An online optimization approach for the Real Time Management of operating rooms

Published version:

DOI:10.1016/j.orhc.2015.08.006

An online optimization approach for the Real Time Management of operating rooms

Davide Duma, Roberto Aringhieri

PII: S2211-6923(14)20071-3
DOI: http://dx.doi.org/10.1016/j.orhc.2015.08.006
Reference: ORHC 87

To appear in: *Operations Research for Health Care*

Received date: 28 November 2014
Accepted date: 26 August 2015

Please cite this article as: D. Duma, R. Aringhieri, An online optimization approach for the Real Time Management of operating rooms, *Operations Research for Health Care* (2015), http://dx.doi.org/10.1016/j.orhc.2015.08.006

An online optimization approach for the Real Time Management of operating rooms

Davide Duma[a], Roberto Aringhieri[a,*]

[a]*Department of Computer Science, University of Torino, Italy*

Abstract

The Real Time Management (RTM) of operating rooms is the decision problem arising during the fulfillment of the surgery process scheduling of elective patients, that is the problem of supervising the execution of such a schedule and, in case of delays, to take the more rational decision regarding the surgery cancellation or the overtime assignment. The main concern of this paper is to propose a model for the RTM and to evaluate its impact on the OR performance assessed by a set of patient- and facility- centred indices. To this end, we consider a generic surgical clinical pathway for elective patients – inspired to a real case study – in which we evaluate the introduction of an online optimization approach for the RTM and some additional optimization modules to deal with the surgery process scheduling problem. To the best of our knowledge, the RTM is not clearly addressed in the literature and this is the first attempt to propose an online approach in the context of surgery process scheduling. We propose a hybrid simulation and optimization model in which simulation is used to model the inherent stochasticity and to replicate the elective patient flow on which the online approach for the RTM and the additional optimization modules operates. We report an accurate computational analysis proving the effectiveness of the proposed approach to the RTM. Finally, we demonstrate the capability and the flexibility of our approach extending our hybrid model to deal with emergency surgeries and different trained surgery teams.

Keywords: Operating Room, Real Time Management, Surgical Pathway,

*Corresponding author

Email addresses: davide.duma@unito.it (Davide Duma), roberto.aringhieri@unito.it (Roberto Aringhieri)

Elective surgery, Emergency surgery

1. Introduction

Problems arising in the Operating Room (OR) planning and scheduling are usually classified into three phases corresponding to three decision levels, namely strategic (long term), tactical (medium term) and operational (short term) [1]. At the operational decision level, the problem arising in the OR management is also called "surgery process scheduling" and is generally divided into two sub-problems referred to as "advanced scheduling" and "allocation scheduling" [2]. The first sub-problem consists in selecting patients from an usually long waiting list and assigning a specific surgery and OR time block to each patient over the planning horizon, which can range from one week to one month [3–10]. Given this advanced schedule, the second sub-problem determines the precise sequence of surgical procedures and the allocation of resources for each OR time block and day combination in order to implement it as efficiently as possible [11–16]. Usually, the two sub-problems have different objectives, that is to maximize the operating room utilization and to minimize the number of surgeries delayed or cancelled, respectively. Furthermore, especially when considering the inherent stochasticity of the problem, the two objectives are conflictual as discussed in [17]. For a complete overview of the problems arising in the OR management, the reader can refer to the papers [18, 19] in which an exhaustive review is reported analyzing in detail multiple fields related to the problem settings and summarizing significant trends in research areas of future interest.

The Real Time Management (RTM) of operating rooms is the decision problem arising during the fulfillment of the surgery process scheduling of elective patients, that is the problem of supervising the execution of such a schedule and, in case of delays, to take the more rational decision regarding the surgery cancellation or the overtime assignment.

The literature reports few attempts to address the problem as shown in [20]. In [21] the authors showed how a computer assisted system could help mitigating the increase of over-utilization of the operating room resources such as overtime. The problem of tardiness from scheduled start times is addressed in [22] comparing the effectiveness of several procedures to reduce tardiness. The authors showed that the generation of a modified or auxiliary OR schedule that compensates for known causes of tardiness can be a good solution to reduce tardiness even if its impact proportionally increases as the number of cases involved. The problem of rescheduling the elective patients upon the arrival of emergency patients is addressed in [23, 24]. The authors proposed a MILP model which considers the overtime cost of the operating rooms and/or the post-anesthesia care units, the cost of postponing or preponing elective surgeries, and the cost of turning down the emergency patients. They proposed a genetic algorithm for its approximate and faster solution. The results of the case study suggest that, instead of shuffling the elective surgeries, it would be worthwhile to consider performing the elective surgeries using the overtime of the operating rooms. Note that the problem of rescheduling patients can be addressed as a particular job shop scheduling problem [25, 26] but these experiences can not directly applied to the operating room context due to its peculiarity in the evaluation of a solution, as we will show in the following. Strategies to move a patient from an operating room to another and based on statistical remarks are proposed in [27–29].

To the best of our knowledge, this is the first attempt to propose an online approach in the context of the operating room management.

The main concern of this paper is to propose a model for the RTM and to evaluate its impact on the OR performance assessed by a set of patient- and facility- centred indices. To this end, we consider a generic surgical clinical pathway for elective patients – inspired to a real case study – in which we evaluate the introduction of an online optimization approach for the RTM and some additional optimization modules to deal with the surgery process scheduling problem.

A Clinical Pathway (CP) can be defined as "health-care structured multidisciplinary plans that describe spatial and temporal sequences of activities to be performed, based on the scientific and technical knowledge and the organizational, professional and technological available resources" [30].

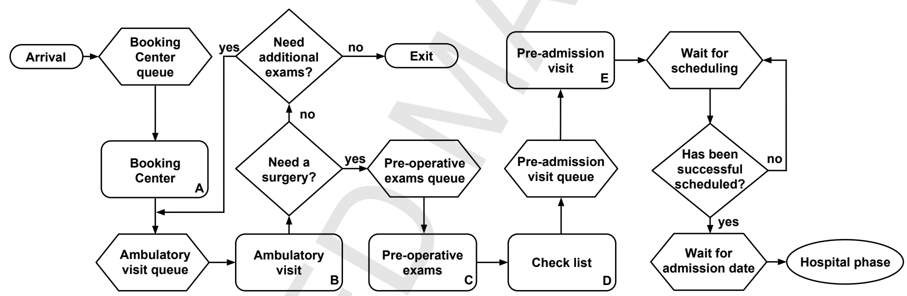
As reported in [31], health care optimization problems are challenging, often requiring the adoption of unconventional solution methodologies. The solution approach proposed herein belongs to this family. We propose a hybrid simulation and optimization model in which simulation is used to model the inherent stochasticity and to replicate the elective patient flow on which the online approach for the RTM and the additional optimization modules operates.

The paper is organized as follows. The three phases of a generic surgical clinical pathway are described in Section 2 pointing out the corresponding optimization problem arising in each phase. Our hybrid simulation and optimization model is discussed in Section 3. In Section 4 we report an accurate computational analysis in order to prove the effectiveness of the proposed approach to the RTM, and to evaluate the impact of the optimization on the management of a surgical pathway. In Section 4.5 and 4.6 we demonstrate the capability and the flexibility of our approach extending our hybrid model to deal with emergency surgeries and different trained surgery teams. Section 5 closes the paper.

2. Surgical clinical pathway and optimization problems

The definition of the surgical pathway is inspired to that presented and analyzed in [32] for the thyroid surgical treatment. The reader can refer to this paper for further details. From a management point of view, a surgical pathway can be seen as made up of three phases.

Figure 1: Pre-admission phase flowchart

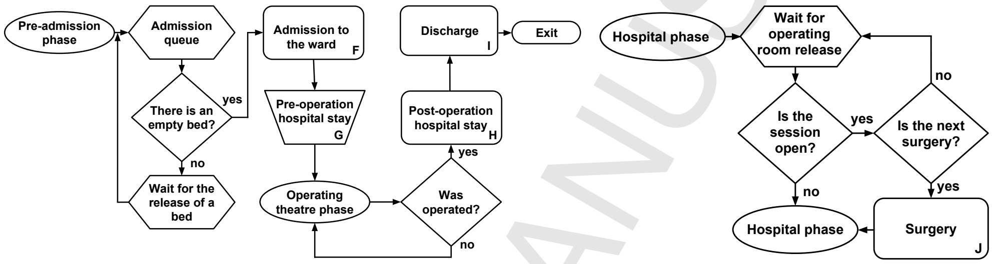


The first phase concerns the *pre-admission phase* and it is related to all the activities regarding patients before their admission (see Figure 1).

