions based on offline statistics. In: Proceedings of the 22nd Annual ACM–SIAM Symposium on Discrete Algorithms, SIAM, pp 1285–1294
23. Mehta A, Saberi A, Vazirani U, Vazirani V (2007) Adwords and generalized online matching. Journal of the ACM 54(5):22
24. Ministère de la santé et des services sociaux (2010) Plan stratégique 2010–2015. Tech. rep.
25. Patrick J, Puterman ML, Queyranne M (2008) Reducing wait times through operations research: Optimizing the use of surge capacity. Healthcare Policy 3(3):75–88
26. Petrovic S, Leite-Rocha P (2008) Constructive and GRASP approaches to radiotherapy treatment scheduling. In: Advances in Electrical and Electronics Engineering - IAENG Special Edition of the World Congress on Engineering and Computer Science, pp 192–200
27. Petrovic S, Leung W, Song X, Sundar S (2006) Algorithms for radiotherapy treatment booking. In: Proceedings of the 25th Workshop of the UK Planning and Scheduling Special Interest Group (PlanSIG'2006), Nottingham, UK
28. Santibáñez P, Begen M, Atkins D (2007) Surgical block scheduling in a system of hospitals: an application to resource and wait list management in a british columbia health authority. Health care management science 10(3):269–282
29. Sauré A, Patrick J, Tyldesley S, Puterman ML (2012) Dynamic multi-appointment patient scheduling for radiation therapy. European Journal of Operational Research 223(2):573–584