European Journal of Operational Research 223 (2012) 573–584

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Innovative Applications of O.R.

Dynamic multi-appointment patient scheduling for radiation therapy

Antoine Sauré [a,*], Jonathan Patrick [b], Scott Tyldesley [c], Martin L. Puterman [a]

[a] Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC, Canada V6T 1Z2
[b] Telfer School of Management, University of Ottawa, 55 Laurier Avenue East, Ottawa, ON, Canada K1N 6N5
[c] British Columbia Cancer Agency, 600 West 10th Avenue, Vancouver, BC, Canada V5Z 4E6

ARTICLE INFO

Article history:
Received 1 July 2011
Accepted 27 June 2012
Available online 9 July 2012

Keywords:
Patient scheduling
OR in health services
Markov decision processes
Linear programming
Approximate dynamic programming

ABSTRACT

Seeking to reduce the potential impact of delays on radiation therapy cancer patients such as psychological distress, deterioration in quality of life and decreased cancer control and survival, and motivated by inefficiencies in the use of expensive resources, we undertook a study of scheduling practices at the British Columbia Cancer Agency (BCCA). As a result, we formulated and solved a discounted infinite-horizon Markov decision process for scheduling cancer treatments in radiation therapy units. The main purpose of this model is to identify good policies for allocating available treatment capacity to incoming demand, while reducing wait times in a cost-effective manner. We use an affine architecture to approximate the value function in our formulation and solve an equivalent linear programming model through column generation to obtain an approximate optimal policy for this problem. The benefits from the proposed method are evaluated by simulating its performance for a practical example based on data provided by the BCCA.

1. Introduction

External beam radiation (hereafter referred to as radiation therapy) is a cancer treatment that uses high-energy rays to kill or shrink tumor cells. It is the principal therapy for some types of cancer, but it is also used in combination with other treatments and therapies (e.g. chemotherapy, hormonal therapy and surgery) for many other types of cancer. When a cure is not possible, radiation therapy can be used for palliative purposes for multiple cancer types. In British Columbia, approximately 52% of cancer patients require radiation therapy some time during the course of their illness and 40% receive radiation therapy within 5 years of diagnosis (Source: BCCA registry and treatment databases).

The urgency of radiation therapy can vary for different cancer types and indications, and can be characterized as urgent or routine. It can also be characterized as curative therapy or palliative therapy. Curative therapy can be grouped further into primary curative therapy (where radiation is the main or only treatment given to cure a patient) or adjuvant (when radiation is added to another primary curative therapy, usually surgery, to reduce the risk of relapse). Indications for radiation arise for almost all types of cancer in all organs systems resulting in hundreds of potential cancer sites and treatment types. However, a few cancer sites and treatment types dominate the overall demand for treatment. Radiation therapy is delivered primarily with external beam radiation machines (linear accelerators). These machines may have different energies of radiation, or additional equipment, making them useful for subtypes of radiation treatments and cancer sites. Most radiation departments have a variety of identical machines that serve a range of patient types, allowing for redundancy in the case of breakdowns. Some smaller centers may have two to four identical machines perhaps catering to patients of particular types with more complex patients sent to larger centers with specialized equipment.

Long standing evidence suggests that delays in radiation therapy are associated with tumor progression, persistence of cancer symptoms, psychological distress and decreased cancer control and survival rates (O'Rourke and Edwards, 2000; Fortin et al., 2002; Waaijer et al., 2003; Coles et al., 2003; Mackillop, 2007; Chen et al., 2008). For this reason, many cancer institutions around the world have adopted wait time benchmarks for the time from when the patient is ready to begin treatment to the start of it. In Canada, the maximum acceptable wait suggested by the Canadian Association of Radiation Oncologists for all non-emergency and non-urgent cases is 14 days (Norris, 2009). Unfortunately, fewer than 75% of the radiation therapy treatments in British Columbia are initiated within this time frame (see Table 1).

Delays in radiation therapy are a direct consequence not only of an imbalance between capacity and demand but also a result of inefficient patient scheduling. Three relevant aspects make scheduling radiation therapy treatments especially challenging. First, radiation

* Corresponding author. Tel.: +1 604 632 4241; fax: +1 604 822 9574.
E-mail addresses: antoine.saure@sauder.ubc.ca (A. Sauré), patrick@telfer.uottawa.ca (J. Patrick), styldesl@bccancer.bc.ca (S. Tyldesley), martin.puterman@sauder.ubc.ca (M.L. Puterman).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.06.046

Table 1
Service levels for patients who received radiation therapy in British Columbia between 2004 and 2008. *Source*: BCCA registry and treatment databases.