In this phase, a relevant information is the Diagnosis Related Group (DRG). A DRG defines a general time limit before which the patient should be operated on. Note that the DRG refers to the access time (i.e., days to surgery) and not to the waiting time on the day of surgery. In our context, a *Urgency Related Group* (URG) is assigned to each patient belonging to the same DRG: the URG states a more accurate time limit called *Maximum Time Before Treatment* (MTBT). In other words, URG allows to define a partition of the patients in the same DRG in order to prioritize their surgical operation. The optimization problem arising in this phase is the advanced scheduling problem, which consists in the selection of patients from the (usually long) waiting list and in their assignment to an *OR session* (i.e., an operating room on a given day) in such a way that several operative constraints are satisfied (number of beds available during the patient stay, total time available for the OR session, and so on). Our objective is to maximize the utilization of the operating rooms in each day in such a way to guarantee that each patient is operated within the time limit defined by the URG. This problem is well known in the literature as Surgical Case Assignment Problem (SCAP) [33].

Figure 2: Flowcharts of the hospital and operating theatre phases

(a) Hospital phase (b) Operating theatre phase



The *hospital phase* is concerned with all the activities involving the admitted patient stay except for those related to the operating theatre as depicted in Figure 2a. The relevant information in this phase is the Length Of Stay (LOS) of each patient, that is the number of days required before the discharge. The optimization problem arising in this phase is the allocation scheduling problem, which consists in finding a sequence of patients to determine the order in which they are operated on. The objective is to minimize the risk of cancellation, while keeping an acceptable utilization rate with respect to the available operating time taking into account a patient-centred point of view (considering waiting time, class of urgency, possible previous referrals).

Figure 2b depicts the *operating theatre phase*, which is a component of the hospital phase, as highlighted in Figure 2a. Due to its importance in a surgical pathway, it requires to be treated separately. Patients assigned to a given OR session will be operated on following the sequence previously defined unless delays imposes to define a new sequence. Patients not operated on will be rescheduled. We could have a delay as soon as the *Estimated Operating Time* (EOT) differs from the *Real Operating Time* (ROT). The RTM operates when such a delay become significant, that is exceeding the total operating time allowed. The following possible decisions should be considered:

- to use some overtime reducing the total amount weekly available;
- to cancel 1 or more surgeries and to re-schedule them, when possible;
- to change the sequence of the remaining patients in order to minimize the cancellation of surgeries whose patients are close to their MTBT while keeping an acceptable level of OR utilization.

The first two choices are generally non-trivial and alternatives requiring to consider several aspects. For instance, the decision of postponing a patient could violate MTBT. Further, it determines an increased patient stay lowering the patient satisfaction and, by consequence, the quality of the service. On the other side, overtime is a scarce resource. So, it seems crucial to establish some criteria driving the decisions of using it to avoid cancellations.

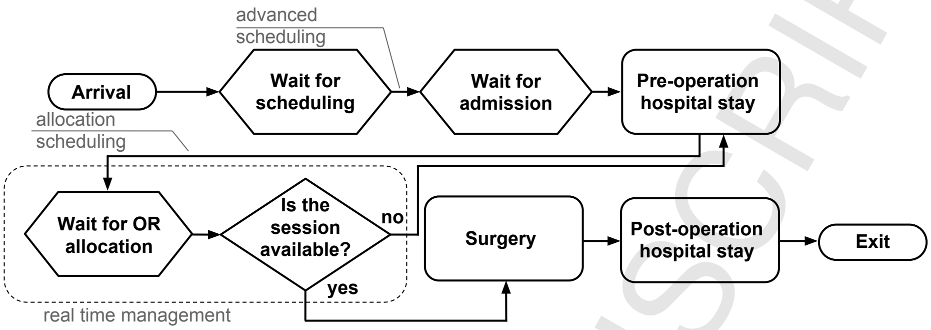
Considering the surgery process scheduling at the operational level means that we are not directly dealing with the "Master Surgical Schedule" (MSS) and, therefore, with the assignment of operating rooms to wards or specialties taking into account also the physician (e.g., surgeons and anesthetists) availability. From our point of view, this will be implicitly considered when defining the scenarios for the quantitative analysis: actually, if we suppose to have an OR available for 300 minutes means that the required physicians are available. Note that this is a quite common assumption when dealing with surgery process scheduling.

3. The Hybrid Model

This section discusses the hybrid simulation optimization model proposed in this paper. Simulation is exploited to model the inherent stochasticity that characterizes the problems arising in the operating room management, that is the arrival of patients, the variability of patient length of stays and the variability of patient operating times (see, e.g., [34–36]). Furthermore, it allows to easily replicate the three phases of the patient flow depicted in Section 2. On this simulated surgical clinical pathway, it is possible to embed the optimization modules to deal with the decision problems described in Section 2.

Figure 3 summarizes how the patient passes through the surgical pathway highlighting when the optimization operates: the advance scheduling

Figure 3: Description of the hybrid simulation and optimization model



manages its admission, the allocation scheduling manages its position in the surgery sequence and, finally, the RTM manages the ongoing operations before the surgery. Summing up, simulation allows to model the operative context required by the optimization modules to operate correctly over the time horizon needed to evaluate the impact of such optimization modules.

In the following, we will briefly describe the hybrid model through the description of its main components, that is the simulation framework and the three optimization modules.

Table 1 introduces the notation of the problem used hereafter in the paper.

Table 1: Notation

N: number of OR sessions	S_j: duration of j-th OR session
d_j: operating day (from Mon to Fri) of the j-th OR session	k: index of the day, $k = 1, \ldots, 7$
B_k: number of beds available the k-th day of the week	Ω: weekly overtime available
I: set of patients in the pre-admission waiting list	L: set of scheduled patients
$L^{(j)}$: set of patients scheduled into the j-th OR session	M_i: MTBT of patient i
t_i: waiting time of the i-th patient	ℓ_i: LOS of patient i
e_i: EOT of patient i	r_i: ROT of patient i

3.1. The Simulation Framework

The simulation framework is based on a Discrete Event Simulation (DES) since it is the most suitable methodology to analyze a discrete and stochastic workflow. Further, DES is the only approach capable to represent the single entities within a CP, which is a necessary condition to apply the proposed optimization planning modules. The proposed simulation model is a straightforward implementation of the surgical pathway depicted in Figure 1 and 2. The main parameters of the simulation model, and their distribution, are depicted in Appendix A.

Note that the hybrid model is implemented using AnyLogic 6.9 [37]: its Enterprise Library is exploited for the implementation of the DES simulation framework whilst the optimization modules are implemented from scratch in Java, which is the native programming language of AnyLogic.

3.2. Solving the Advanced Scheduling Problem

We propose a metaheuristic based on a greedy construction of an initial solution and then a local search to improve that solution. The proposed algorithm is a simplified version of that discussed in [38]. The operative context is represented by a long queue of patients from which we would like to select a subset of patients to be admitted taking into account the fact that the resources available can be reduced since patients admitted the previous week are already in the hospital phase, usually waiting for the discharge but also for their surgery. From a temporal point of view, we suppose to plan the next week of surgeries at the end of the current week, that is on Friday.

3.2.1. Constructive greedy algorithm

The algorithm associates to each patient $i \in I$ the following values

$$w_i = \frac{t_i + \min_{1 \le j \le N} d_j}{M_i}, \tag{1}$$

$$\tilde{w}_i^k = \frac{t_i + \min_{1 \le j \le N} d_j + \pi(k)}{M_i} = w_i + \frac{\pi(k)}{M_i}, \tag{2}$$

where $\pi(k)$ measures the distance of the current day k to the next Friday: for instance, at the moment of determining a solution for the advance scheduling problem $\pi(k)$ is equal to 7. The value w_i measures the ratio of the time elapsed before the surgical operation and the MTBT associated to the URG of the patient $i \in I$ whilst $\tilde{w}_i^k$ is a projection of w_i referred to the next week.

Starting from the schedule containing the patients planned the previous week, patients to be admitted and belonging to the admission queue are ordered by decreasing value of w_i in such a way to promote the scheduling of those patients which are close to their MTBT. Then, each patient is considered for the scheduling. A patient will be inserted in the current schedule if there exists an OR session available with enough free operating time in such a way to satisfy the operative constraints regarding the bed occupation.

Among different possible OR sessions, the algorithm tries to schedule the patient first in the day k such that $k + \ell_i \leq 5$. If it is not possible, the algorithm tries the insertion in the day k such that $k + \ell_i > 5$. The rationale here is to avoid the use of the weekend stay beds which could be a limited resource. This rule can be overridden when $\tilde{w}_i^k \geq 1$ assigning the patient to the first day $k = 1$, if possible, or to the second day $k = 2$, and so on. In this case, we would like to reduce the probability of not satisfying the URG requirements in case of cancellation. Finally, if a patient cannot be scheduled, the algorithm will consider the next patient. The algorithm terminates when all patients in the queue have been considered for the insertion in the current schedule.

3.2.2. Improvement local search algorithm

The Local Search tries to improve the greedy solution by exchanging pairs of patients already scheduled in such a way to cluster them in a reduced number of OR sessions and, by consequence, to allow the insertion of new patients previously not scheduled. Let j^* be the OR session having the maximum operating time yet available, that is the one having the minimal utilization. The Local Search algorithm follows these criteria to select the new incumbent solution:

- the new solution will be that providing the maximal increase of the time yet available of j^*;
- otherwise, if the two schedules are equivalent in j^*, the algorithm will consider the second least utilized OR session, and so on;
- otherwise, if the two schedules are equivalent in all OR sessions, the algorithm selects those solutions having OR sessions less utilized at the end of the week.

