Computers & Industrial Engineering 78 (2014) 235–248

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

The planning and scheduling of operating rooms: A simulation approach

R. M'Hallah *, A.H. Al-Roomi

Department of Statistics and Operations Research, College of Science, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait

ARTICLE INFO

Article history:
Available online 1 August 2014

Keywords:
Operating room
Stochastic off-line planning
Stochastic on-line scheduling
Operational level
Operating theater
Under and over utilization

ABSTRACT

This paper considers the stochastic off-line planning and on-line scheduling of operating rooms (OR) at the operational level. It studies the problem for a local public hospital that adopts a closed block scheduling strategy for elective surgical cases. Emergency cases have their dedicated ORs; thus are not considered. At the off-line level, the coordinating surgeon of a given department assigns a subset of elective surgeries to each OR of the operating theater (OT). At the on-line level, unless mandated by extraneous factors, a surgical crew performs all surgical cases assigned to the OR, sequencing them in a non-increasing order of their largest expected surgical time.

This paper investigates how to enhance the OT's expected under and over utilization while maintaining the current average number of surgical cases in general and of major ones in particular. Specifically, it considers two methods that reduce the over utilization of the OT: canceling surgical cases at the on-line operational level, and limiting the workload planned at the off-line operational level. In addition, it considers three management strategies that monitor and control the flow of surgical cases with the objective of diminishing the variability of OR completion times. These strategies are: transferring a surgical case from a busy OR to a free one, using a single queue for all ORs, and adopting an alternative set up of the OT where surgical cases are separated by type.

The paper builds a simulation model for each strategy, and compares the model's output to that of the current situation. It assesses the utility of each strategy based on statistical inferential techniques. Regarding the OR over and under utilization, it is recommended to cancel surgical cases that start after the closing time of the OT. Since this may be difficult to implement, it is advisable to reduce the workload planned at the off-line operational level to 90% of the OT's capacity. Regarding the variability of the OR completion times, it is judicious to adopt a single queue for all surgical cases. When this is infeasible from a technical or a managerial point of view, mixing surgical cases is preferable to separating them by type unless the hospital further decreases the number of minor cases assigned to the OT. Finally, the transfer of the last surgical case from a busy OR to a free one reduces the range of OR completion times.

The proposed simulation model can be easily extended to other hospitals and/or to the strategic and tactical managerial levels. It can account for different constraints and/or managerial procedures. It can be used in conjunction with optimization techniques. Its implementation requires limited knowledge of basic simulation techniques while it offers a simple, user friendly, interactive, decision support system that can be used by coordinating surgeons and management.

1. Introduction

This paper addresses a critical health care problem: the stochastic off-line planning and on-line scheduling of the operating rooms (ORs) of a local hospital at the operational level. ORs "contribute the largest amount of revenues to most hospitals" (Liu, Chu, & Wang, 2011). Yet, they consume 40% of a hospital's resources (Macario, Vitez, Dunn, & McDonald, 1995) and constitute the most expensive resource (van Oostrum et al., 2008). However, ORs draw their importance not only from their potential revenues and prohibitive costs for the society, patients and health care services, but also from their sizeable impact on the safety of patients and on the work flow of other departments (Bosman, 2009; Belien, Demeulemeester, & Cardoen, 2009; Dhupar, Evankovich, Klune, Vargas, & Hughes, 2011). The abundant literature on OR planning and scheduling supports this claim, and emphasizes the role of OR management on the performance of a health care system (Guerriero & Guido, 2010). This role is more obvious in public hospitals, where "surgery brings very little to no revenue...and is

* Corresponding author. Tel.: +965 669 14150; fax: +965 248 37332.
E-mail addresses: rymmha@yahoo.com, rym.mhallah@ku.edu.kw (R. M'Hallah), abrarhhr@yahoo.com (A.H. Al-Roomi).

http://dx.doi.org/10.1016/j.cie.2014.07.022
0360-8352/

often considered an expense. Optimization of surgical services and resources. . .enables these hospitals to provide an overall higher quality of care" (Vijayakumar, Parikh, Scott, Barnes, & Gallimore, 2013).

This paper proposes methods that enhance the mean of the utilization-related performance measures and decrease the variability of the OR completion times while maintaining or improving the current numbers of surgical cases. For the first aspect, it considers case cancelation and the reduction of the workload planned at the off-line operational level. For the second aspect, it studies three on-line operational management strategies, which improve the usage of the OT's open time. These strategies are: (i) the transfer of a surgical case from a busy OR to a free one, (ii) the use of a single queue for all ORs, and (iii) the mixing of different types of surgical cases.

This paper is organized as follows. Section 2 provides a brief overview of stochastic OR planning and scheduling, and motivates the use of simulation modeling. Section 3 describes the OT, defines the off-line planning and on-line scheduling of the ORs as a two-stage stochastic mixed-integer program, details the simulation-based proposed model, validates it, discusses the OT's performance, and pinpoints areas of improvement. Section 4 presents methods that enhance the OT's performance, emphasizing the feasibility of their real life applicability and implementation. Finally, Section 5 summarizes the study and gives possible extensions.