Year	Patients	% initiated within		
		14 days	28 days	56 days
2004	9834	76.4	97.5	100.0
2005	10144	71.3	96.2	99.9
2006	10168	73.8	97.5	100.0
2007	10487	77.0	96.3	99.9
2008	10318	70.0	96.2	99.9

therapy treatments can be classified into multiple types. The classification is usually made on the basis of cancer site, treatment intent and urgency level. Second, radiation therapy treatments are spread out over time. For most types of cancer, radiation therapy is delivered in daily consecutive sessions for a time period that may vary between 1 day and 8 weeks, with breaks on weekends. Third, radiation therapy sessions do not necessarily have the same duration. Each session is scheduled for a time period ranging from 12 to 60 minutes. The combination of these three aspects of radiation therapy means that a simple first-come-first-served policy will inevitably perform very poorly. This is either because later arriving higher priority demand would be forced to wait longer or else because the uneven session lengths would create unusable gaps in the system resulting in wasted capacity (not unlike in the game Tetris). Thus, due to the importance of timely access to care (as mentioned above) as well as the difficulty of determining intelligent schedules through simpler means, more sophisticated mathematical models are necessary than have hitherto been used for this problem.

To that end, we formulate the radiation therapy appointment scheduling problem as a discounted infinite-horizon Markov decision process (MDP). To deal with an intractable number of variables and constraints, we first approximate the value function in the equivalent linear programming formulation of the MDP using an affine architecture and then solve the dual of the resulting approximate linear programming model through column generation. From the solution we derive an approximate optimal booking policy which we test via simulation. We assume that the machines used for radiation therapy do not differ significantly and thus that treatments can be delivered by any treatment unit. The total treatment capacity is determined by aggregating individual capacities from multiple machines. This assumption is realistic for some small centers equipped with three to four identical machines and focusing on a particular subtype of cancer patients and for subcomponents of larger cancer centers with three to four redundant machines that cater to four to five different cancer sites. For simplicity, we assume treatment requests are observed when patients are ready to begin treatment and not earlier. In this way, we implicitly take into account patient availability and the time needed for any preparatory activities such as imaging and treatment planning. The model could be easily adapted to include a delay between the date of request and the first potential day of service. This delay could be stochastic and differ from one treatment type to another.

From a methodological point of view, our work represents a significant extension of the dynamic multi-priority patient scheduling model and solution approach developed by Patrick et al. (2008). In addition to multiple priority types, we consider patients who receive treatment across multiple days and for irregular lengths of time (see Fig. 1). Furthermore, we allow the possibility of using overtime on different days of the planning horizon, and not necessarily for entire treatments. These additional complications are essential for any realistic attempt to model the scheduling of radiation therapy treatments. The new dimensions to the problem and the possibility of enlarging the system capacity through the use of overtime, together with the non-convex nature of the overtime cost, make this problem much more difficult to model and solve. To the best of our knowledge, ours is the first paper to incorporate multi-appointment requirements when optimizing advance multi-priority patient scheduling policies in a dynamic setting. Our work also constitutes a novel application of approximate dynamic programming to a problem that has received very limited attention in the operations research literature.

The paper is organized as follows. Section 2 summarizes the literature relevant to our work. Section 3 provides a detailed description of the radiation therapy appointment scheduling problem. Sections 4 and 5 describe the proposed MDP model and solution approach, respectively. In Section 6 we provide some managerial insights based on applying our methodology in a small sample problem. We also evaluate the benefits of the proposed method by simulating its performance in a more practical example based on BCCA patient data. Finally, Section 7 states our main conclusions and suggests possible extensions.

2. Literature review

The literature on patient scheduling is usually classified as either *allocation* scheduling or *advance* scheduling. Allocation scheduling refers to methodologies for assigning specific resources and starting times to patients, but only once all patients for a given service day have been identified. Advance scheduling, on the other hand, refers to methodologies for scheduling patient appointments in advance of the service date, when future demand is still unknown. Most studies in the patient scheduling literature address allocation scheduling problems. Magerlein and Martin (1978), Cayirli and Veral (2003), Mondschein and Weintraub (2003), Gupta and Denton (2008), Cardoen et al. (2010) and Begen and Queyranne (2011) provide comprehensive reviews. Our work, however, falls into the second class of problems.

Even though patient scheduling problems have been studied extensively over the last decade, the dynamic allocation of medical capacity in advance of the service date and in the presence of multiple types of patients has received limited attention. Rising et al. (1973) describe a case study where simulation models are used to evaluate alternative decision rules for scheduling appointments in an outpatient clinic. The authors consider two types of patients, walk-ins and patients with advance appointments, and focus on the effects of different booking rules on patient throughput and physician utilization. Gerchak et al. (1996) study the advance scheduling of elective surgeries at an operating room when a random portion of the daily capacity is devoted to emergency surgeries. They analyze the tradeoffs between capacity utilization and delays and prove that the optimal capacity allocation policy is

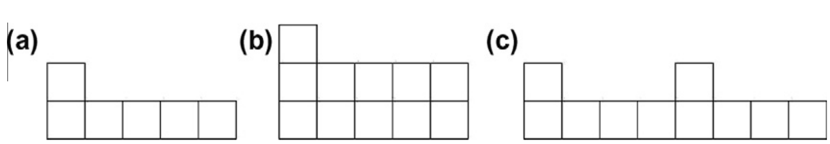

Fig. 1. Graphical representation of some possible treatment patterns. Pattern (b), for example, represents a treatment consisting of a first session of three appointment slots and four additional sessions of two appointment slots each.