3.3. Solving the Allocation Scheduling Problem

In our settings, the allocation scheduling problem consists in establishing the order in which patients $i \in L^{(j)}$ will be operated on in such a way to minimize the inefficiency due to possible cancellations.

Considering a given schedule, there is a set of patients for which is better to avoid the cancellation of their surgery, that is those patients whose $\tilde{w}_i^k$ is greater than or equal to 1 and those patients whose surgery was already postponed. To deal with these special cases, let us introduce the following values:

$$W_i = \begin{cases} \tilde{w}_i^k & \text{if } \tilde{w}_i^k > 1 \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

and let $D_i > 0$ be the number of days elapsed after a cancellation, 0 otherwise. Finally, we define the value

$$s_i = \alpha_1 W_i + \alpha_2 D_i + \alpha_3 e_i \tag{4}$$

for each $i \in L^{(j)}$ where α_1, α_2 and α_3 are parameters. Setting

$$\alpha_1 \gg \alpha_2 \gg \alpha_3 = \begin{cases} 1 & \text{case (A)} \\ -1 & \text{case (B)} \end{cases},$$

the sequencing of patients $i \in L^{(j)}$ is simply obtained by ordering them by decreasing order of their s_i.

Within such a ordering, the use of α imposes three priority levels. First we schedule patients close to their MTBT. Then we schedule those whose surgery was previously postponed in such a way to foster those waiting for more days after the cancellation. Finally, when the first two components of s_i, that is $\alpha_1 W_i$ and $\alpha_2 D_i$, yield to the same value for two different patients, we break ties by ordering them following a LPT or a SPT policy (with respect to EOT) in the case (A) and in the case (B), respectively.

3.4. An online approach to the Real Time Management

The solutions discussed in the previous sections provide a schedule based on the EOT, which is usually an estimate of the surgeons. Unfortunately, it is possible that the ROT differs from the EOT. Given $L^{(j)}$ and a patient $i \in L^{(j)}$, the whole schedule could be delayed if $r_i > e_i$. When the overall delay could determine the exceeding of the jth OR session duration S_j, the RTM should deal with the problem of postponing a surgery or using a part of the overtime available. Such a decision poses the problem of evaluating the impact of consuming overtime or to have a cancellation.

Let us consider the jth OR session on day $k = d_j$ having duration S_j and a list $L^{(j)}$ of scheduled and sequenced patients. Suppose that $m < \left|L^{(j)}\right|$ patients are already operated on. Let ρ_m the effective time elapsed to operate on the m patients, that is

$$\rho_m = \sum_{i=i_1,\ldots,i_m} r_i. \tag{5}$$

Let us introduce the following parameter:

$$\beta^j_{km} = 1 + \frac{N_k}{N} - \frac{\Omega^j_{km}}{\Omega} \qquad (6)$$

where Ω^j_{km} is the residual overtime after the surgery of patient i_m and N_k is the number of the remaining OR sessions scheduled from the day $k+1$. Note that $N_k = 0$ if k corresponds to Friday.

The value β^j_{km} would measure the overtime still available with respect to the number of OR sessions to be still performed. Actually, β^j_{km} is closed to 1 when the overtime was used proportionally; it is between 0 and 1 or it is greater than 1 when it was underused or overused, respectively. Because of N_k is equal to 0, we remark that the last day of the week it is always less than or equal to 1 hence promoting the use of the residual overtime. The online algorithm starts every time a surgery ends and $\rho_m > \sum_{i=i_1,\ldots,i_m} e_i$. It consists of three procedures.

Sequencing check. The sequencing of the remaining patients is checked in such a way to ensure that (i) all the remaining patients having $\tilde{w}^k_i > 1$ are scheduled prior to the other patients and (ii) those having $\tilde{w}^k_i > 1$ are ordered by decreasing value of $\tilde{w}^k_i$; if those patients run out the available operating time S_j, the patients having $\tilde{w}^k_i \leq 1$ maintain the same original ordering; otherwise, the free operating time is filled selecting a subset of the patients having $\tilde{w}^k_i \leq 1$ according to the Bin Packing Best Fit rule;

Patient postponing. Let i_{m+1} be the next patient in the schedule. Then, if

$$e_{i_{m+1}} > S_j - \rho_m \; ,$$

the patient i_{m+1} could incur in a cancellation. Therefore, the algorithm checks if

$$\beta^j_{km} \left(\frac{e_{i_{m+1}} + \rho_m}{S_j} \right) \leq 1 \qquad (7)$$

and if (7) is satisfied, the required overtime is assigned to the patient i_{m+1}.

Rescheduling. At the end of the day, all the postponed surgeries must be rescheduled on OR sessions having enough free operating time. First the algorithm considers all the patients having $\tilde{w}^k_i > 1$ trying to insert each patient in the first OR session available. Then, the algorithm tries to insert iteratively subsets of patients having $\tilde{w}^k_i \leq 1$ according to the

Bin Packing Best Fit rule. If an insertion is not possible, the patient will be scheduled on the first day available in the next week.

Finally, we remark that the algorithm for the insertion of a subset of patients, used both in the sequencing check and in the rescheduling procedures, is and adaptation of the dynamic programming discussed in Section 3.4.1 of [39]. For a description of the Best Fit rule for the Bin Packing, the reader can refer to Section 8.2 of [39].

4. Quantitative analysis

This section reports the quantitative analysis performed in order to evaluate the impact of the online approach to the RTM and the additional optimization modules on the management of a surgical clinical pathway.

The main idea behind the proposed quantitative analysis is to evaluate their impact week by week, that is how the previous decisions (e.g., determining less or more cancellations) can impact on the current decisions.

In our work, we are considering the surgery process scheduling problems arising at the operational level, which has usually a planning horizon of a week. The idea behind our quantitative analysis is therefore to evaluate the impact of such plannings week by week, that is how the previous decisions (e.g., determining less or more cancellations) can impact on the current decisions.

Section 4.1 describes how the computational experiments are carried out reporting the possible configurations of the optimization modules, the performance indices and the different evaluation scenarios. Section 4.2 reports about the logical validation of the simulation model discussed in Section 3.1. Section 4.3 and Section 4.4 report the results of the computational tests made on two different evaluation scenarios. Finally, Section 4.5 and 4.6 extend the original hybrid model to deal with the emergency surgeries and different trained surgery teams in order to prove the capability and the flexibility of our approach.

The results reported in the following sections are the average value among those obtained by running the hybrid model 30 times on a given configuration and, each time, starting from a different initial conditions. On average, one single run requires from 1.3 to 4.3 seconds when running with all the optimization approaches turned off or turned on, respectively. This means that no more than $4.3 \times 30 = 129$ seconds are needed to simulate two years of operating room management. Finally, we remark that the algorithms for the advanced scheduling are the most time consuming components while the running time required by the other optimization algorithms are negligible. Finally, we remark that all the simulation parameters are depicted in Appendix A. The Appendix describes and reports the parameters regarding the patient flow characteristics, the duration of the activities and their distributions, and all the other parameters characterizing our simulation such as the values for each class of URG and the number of beds available.

4.1. Test configurations, performance indices and scenarios

The optimization algorithms described in Section 3.2, 3.3 and 3.4 can be combined in different ways In order to evaluate their actual impact, we define a *baseline configuration* with respect to the three phases as follows:

Phase 1: advanced scheduling performed by a first-fit algorithm, that is (i) it considers patients by decreasing order of w_i, (ii) it scans the OR session from Monday to Friday and assigns the selected patient to the first one having enough operating time available (if possible);

Phase 2: the patient sequencing is that resulting from the patient assignment, that is, the first assigned to an OR session will be the first in the sequence, and so on;

Phase 3: overtime is assigned *a priori* uniformly to all OR sessions in an amount equal to $\frac{\Omega}{N}$;

Phase 3: all the surgeries are rescheduled only at the end of the day using the first-fit algorithm, that is the first phase of the RTM rescheduling algorithm.

Besides the baseline configuration, we define further configurations to evaluate the impact of the optimization modules. Each configuration is defined with respect to the baseline configuration.

- **Phase 1**:
  - **option 1:** computing w_i w.r.t Monday instead of the previous Friday (in the simulation model, Friday is the day in which the advance scheduling is performed);
  - **option 2:** adopting the greedy explained in Section 3.2.1 (instead of the first-fit algorithm);
  - **option 3:** adopting the Local Search depicted in Section 3.2.2;
- **Phase 2**:
  - **LPT/SPT:** use LPT or SPT rules in sequencing (case (A) or (B) in Section 3.3), respectively;
- **Phase 3**:

option A: adopting the RTM online algorithm after each surgery;
option B: adopting the rescheduling algorithm depicted in Section 3.4 at the end of the day (instead of the first-fit algorithm).

Table 2 reports the two types of indices adopted to evaluate the impact of the optimization modules. We define a set of patient-centred indices in such a way to evaluate the performance from a patient point of view. We also define a set of facility-centred indices in such a way to evaluate them against to the patient-centred ones.