2. Literature review

Because of its practical relevance, financial impact, and human importance, OR planning and scheduling have drawn a lot of attention, quantified by the abundant literature on the subject. Several surveys are available. For example the survey of Cardoen, Demeulemeester, and Belien (2010) classifies the problems by: patient's condition (elective or non-elective), measure of performance (waiting time, throughput, utilization, leveling, makespan, patient deferrals, financial measures, and preference), decision type (date, time, duration, etc.), research methodology (problem considered, solution technique, and constraints), uncertainty (deterministic versus stochastic), and efficiency/efficacy of the technique. Using the above classification, this paper deals with the stochastic problem of determining the workload and utilization of an OT that is dedicated to elective surgical cases, using simulation modeling. Guerriero and Guido (2010) give another structured surgical planning and scheduling literature review with emphasis on operational research techniques and on the most significant manuscripts that apply mathematical and simulation techniques. Mc Intosh, Dexter, and Epstein (2006) give a thorough discussion of simulation-based modeling for OR management. May, Spangler, Strum, and Vargas (2011) classify the literature on surgical scheduling according to its planning horizon and indicate potential areas of research for each of the six classes. Due to the large panoply of manuscripts on the subject, this section focuses on stochastic OR planning and scheduling.

Unlike the current problem where the emphasis is on the OT's performance and where an OR can handle any type of surgical cases, Marcon and Dexter (2006) focus on the effect of the OR's over utilization and sequencing rules on the variation of the workload of the post anesthesia care unit (PACU) in terms of occupied beds and nurses. Their OT has three subsets of ORs, where subset k, $k = 1, 2, 3$, handles only type k surgical cases. They assume that the mean surgical times follow Lognormal distributions and that the actual surgical times follow Lognormal distributions whose mean and standard deviations are derived from the mean surgical times. They consider seven sequencing rules: random, expected longest case first, expected shortest case first, Johnson's rule on expected surgery times, half increase and half decrease of expected surgery time, half decrease and half increase of expected surgery time, and mixed. In addition, they study the possibility of assigning two surgeons per OR per day and of adding urgent cases. For the problem at hand, the mean surgical times follow Uniform distributions and the actual times truncated Normal distributions, with the sequencing rule being the largest expected surgical time first. In addition, the problem at hand deals with elective cases only; i.e., no add-on case is permitted.

Saremi, Jula, ElMekkawy, and Wang (2013) address the appointment scheduling of outpatient surgeries in a Canadian OR department that serves multiple patient types. That is, they consider the second stage of the problem at hand; i.e., the on-line operational scheduling problem with the number of surgical cases pre-determined. Every patient undergoes three steps: preparation, surgery, and recovery where the service time of each of these steps is stochastic. Patients are treated first come first serve. The OT has two beds in surgery preparation, two ORs with two beds each, and two PACU beds for recovery. Their setup is similar to the local hospital except that (i) the proposed simulation model omits the preparation stage, which does not constitute a bottleneck or a delay according to management; (ii) has more ORs; and (iii) has no limit on the number of PACU beds. In addition, they test different time distributions whereas the current paper uses time distributions that are fitted from data collected from the OT. The authors design three tabu search simulation-based models to minimize three conflicting objectives: the waiting time of patients, the makespan of the OT, and the number of surgery cancelations. The first model uses an arbitrary schedule of surgeries. The second applies a deterministic integer program to sequence the surgical cases. The third sequences the cases using a combination of a binary program and a heuristic. To test the validity of the claim of Gul, Denton, Fowler, and Huschka (2011) regarding the usefulness of optimized sequencing, they also test more standard scheduling rules: shortest processing time, longest processing time, increasing variance of service time, increasing coefficient of variability of service time, and the dome-shape rule. Despite its potential benefits, sequencing is generally determined by surgeons based on health-related criteria. Thus, the paper at hand focuses on generally applied sequencing rules rather than on optimized ones. When simulation optimization is the main goal, approaches of Lin, Mustafa, and Pasupathy (2013) or Klassen and Yoogalingam (2009) become useful. The former uses a combined genetic algorithm data envelopment analysis to assess the simulation results and guide the search process while sampling as few points as possible. The second uses simulation optimization to confirm that the dome-shaped sequencing rule is the most robust in terms of many performance criteria but that a flatter dome-shaped rule could be more useful.

Azari-Rad, Yontef, Aleman, and Urbach (2014) apply simulation modeling to determine the impact of peri-operative processes, scheduling and sequencing rules on the proportion of canceled surgical cases in an OT that handles both elective and emergency surgical cases. van Essen, Hans, Hurink, and Oversberg (2012) use simulation modeling to illustrate the effect of inserting emergency surgical cases on the on-line operational level where emergency cases arrive according to a Poisson process and the surgical times of elective cases follow a Lognormal distribution. They investigate the insertion procedure that minimizes the waiting time of emergency cases. Denton, Miller, Balasubramanian, and Huschka (2010) balance the fixed cost of opening ORs and the total cost of overtime utilization of the opened ORs via a deterministic integer programming model, a stochastic recourse model and a robust formulation for setups where the distribution of mean and actual surgical times are unknown. Even though they consider the on-line scheduling problem, their focus is geared towards determining the optimal number of ORs to open rather than on using all ORs and minimizing over-utilization.