Table 2: Patient-centred and facility-centred indices

Index	Definition
	Patient-centred
C	number of cancellations
f_{MTBT}	percentage of patients operated within the MTBT
I_{avg}	average length (number of patients) of the waiting list
t_{avg}	average waiting time spent in the waiting list
w_{avg}	average value of patient's w_i at the time of their surgery
w_{max}	maximum value of patient's w_i at the time of their surgery
	Facility-centred
u_{bed}	bed utilization
u_{OR}	OR session utilization

It is quite evident that different indices can affect each other. For instance, the increase of the number of cancellations can affect the bed utilization and, in its turn, can reduce the percentage of patients operated within the MTBT.

Table 3 describes the two different scenarios in which we evaluate the optimization solutions on different operating contexts. The two scenarios are characterized by about the same overall amount of operating time available (7920 vs. 7980 minutes) distributed in a different way with respect to the number N of available OR sessions (21 vs. 15) and their duration S_j, $j = 1, \ldots, N$.

4.2. Simulation Model Validation

The validation of a simulation model requires a quite complex analysis. In our case, we are only interested in the logical correctness of the simulation model representing the surgical pathway. On the other side, we are not interested in the replication of a real system.

Table 3: Scenarios

(a) Scenario 1

	OR 1	OR 2	OR 3	OR 4	OR 5
mon	300	360	420	420	420
tue	300	360	420	420	
wed	300	360	420		
thu	300	360	420	420	420
fri	300	360	420	420	

(b) Scenario 2

	OR 1	OR 2	OR 3
mon	540	540	540
tue	540	540	540
wed	540	540	540
thu	540	540	480
fri	540	540	480

To this end, we adapted our simulation model to represent the inspiring case, that is that reported in [32]. In that paper, the proposed model dealt with two patient flows having similar EOT but different LOS. Note that the LOS of the second flow is roughly the double of the first one while the number of patients in the first flow is roughly the double of the second flow. Since our model can generate only one type of patient flow, we adapted our patient flow generator in such a way to have, on average, the same number of patients having the LOS of the first flow which is the most numerous. In this validation scenario, we have $N = 7$ OR sessions having the same duration equal to 360 minutes. Two OR sessions are scheduled on from Tuesday to Thursday and one on Friday. The other parameters are set to the same value reported in Appendix A. Furthermore, we turn off all the optimization during the three phases. In Table 4 we compare the results of our adapted simulation model with those reported in [32].

Table 4: Model validation: comparison with real measures

	u_{bed}	u_{OR}
Real measures	51.1%	77.3%
Simulation model	49.1%	80.8%
Difference	2.0%	3.5%

The differences in the two performance indices can be accounted to the different composition of the patient flow as depicted above. For instances, the gap of 3.5% for u_{OR} expressed in minutes corresponds to the execution of one surgery having average duration. On the basis of these considerations, the comparison is satisfactory with respect to our objective, which is the validation of the logical correctness of our simulation model.

4.3. Scenario 1: analysis

We tested all the possible configurations that can be obtained combining the options defined in Section 4.1. Our aim is to identify the best configuration which increases the patient-centred indices without deteriorating the facility-centred ones. First, the impact of each optimization modules is evaluated through the quantitative analysis. Based on these results, two further configurations have been studied. The results are summarized in Table 5, which reports the value of the performance indices for each test configuration denoted by the value in the first column "id". Note that the column reporting the number of cancellations also reports in brackets the total number of patients operated on. All the results are compared with those obtained for the baseline configuration.

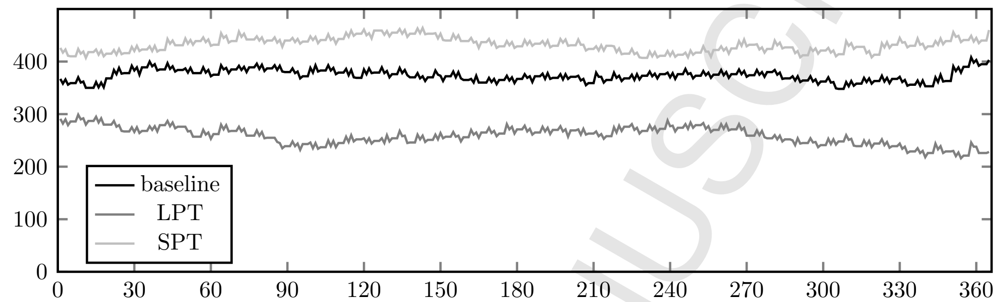
Table 5: Performance indices for each test configuration

Option(s)							**Performance indices**							
id	1	2	3	seq.	A	B	C	$f_{\mathbf{MTBT}}$	$I_{\mathbf{avg}}$	$t_{\mathbf{avg}}$	$U_{\mathbf{bed}}$	$U_{\mathbf{OR}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$
(0)	baseline configuration						234 (2348)	32.6%	338	55	63.6%	89.9%	1.17	4.05
(1)	√						235 (2347)	31.9%	346	56	60.2%	89.8%	1.11	3.29
(2)	√	√					226 (2340)	26.0%	360	58	60.6%	89.3%	1.16	3.27
(3)			√				252 (2346)	36.0%	324	52	60.4%	89.6%	1.12	3.61
(4)	√		√				246 (2349)	35.3%	330	53	60.3%	89.8%	1.06	3.41
(5)	√	√	√				230 (2338)	27.2%	355	58	60.8%	90.0%	1.17	3.10
(6)				LPT			236 (2367)	47.9%	292	48	60.5%	90.8%	1.03	3.79
(7)				SPT			240 (2261)	12.1%	452	72	58.6%	86.4%	1.51	4.91
(8)					√		197 (2384)	74.6%	213	35	59.3%	91.3%	0.80	2.64
(9)						√	236 (2315)	30.7%	339	55	72.6%	88.8%	1.18	3.79
(10)					√	√	222 (2372)	73.0%	223	37	64.0%	90.7%	0.83	2.68
(11)			√	LPT	√		239 (2389)	79.9%	192	32	60.3%	91.8%	0.73	2.62
(12)	√		√	LPT	√		248 (2390)	85.5%	207	34	60.6%	91.8%	0.71	1.87

Regarding the impact of the advanced scheduling optimization module, we can observe a lower waiting time in the waiting list and an improvement of the performance indices related to MTBT in test configurations (3) and (4). On the other side, the minimal number of cancellations is obtained with configuration (2) but, at the same time, the percentage of patients operated on before their MTBT decreases consistently. Note that the use of Local Search allows to insert more patients determining the improvement measured in (3) and (4).

Regarding the impact of the allocation schedule optimization module, we can observe significantly better performances when LPT policy is adopted. Figure 4 shows the trend of I_{avg} under the baseline, (6) and (7) configurations.

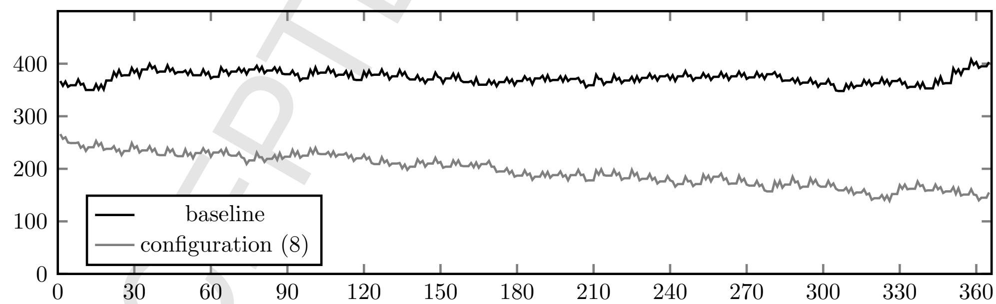
Figure 4: Trend of I_{avg} (data referred to the 2nd year, days on x-axis, patients on y-axis)



Regarding the impact of the online approach for the RTM, we observe a remarkable improvement of all the performance indices (see configurations (8) and in particular f_{MTBT}). On the other side, we observe the negligible impact of the algorithm for the rescheduling postponed patients at the end of the day (see configurations (9) and (10)).

Figure 5 and 6 show respectively the trend of I_{avg} and w_{avg} under the baseline and (8) configurations. Note that it is positive when $w_{\mathrm{avg}} < 1$ which means that all the patients are operated on before their MTBT, on average.

Figure 5: Trend of I_{avg} (data referred to the 2nd year, days on x-axis, patients on y-axis)



Finally, configurations (11) and (12) report about the combination of the different best options. We note a further improvement of the performance indices except for that related to the number of cancellations if compared with configuration (8). This is due to the fact that Local Search allows to insert more patients in the advanced scheduling thus reducing the waiting time in the waiting list but increasing the probability of incurring in a cancellation. Figure 7 shows the trend of w_{avg} under the baseline, (11) and (12) configurations. While baseline configuration shows a value of w_{avg} always greater than 1, we remark that both configurations (11) and (12) tend to be less than 1. Further, configuration (12) seems more stable and powerful in reducing this index.

4.4. Scenario 2: analysis

The second scenario differs from the first one in terms of the schedule of the OR sessions. As for scenario 1, the impact of each possible configuration is evaluated and then, based on these results, four further configurations have been studied. The results are summarized in Table 6. All the results are compared with those obtained for the baseline configuration.

Comparing the results for the two scenarios, we can observe that the number of cancellations with respect to the number of operated patients is almost the same. Further, the utilization indices (U_{OR} and U_{bed}) ranges around the same values, that is 60% and 90% for beds and OR sessions, respectively. The comparison of the results reported for configurations (6) and (7) confirms the fact that LPT can provide better results than SPT. The significant reduction of the waiting list length and of the patient waiting time is confirmed also in the analysis of the second scenario. Similarly, the results confirm the significant reduction of w_{avg} and w_{max}.

Figure 6: Trend of w_{avg} (data referred to the 2nd year, days on x-axis, patients on y-axis)

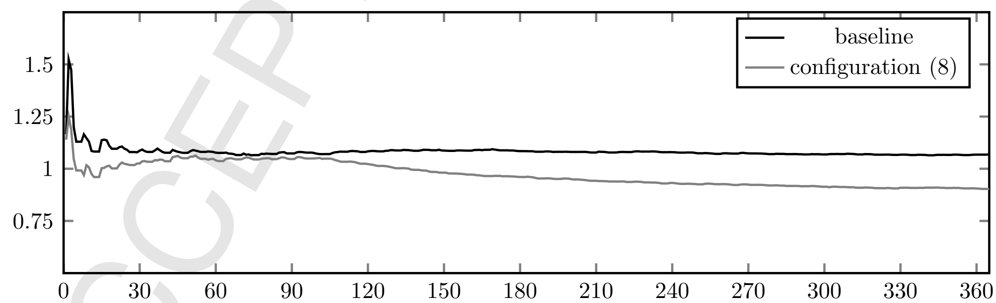


Figure 7: Trend of w_{avg} (data referred to the 2nd year, days on x-axis, patients on y-axis)

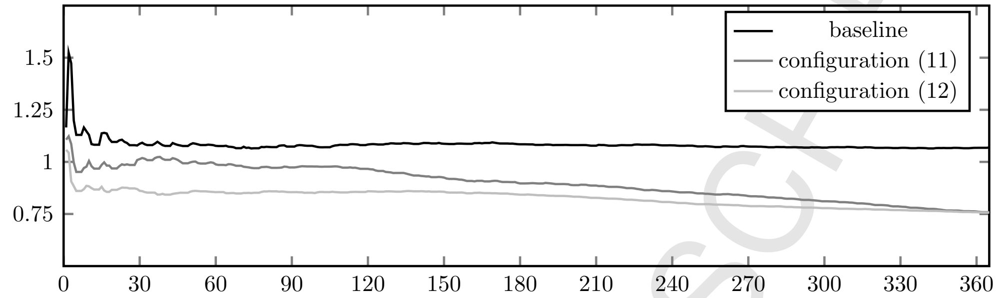


Table 6: Performance indices for each test configuration

Option(s)							Performance indices								
id	1	2	3	seq.	A	B	C		$f_{\mathbf{MTBT}}$	$I_{\mathbf{avg}}$	$t_{\mathbf{avg}}$	$U_{\mathbf{bed}}$	$U_{\mathbf{OR}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$
(0)	baseline configuration						187	(2411)	57.7%	269	44	61.5%	92.0%	0.97	3.80
(1)	√						199	(2408)	60.0%	260	43	62.0%	92.1%	0.95	3.61
(2)	√	√					181	(2403)	70.5%	250	41	61.0%	91.8%	0.85	3.12
(3)			√				200	(2404)	61.1%	250	41	61.9%	92.0%	0.92	3.66
(4)	√		√				203	(2409)	67.7%	256	42	62.3%	92.0%	0.87	2.94
(5)	√	√	√				198	(2398)	69.2%	250	42	61.7%	91.9%	0.86	3.22
(6)				LPT			174	(2462)	85.1%	151	25	61.7%	94.3%	0.62	3.16
(7)				SPT			202	(2346)	27.0%	355	58	60.8%	89.5%	1.24	3.79
(8)					√		177	(2422)	84.7%	165	27	61.0%	92.6%	0.67	2.52
(9)						√	189	(2405)	62.1%	252	41	62.6%	92.0%	0.92	3.83
(10)					√	√	176	(2411)	84.4%	164	27	64.0%	92.5%	0.66	2.51
(11)	√	√		LPT	√		155	(2430)	96.0%	123	21	59.8%	92.8%	0.49	2.16
(12)	√	√	√	LPT	√		205	(2434)	95.3%	127	21	62.4%	93.1%	0.50	2.15
(13)	√	√		LPT	√	√	159	(2426)	96.5%	123	21	60.2%	92.8%	0.49	2.16
(14)	√	√	√	LPT	√	√	201	(2419)	96.8%	119	20	61.6%	92.6%	0.48	2.18

In both scenarios, configurations (11)–(12) and (11)–(14) are those providing the best overall performances. Configurations (11)–(14) confirm the impact of Local Search during the pre-admission phase: Local Search is capable to insert more patients in the scheduling thus reducing the average length of the waiting list but increasing the number of cancellations.

4.5. Dealing with a flow of emergency surgeries

The management of emergency surgery is quite a complex task: actually, delaying emergency surgery may increase the risk of postoperative complications and morbidity. Therefore, the responsiveness of the surgical pathway, that is the speed at which an OR is available for that surgery, is the crucial factor to guarantee a positive final outcome.

To deal with emergency surgeries, a common approach is to reserve OR capacity since it is believed to increase the responsiveness. This approach poses a question, that is if it is better to have dedicated emergency ORs (see, e.g.,[40]) or, alternatively, to reserve capacity in the elective ORs (see, e.g.,[41]).

In this section, we would like to evaluate the impact of introducing a patient flow of emergency surgeries within an optimized surgical pathway. Basically, we would evaluate the RTM capability of dealing with the emergencies. To this end, we modified our simulation model adding an emergency patient flow sharing only the ORs with the elective patient flow, that is the emergency patients have dedicated stay beds. The emergency patient flow is generated in such a way to have, on average, one emergency patient each day having the same EOT and ROT of an elective patient with the highest level of URG. In our setting, a patient requiring an emergency surgery is operated as soon as an OR becomes available. This means that no changes are considered in the algorithms for determining a solution for the advanced and the allocation scheduling. We tested the modified model on the second scenario (Table 3b) taking into account the baseline, (13) and (14) configurations. Here we did not consider the first scenario (Table 3a) since it has more ORs and therefore, it is easier to assign an emergency patient without worsening the solution.

Table 7 reports the value of f_{MTBT} for the emergency patients varying MTBT value between 30 and 240 minutes.

Table 7: f_{MTBT} for emergency patients w.r.t. different MTBT

id	30	60	90	120	150	180	210	240
(0)	55.7%	75.4%	88.6%	95.1%	98.3%	99.7%	100.0%	100.0%
(13)	52.2%	72.9%	84.1%	88.3%	95.5%	98.1%	99.4%	100.0%
(14)	53.5%	73.1%	83.3%	89.7%	95.0%	98.1%	99.6%	99.9%

Table 8 reports the performance of the whole surgical pathway after introducing the emergency patient flow. Note that the last three columns report the value for the indices w_{avg}, w_{max} and f_{MTBT} referred to the emergency patients with a MTBT threshold set to 60 minutes.

As one might expect, it may be noted a general worsening of the patient-centred indices for the elective patients. On the other side, the three indices

Table 8: Evaluating performance of best configurations in Table 6

	Elective							Emergency		
id	C	$I_{\mathbf{avg}}$ ($t_{\mathbf{avg}}$)	$U_{\mathbf{bed}}$	$U_{\mathbf{OR}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$	$f_{\mathbf{MTBT}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$	$f_{\mathbf{MTBT}}$
(0)	294 (2302)	410 (66)	62.9%	95.1%	1.4	4.8	16.4%	0.60	3.04	75.4%
(13)	370 (2271)	302 (50)	73.9%	94.4%	1.0	3.3	48.2%	0.68	3.59	72.9%
(14)	367 (2300)	323 (52)	71.0%	95.6%	1.1	3.4	41.1%	0.68	3.47	73.1%

referred to the emergency patients shown quite satisfactory results considering the really tight MTBT threshold and the absence of any optimization.

Recalling the model depicted in Section 3.4, RTM decisions largely depend on the ratio $\frac{\Omega^j_{km}}{\Omega}$, that is from the amount Ω of overtime available each week. Therefore, we would evaluate the overtime available (and the overtime really used) to guarantee the same performance before the introduction of the emergency patient flow as suggested in [24]. The quantity of overtime available can be interpreted as the hours available of a dedicated operating room for emergency surgeries.

Table 9 reports about such tests. The first column reports the extra overtime available each week, that is the number of overtime hours added to the initial overtime of five hours (see Table A.15), that is one hour per day. The last two columns reports the overtime actually used and its percentage with respect to the total overtime available, respectively.

Table 9: Overtime estimation

	Elective							Emergency			overtime used	
Ω	C	$I_{\mathbf{avg}}$ ($t_{\mathbf{avg}}$)	$U_{\mathbf{bed}}$	$U_{\mathbf{OR}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$	$f_{\mathbf{MTBT}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$	$f_{\mathbf{MTBT}}$	minutes	%
0	370 (2271)	302 (50)	73.9%	94.4%	1.03	3.32	48.2%	0.68	3.59	72.9%	295	98.5%
5	333 (2290)	248 (41)	73.6%	95.1%	0.90	3.19	71.9%	0.69	3.59	72.1%	543	90.5%
10	318 (2352)	212 (35)	72.6%	96.8%	0.75	2.30	84.6%	0.70	3.53	71.6%	689	76.5%
15	291 (2231)	210 (35)	74.2%	92.7%	0.80	4.03	84.8%	0.67	3.61	73.2%	792	66.0%
20	273 (2250)	205 (33)	72.6%	93.6%	0.79	3.52	86.9%	0.69	3.43	72.1%	823	54.8%
25	260 (2361)	176 (29)	68.9%	97.5%	0.65	2.22	90.6%	0.71	3.72	72.4%	797	44.3%
30	244 (2397)	157 (26)	66.6%	98.6%	0.59	1.85	93.7%	0.73	3.65	70.8%	785	37.4%
35	233 (2403)	166 (28)	66.8%	98.9%	0.61	1.88	93.6%	0.72	3.68	71.7%	811	33.8%
40	209 (2424)	164 (27)	64.5%	99.5%	0.61	1.83	92.8%	0.74	3.71	70.7%	791	29.3%
45	205 (2422)	140 (23)	64.3%	99.5%	0.54	1.82	95.2%	0.73	3.63	70.9%	806	26.9%
50	194 (2413)	136 (23)	62.9%	99.5%	0.53	1.83	96.0%	0.72	3.67	71.3%	817	24.8%
55	183 (2406)	127 (21)	63.5%	99.1%	0.50	1.94	96.8%	0.74	3.74	70.1%	817	22.7%

The first remark is concerned with the overtime percentage effectively used, which decreases as soon as the number of hours weekly available increases. On the other side, it seems that about 800 minutes of overtime are those really used to deal with the emergency surgery flow under the second scenario.

The available overtime seems the more influencing factor. Actually, we can reach about the 90% of elective patients operated within their MTBT by making available 25 hours of overtime but using only the 44.3%. On a schedule of five days, the 25 hours of overtime can correspond to the availability for five hour each day of one dedicated operating room for emergency surgery. Therefore, our results suggest that it could be better to avoid the use of dedicated operating room for emergency surgery and to use those resources in a more flexible way within the elective patient flow as suggested in [41]. Furthermore, it seems convenient also from a budget point of view due to the fact since the overtime really used is a fraction of the whole available. Finally, we can observe that the overtime can be reduced adopting ad-hoc scheduling as shown in [42].

4.6. Dealing with differently trained surgical teams

A surgical team is a set of experts who perform surgery activities and related tasks together usually including surgeons, assistants, nurses, anaesthetists and surgical technologists. Such roles require a long period of training to be specialized (especially for surgeons and anaesthetists), with a significant impact on the variability of surgery duration.

Even if our focus is at the operational level to deal with the resource management, we would provide an evaluation of having surgical teams with different level of training. We suppose that a surgical team having less trained components could require additional time to accomplish their tasks.

The additional time added to the ROT is generated through an exponential distribution of parameter $\lambda_j > 0$ set to have average delay $\frac{1}{\lambda_j}$ (minutes). The exponential distribution has been chosen because such a probability function is positive and quickly decreasing.

We considered the new scenarios 1b and 2b obtained from the original one simply adding the value for parameter $\frac{1}{\lambda_j}$, as reported in Table 10a and 10b.

Table 10: Scenarios with additional delay

(a) Scenario 1b

	OR 1	OR 2	OR 3	OR 4	OR 5
$\frac{1}{\lambda_j}$	best	5	10	15	20

(b) Scenario 2b

	OR 1	OR 2	OR 3
$\frac{1}{\lambda_j}$	best	10	20

Tables 11 and 12 show the results for the new scenarios corresponding to the most representative configurations. Comparing them with those obtained for the baseline configurations, adding further delay causes a substantial deterioration of the performance indices, specially in correspondence of cancellations and number of patients operated on according to their MTBT.

Table 11: Performance indices of scenario 1b for best configurations in Table 5

id	C	$I_{\mathbf{avg}}$ ($t_{\mathbf{avg}}$)	$U_{\mathbf{bed}}$	$U_{\mathbf{OR}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$	$f_{\mathbf{MTBT}}$
(0)	378 (2162)	597 (93)	62.7%	87.5%	1.9	6.2	10.9%
(8)	281 (2298)	401 (64)	60.5%	93.5%	1.4	3.3	18.0%
(11)	356 (2345)	319 (52)	63.5%	95.4%	1.1	3.0	38.1%
(12)	368 (2342)	311 (51)	63.7%	95.1%	1.0	2.4	43.7%

Table 12: Performance indices of scenario 2b for best configurations in Table 6

id	C	$I_{\mathbf{avg}}$ ($t_{\mathbf{avg}}$)	$U_{\mathbf{bed}}$	$U_{\mathbf{OR}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$	$f_{\mathbf{MTBT}}$
(0)	344 (2248)	496 (78)	63.5%	90.6%	1.65	5.27	9.4%
(8)	266 (2354)	310 (50)	62.6%	95.0%	1.09	3.03	43.1%
(11)	332 (2400)	240 (39)	64.1%	96.5%	0.83	2.32	75.6%
(12)	350 (2396)	245 (40)	64.6%	96.4%	0.84	2.24	75.7%
(13)	341 (2375)	231 (38)	68.3%	95.6%	0.81	2.37	78.2%
(14)	355 (2386)	245 (40)	67.7%	96.3%	0.84	2.30	73.4%

The positive impact of the optimization persists: actually, the value of f_{MTBT} ranges from 10.9% to 43.7% in the scenario 1a, and from 9.4% to 73.4% in the scenario 1b. On the contrary, the value of C is almost stable, probably due to the nature of the new delays.

As already done in Section 4.5, we evaluate the additional number of overtime hours Ω necessary to regaining roughly the same overall performance that had been obtained for scenario 2. Results are reported in Table 13.

The results suggest that 15 hours of additional overtime per week would be enough to bring the performance indices to the best values of the scenario 2. Further, we observe that the overtime actually used is about the same of the scenario 2 (304 vs. 295 minutes).

Table 13: Overtime estimation

	Elective							overtime used	
Ω	C	$I_{\mathbf{avg}}$ ($t_{\mathbf{avg}}$)	$U_{\mathbf{bed}}$	$U_{\mathbf{OR}}$	$w_{\mathbf{avg}}$	$w_{\mathbf{max}}$	$f_{\mathbf{MTBT}}$	minutes	%
0	341 (2375)	231 (38)	68.3%	95.6%	0.81	2.37	78.2%	225	74.8%
5	282 (2410)	168 (28)	64.7%	96.8%	0.62	1.88	92.9%	256	42.7%
10	227 (2411)	146 (24)	62.6%	97.2%	0.56	1.82	95.0%	281	31.2%
15	169 (2414)	119 (20)	60.3%	97.4%	0.48	1.81	96.8%	304	25.3%

5. Conclusions

In this paper we proposed a model for the Real Time Management of operating rooms. Given an OR schedule, it consists in a sort of centralized surveillance system whose main task is to supervise the execution of such a schedule and, in the case of delays, to take the more rational decision regarding the surgery cancellation or the overtime assignment. We evaluated its impact on the performance of a generic surgical clinical pathway for elective patients. To this end, we developed a hybrid simulation and optimization model.

The extensive quantitative analysis discussed in Section 4 showed the positive impact of the optimization in the management of a surgical pathway through the evaluation of a set of patient-centred and facility-centred indices.

The online algorithm developed for the RTM is capable to determine a general improvement of all the performance indices. Comparing the baseline configuration with the best configuration in the two scenarios considered, we observed a vast improvement of the performance indices related to the waiting list in terms of its length and the waiting time. This allow to almost double the percentage of the patients operated on before their MTBT time limit. These improvements can determine a general improvement of the quality of service from a patient-centred point of view without deteriorating the facility-centred performance indices. The quantitative analysis confirms the trade-off between the number of cancellations and the number of operated patients (or, equivalently, the OR session utilization) as discussed in [17].

The analysis provided in Section 4.5 and 4.6 demonstrate the capability and the flexibility of our hybrid model to deal with different OR settings. These analysis also showed how the overtime could be interpreted as a really flexible resources that can be used to bring under control challenging situations.

From an OR management point of view, the quality of the provided results and the low computation time suggest the development of a decision support system based on the online algorithm for the RTM powered by an ICT infrastructure to track the surgeries within the operating rooms. Such a system could support the OR supervisor(s) in the management of the current schedule optimizing the use of the overtime.

Future research avenues could consider a more systematic analysis of the management of a joint elective and emergency patient flow. Another avenue could be the extension of the proposed approach to the analysis of multiple surgical pathways and their shared resources.

Furthermore, in our approach we are not directly taking into account the physician availability (e.g., surgeons and anesthetists) as is common when dealing with surgery process scheduling. Therefore, a future research avenue could include the availability of the physicians by introducing an optimization module dealing with the Master Surgical Schedule problem.

Acknowledgements

The authors acknowledge support from the Italian Ministry of Education, University and Research (MIUR), under the grant n. RBFR08IKSB, “Firb – Futuro in Ricerca 2008”. The authors wish to thank Elena Tànfani for fruitful discussions, which gave rise to this work whose preliminary results are reported in [43]. Finally, the authors wish to thank the anonymous referees for their valuable comments.

Appendix A. Parameters

In this appendix, we report the parameters of the simulation model and its setting both for the model validation (Section 4.2) and for the quantitative analysis (Sections 4.3–4.6). In brackets, the unit of measure.

Flow and patient characteristics:

- r_0**:** patient interarrival rate [patients/minutes],
- R_0**:** initial length of the pre-admission waiting list [patients],
- p_1**:** patient probability to require a surgical treatment during the ambulatory visit (see Fig. 1),
- p_2**:** patient probability to do not require a surgical treatment but requiring further exams during the ambulatory visit (see Fig. 1),
- $p_{\mathbf{A}}, \ldots, p_{\mathbf{G}}$**:** urgency class A,...,G patient probability.

Duration of activities:

$T^{min,avg,mod}_{\mathbf{A},\ldots,\mathbf{F},\mathbf{I}}$**:** minimum, average and modal time for the execution of A,. . .,F,I [minutes] (see Figures 1–2b),

$\ell_{\mathbf{min,\ max,\ mod;\ A,\ldots,G}}$**:** minimum, maximum and modal LOS for patients belonging to the urgency class A,. . .,G [days],

$\bar{\epsilon}_{\mathbf{A},\ldots,\mathbf{G}}$**:** average EOT for the surgery of a patient belonging to the urgency class A,. . .,G [minutes],

$e_{\mathbf{max}}$**:** maximum duration of a surgery [minutes],

$\sigma_{\mathbf{A},\ldots,\mathbf{G}}$**:** EOT standard deviation [minutes],

σ**:** ROT standard deviation for each patient [minutes],

τ**:** tolerance time within which the surgical team operates a patient at the end of OR session without resorting to the overtime [minutes].

Table A.14 shows the distributions used to generate the required time for the execution of the activities A,. . .,J. Table A.15 reports the values assigned to the parameters for the model validation and for the quantitative analysis.

Table A.14: Distribution of the activity durations

Activities	Durations	Parameters
A,. . ., F, I	$T^{\mathrm{A},\ldots,\mathrm{F,I}}_{\min} + T$, $T \sim \mathrm{Gamma}(k, \vartheta)$	$k = T^{\mathrm{A},\ldots,\mathrm{F,I}}_{avg} - T^{\mathrm{A},\ldots,\mathrm{F,I}}_{mod}$, $\vartheta = \frac{T^{\mathrm{A},\ldots,\mathrm{F,I}}_{avg} - T^{\mathrm{A},\ldots,\mathrm{F,I}}_{min}}{T^{\mathrm{A},\ldots,\mathrm{F,I}}_{avg} - T^{\mathrm{A},\ldots,\mathrm{F,I}}_{mod}}$
H (LOS)	$\lfloor \mathrm{Triangular}(l_{\mathrm{min;\ A,\ldots,G}}, l_{\mathrm{max;\ A,\ldots,G}}, l_{\mathrm{mod;\ A,\ldots,G}}) + \frac{1}{2} \rfloor$	
J (EOT)	$\min\left\{\max\left\{\lfloor \frac{T}{u} + \frac{1}{2} \rfloor u, 0\right\}, e_{\max}\right\}$, $T \sim \mathrm{Lognormal}(\mu, s^2)$	$\mu = \log \epsilon_{\mathrm{A,\ldots,G}} - \frac{1}{2}\log\left(\frac{\sigma^2_{A,\ldots,G}}{\epsilon^2_{A,\ldots,G}} + 1\right)$, $s = \sqrt{\log\left(\frac{\sigma^2_{A,\ldots,G}}{\epsilon^2_{A,\ldots,G}} + 1\right)}$
J (ROT)	$\min\{\max\{0, T\}, e_{\max}\}$, $T \sim \mathrm{Gaussian}(\mathrm{EOT}, \sigma^2)$	

Starting from the values reported in [32], that is minimum, maximum, average and modal values, we use a Gamma distribution because, empirically, those values suggested a distribution whose shape recalls the Gamma. The parameters k and ϑ were obtained in such a way to equal the expected and the modal values reported in [32]. Further, we compute the value of the survival function on the maximum time for the execution of activities (always reported in the paper), obtaining a value less than 10%.

The EOT of the patient i represents a prediction of the surgery duration performed by the surgeons at the moment of the ambulatory visit who indicates the mean duration of similar surgeries on the basis of the own personal experience (in absence of historical data). In the literature, *a priori*

Table A.15: Parameters used in the simulation framework

Parameters	unit of measure	Validation	Quantitative analysis
r_0	patients/minutes	$5.8 \cdot 10^{-3}$	$2.0 \cdot 10^{-2}$
R_0	patients	140	420
p_1, p_2		0.2, 0.1	0.2, 0.1
$p_{\mathrm{A}}, \ldots, p_{\mathrm{G}}$		$0.0245, 0.1401, 0.4136, 0.1785$ $0.1140, 0.0749, 0.0544$	$0.0245, 0.1401, 0.4136, 0.1785$ $0.1140, 0.0749, 0.0544$
$T_{min}^{\mathrm{A},\ldots,\mathrm{F},\mathrm{I}}$	minutes	$5, 25, 25, 25, 40, 25, 35$	$5, 25, 25, 25, 40, 25, 35$
$T_{avg}^{\mathrm{A},\ldots,\mathrm{F},\mathrm{I}}$	minutes	$7.5, 31.5, 31, 28, 62.5, 32, 41$	$7.5, 31.5, 31, 28, 62.5, 32, 41$
$T_{mod}^{\mathrm{A},\ldots,\mathrm{F},\mathrm{I}}$	minutes	$6, 30, 26, 25, 50, 30, 40$	$6, 30, 26, 25, 50, 30, 40$
$\ell_{\mathrm{min;\ A},\ldots,\mathrm{G}}$	days	$2, 1, 1, 1, 1, 1, 1$	$2, 1, 1, 1, 1, 1, 1$
$\ell_{\mathrm{max;\ A},\ldots,\mathrm{G}}$	days	$29, 16, 7, 9, 5, 5, 5$	$29, 16, 7, 9, 5, 5, 5$
$\ell_{\mathrm{mod;\ A},\ldots,\mathrm{G}}$	days	$3, 2, 2, 2, 2, 2, 2$	$3, 2, 2, 2, 2, 2, 2$
$e_{\max}$	minutes	360	420
$\bar{\epsilon}_{\mathrm{A},\ldots,\mathrm{G}}$	minutes	$145, 171, 149, 153, 171, 164, 166$	$145, 171, 149, 153, 171, 164, 166$
$\sigma_{\mathrm{A},\ldots,\mathrm{G}}$	minutes	$85, 85, 66, 60, 61, 51, 60$	$85, 85, 66, 60, 61, 51, 60$
σ	minutes	0	30
τ	minutes	30	10
Ω	minutes	0	300
u	minutes	30	30
$B_1, \ldots, B_7$	beds	$18, 18, 18, 18, 18, 18, 18$	$50, 50, 50, 50, 50, 35, 35$
$M_{\mathrm{URG\ A},\ldots,\mathrm{URG\ G}}$	days	$8, 15, 30, 60, 90, 120, 180$	$8, 15, 30, 60, 90, 120, 180$

surgery duration generally follows a Log-Normal distributions (see, e.g.,[44–46]). Then, the ROT of the patient i has been generated in such a way to replicate the uncertainty pertaining the prediction made by the surgeons: we generate a value ν using a Gaussian distributions with average 0 and standard deviation σ; then, the ROT value r_i is computed as $e_i + \nu$.

We observe that the simulation model generates activity durations on the basis of few information reported in [32]. In presence of historical data about surgery durations, it is more reasonable to use them for replicating and predicting surgery durations depending on the characteristics of the patient, learning from the previous experiences. For this purpose several Bayesian methods can be used as reported in the literature (see, e.g.,[47–49]).

References

[1] A. Testi, E. Tànfani, G. Torre, A three-phase approach for operating theatre schedules, Health Care Management Science 10 (2007) 163–172.

[2] J. Magerlein, J. Martin, Surgical demand scheduling: A review, Health Services Research 13 (1978) 418–433.

[3] I. Ozkarahan, Allocation of surgeries to operating rooms using goal programming, Journal of Medical Systems 24 (6) (2000) 339–378.

[4] A. Guinet, S. Chaabane, Operating theatre planning, International Journal of Production Economics 85 (2003) 69–81.

[5] M. Lamiri, X. Xie, A. Dolgui, F. Grimaud, A stochastic model for operating room planning with elective and emergency demand for surgery, Journal of Operational Research Society 185 (2008) 1026–1037.

[6] H. Fei, C. Chu, N. Meskens, A. Artiba, Solving surgical cases assignment problem by a branch-and-price approach, International Journal of Production Economics 112 (2008) 96–108.

[7] E. Hans, G. Wullink, M. van Houdenhoven, G. Kamezier, Robust surgery loading, European Journal of Operational Research 185 (2008) 1038–1050.

[8] I. Marques, M. Captivo, M. Pato, An integer programming approach to elective surgery scheduling, OR Spectrum 34 (2012) 407–427.

[9] C. Rizk, J. Arnaout, ACO for the surgical cases assignment problem, Journal of Medical Systems 36 (2012) 1191–1199.

[10] R. Aringhieri, P. Landa, E. Tanfani, Assigning surgery cases to operating rooms: A VNS approach for leveling ward beds occupancies, in: The 3rd International Conference on Variable Neighborhood Search (VNS'14), Electronic Notes in Discrete Mathematics, 2015, pp. 173–180.

[11] D. Pham, A. Klinkert, Surgical case scheduling as a generalized job shop scheduling problem, European Journal of Operational Research 185 (2008) 1011–1025.

[12] B. Cardoen, E. Demeulemeester, J. Beliën, Sequencing surgical cases in a day-care environment: An exact branch-and-price approach, Computers & Operations Research 36 (9) (2009) 2660–2669.

[13] B. Roland, C. Martinelly, F. Riane, Y. Pochet, Scheduling an operating theatre under human resource constraints, Computers & Industrial Engineering 58 (2010) 212–220.

[14] A. Riise, E. Burke, Local search for the surgery admission planning problem., Journal of Heuristics 17 (4) (2011) 389–414.

[15] W. Herring, J. Herrmann, The single-day surgery scheduling problem: sequential decision-making and threshold-based heuristics, OR Spectrum 34 (2012) 429–459.

[16] N. Meskens, D. Duvivier, A. Hanset, Multi-objective operating room scheduling considering desiderata of the surgical teams, Decision Support Systems 55 (2013) 650–659.

[17] I. Beaulieu, M. Gendreau, P. Soriano, Operating rooms scheduling under uncertainty, in: E. Tànfani, A. Testi (Eds.), Advanced Decision Making Methods Applied to Health Care, Vol. 173 of International Series in Operations Research & Management Science, Springer Milan, 2012, pp. 13–32.

[18] B. Cardoen, E. Demeulemeester, J. Beliën, Operating room planning and scheduling: A literature review, European Journal of Operational Research 201 (2010) 921–932.

[19] F. Guerriero, R. Guido, Operational research in the management of the operating theatre: a survey, Health Care Management Science 14 (2011) 89–114.

[20] E. W. Hans, P. T. Vanberkel, Operating theatre planning and scheduling, in: R. Hall (Ed.), Handbook of Healthcare System Scheduling, Vol. 168 of International Series in Operations Research & Management Science, Springer US, 2012, pp. 105–130.

[21] F. Dexter, A. Willemsen-Dunlap, J. Lee, Operating room managerial decision-making on the day of surgery with and without computer recommendations and status displays, Anesthesia and Analgesia 105 (2) (2007) 419–429.

[22] R. Wachtel, F. Dexter, Reducing tardiness from scheduled start times by making adjustments to the operating room schedule, Anesthesia and Analgesia 108 (6) (2009) 1902–1909.

[23] E. Erdem, X. Qu, J. Shi, S. Upadhyaya, A mathematical modeling approach for rescheduling elective admissions upon arrival of emergency patients, in: 61st Annual IIE Conference and Expo Proceedings, 2011.

[24] E. Erdem, X. Qu, J. Shi, Rescheduling of elective patients upon the arrival of emergency patients, Decision Support Systems 54 (1) (2012) 551–563.

[25] D.-N. Pham, A. Klinkert, Surgical case scheduling as a generalized job shop scheduling problem, European Journal of Operational Research 185 (3) (2008) 1011–1025.

[26] K. Stuart, E. Kozan, Reactive scheduling model for the operating theatre, Flexible Services and Manufacturing Journal 24 (4) (2012) 400–421.

[27] F. Dexter, A strategy to decide whether to move the last case of the day in an operating room to another empty operating room to decrease overtime labor costs, Anesthesia and Analgesia 91 (4) (2000) 925–928.

[28] F. Dexter, T. C. Smith, D. J. Tatman, A. Macario, Physicians' perceptions of minimum time that should be saved to move a surgical case from one operating room to another: Internet-based survey of the membership of the Association of Anesthesia Clinical Directors (AACD), Journal of clinical anesthesia 15 (3) (2003) 206–210.

[29] V. Tiwari, F. Dexter, B. Rothman, J. Ehrenfeld, R. Epstein, Explanation for the near constant mean time remaining in surgical cases exceeding their estimated duration, necessary for appropriate display on electronic white boards, Anesthesia and Analgesia 117 (2013) 487–493.

[30] H. Campbell, N. Bradshaw, M. Porteous, Integrated care pathways, British Medical Journal 316 (133-144).

[31] R. Aringhieri, E. Tànfani, A. Testi, Operations Research for Health Care Delivery, Computers & Operations Research 40 (9) (2013) 2165–2166.

[32] Y. Ozcan, E. Tànfani, A. Testi, A simulation-based modeling framework to deal woth clinical pathways, in: S. Jain, R. Creasey, J. Himmelspach, K. White, M. Fu (Eds.), Proceedings of the 2011 Winter Simulation Conference, 2011, pp. 1190–1201.

[33] E. Tànfani, A. Testi, A pre-assignment heuristic algorithm for the master surgical schedule problem (mssp), Annals of Operations Research 178 (1) (2010) 105–119.

[34] M. Lamiri, X. Xie, A. Dolgui, F. Grimaud, Optimization methods for a stochastic surgery planning problem, International Journal of Production Economics 120 (2009) 400–410.

[35] W. Herring, J. Herrmann, A stochastic dynamic program for the single-day surgery scheduling problem, IIE Transactions on Healthcare Systems Engineering 4 (2011) 213–225.

[36] E. Tànfani, A. Testi, R. Alvarez, Operating room planning considering stochastic surgery durations, International Journal of Health Management and Information 1 (2) (2010) 167–183.

[37] A. Borshchev, The Big Book of Simulation Modeling. Multimethod Modeling with AnyLogic 6, 2013, iSBN 978-0-9895731-7-7.

[38] R. Aringhieri, P. Landa, P. Soriano, E. Tànfani, A. Testi, A two level Metaheuristic for the Operating Room Scheduling and Assignment Problem, Computers & Operations Research 54 (2015) 21–34.

[39] S. Martello, P. Toth, Knapsack Problems: Algorithms and Computer Implementations, Wiley-Intersci. Ser. Discrete Math. Optim., John Wiley and Sons, 1990.

[40] M. Heng, J. G. Wright, Dedicated operating room for emergency surgery improves access and efficiency, Canadian Journal of Surgery 56 (3) (2013) 167–174.

[41] G. Wullink, M. Van Houdenhoven, E. W. Hans, J. M. Van Oostrum, M. Van Der Lans, G. Kazemier, Closing emergency operating rooms improves efficiency, Journal of medical systems 31 (6) (2007) 543–546.

[42] J. van Essen, E. Hans, J. Hurink, A. Oversberg, Minimizing the waiting time for emergency surgery, Operations Research for Health Care 1 (23) (2012) 34 – 44.

[43] R. Aringhieri, D. Duma, A hybrid model for the analysis of a surgical pathway, in: Proceedings of the 4th International Conference on Simulation and Modeling Methodologies, Technologies and Applications (HA-2014), 2014, pp. 889–900.

[44] D. Strum, J. May, L. Vargas, Modeling the uncertainty of surgical procedure times: Comparison of lognormal and normal models, Anesthesiology 92 (4) (2000) 1160–1167.

[45] J. May, D. Strum, L. Vargas, Fitting the lognormal distribution to surgical procedure times, Decision Sciences 31 (1) (2000) 129–148.

[46] W. Spangler, D. Strum, L. Vargas, H. Jerrold, Estimating procedure times for surgeries by determining location parameters for the lognormal model, Health Care Management Science 7 (2004) 97–104.

[47] F. Dexter, J. Ledolter, Bayesian prediction bounds and comparisons of operating room times even for procedures with few or no historic data, Anesthesiology 103 (6) (2005) 1259–1267.

[48] F. Dexter, R. Epstein, E. Bayman, J. Ledolter, Estimating surgical case durations and making comparisons among facilities: Identifying facilities with lower anesthesia professional fees, Anesthesia and Analgesia 116 (5) (2013) 1103–1115.

[49] F. Dexter, J. Ledolter, V. Tiwari, R. Epstein, Value of a scheduled duration quantified in terms of equivalent numbers of historical cases, Anesthesia and Analgesia 117 (1) (2013) 205–210.