Next-Day Operating Room Scheduling with Time-Dependent Stochastic Surgery Durations

Enis Kayış, Tuğçe Karataş, and Refik Güllü

Abstract Operating rooms (ORs) are the most costly part of hospitals, thus a priority for hospital administrations. In this paper, we consider the next-day OR scheduling problem for multiple operating rooms. We assume that surgeries have uncertain durations, and distributions of surgery durations are time-dependent. Our aim is to find the assignment of surgeries to the available ORs, the sequence, and the planned starting times of surgeries in order to minimize the weighted sum of expected waiting time of patients, idle time of ORs, and overtime of the hospital staff. In order to find solutions to the problem, we propose an L-Shaped method, customized to our problem formulation. We quantify the penalty of ignoring the time-dependency of surgery durations within a numerical study. We find that the penalty of ignoring the time-dependency increases with the overtime cost, average surgery durations, and decreases with surgery variability.

Keywords Operating rooms · OR scheduling · L-shaped method

Introduction

Healthcare expenditures are increasing with an ever-increasing rate [14]. According to [7], the proportion of healthcare expenditures in total GDP in the USA increased from 5.0% in 1960 to 17.8% in 2015. Hospital expenditures are responsible for a third of this spending. To reduce costs, many hospitals have been looking for ways to increase the efficient use of their resources. According to the Health Care Financial

E. Kayış (✉)
Industrial Engineering Department, Ozyegin University, Istanbul, Turkey
e-mail: enis.kayis@ozyegin.edu.tr

T. Karataş
Industrial Engineering and Operations Research Department, Columbia University, New York, NY, USA

R. Güllü
Industrial Engineering Department, Bogazici University, Istanbul, Turkey
e-mail: refik.gullu@boun.edu.tr

Management Association, operating rooms (ORs) account for more than 40% of a hospital's revenues and costs [18]. Therefore, it is crucial to increase the efficiency of ORs in order to better manage the increasing hospital costs.

OR planning and scheduling problems are especially challenging. First, many factors, including the ability of hospital staff, type of surgeries, and possible complications encountered during surgeries, result in the inherent uncertainty of surgery durations. Yet, the accuracy of surgery duration predictions is critical: If surgeries last shorter than predicted, ORs remain idle, and utilization of ORs decreases. If surgeries last longer than predicted, patients will wait for their surgeries, and this will result in dissatisfaction of patients. Moreover, the computational burden of the problem increases drastically as the number of ORs and thus the number of the surgeries increase, since there are too many possible assignments of surgeries to ORs and sequences of surgeries within ORs.

OR planning problems are divided into three categories in [1, 17]: strategic, tactical, and operational. Strategic level problems have long-term planning horizons. The long-term capacity of operating rooms and the allocation of operating rooms over different specialties are determined at the strategic level. The temporary capacity of operating rooms and the allocation of surgeries to different days in a week are determined at the tactical level. At the operational level, daily schedules of operating rooms are determined. Our research problem lies in this last category: What is the best next-day OR schedule for a list of surgeries given the time-dependent stochastic surgery durations?

In this chapter, we consider a next-day OR scheduling problem for multiple ORs. Emergent surgeries are ignored, and only elective surgeries are considered. It is assumed that surgeries have uncertain durations, and distributions of surgery durations are time-dependent. Furthermore, it is assumed that a surgery cannot start before its planned time, even if the previous surgery is completed earlier than its planned time. On the other hand, it may start later than the planned time if the previous surgery is completed after the planned time. The objective is to minimize the expected weighted sum of waiting time of patients, idle time of operating rooms, and overtime of the hospital staff. We provide a mathematical model formulation of the problem, and the L-Shaped method is used to solve the resulting formulation.

The vast literature on OR planning assumes that the duration of surgery does not depend on the scheduled time of the day. However, recent studies show that the performance of hospital staff (surgeons, nurses, anesthesiologists, technicians) could change according to the time of the day (e.g., [19, 19]). Most of the hospital staff are energetic in the morning, and this may result in a decrease in surgery durations of early surgeries. Moreover, hospital staff would not want to stay beyond planned working hours and be more effective in parallel processing and communication with other team members, which translates into reduced durations for late surgeries. We consider time-dependent surgery durations and quantify the effect of ignoring the time-dependency of stochastic surgery durations in OR planning.

A related avenue of research is scheduling papers that consider sequence-dependent setup times. In these papers, job duration is assumed to depend on the

preceding job(s). In our study, we model time dependency such that the distribution of surgery duration depends on whether it is the first, middle, or last scheduled surgery. Apart from [35], which study the deterministic OR scheduling problem with sequence-dependent setup times, our paper is the first study that considers time-dependent stochastic surgery durations in next-day OR scheduling.

Literature Review

OR planning and scheduling problems have drawn great attention during the last 50 years. For a comprehensive review, we refer the reader to [6, 6]. In this section, we briefly discuss papers on stochastic OR planning grouped into two categories: papers that study only scheduling and those that consider both sequencing and scheduling of surgeries. Then, we highlight papers on machine scheduling with sequence-dependent setup times, which is a related stream for our study.

Papers in the first category assume that sequence of surgeries is known at the beginning and aims to find the optimally planned surgery durations. Reference [33] finds that the problem is of a "Newsvendor" type when there are two surgeries. If the number of surgeries is more, a myopic heuristic is proposed. Reference [31] aims to minimize the weighted customer delay time and system completion time. Under exponentially and independently distributed surgery durations, it proves the convexity of the objective function. It is argued that the optimal durations exhibit a dome-shaped pattern: the planned surgery durations first increase and then decrease as one moves toward the later surgeries in the sequence. Reference [8] formulate the problem as a two-stage stochastic linear program. It finds the lower and upper bounds of the objective function and then employs the L-Shaped method with sequential bounding to solve this problem. Reference [4] study the case where the surgery durations are distributed with a joint discrete probability function. It is shown that the optimal surgery durations can be calculated in polynomial time. Since the exact solution is hard to find, some papers focus on developing effective heuristics. Reference [29] propose a robust heuristic to minimize the weighted sum of the waiting time of the patients and the idle time of surgeons. Finally, [21] develop a method to find the exact solution to the problem and propose a hybrid heuristic motivated by real practices.

The problem of finding the optimal sequence as well as the planned starting times of surgeries within a single OR, is even more challenging. Reference [32] proposes a sequential two-phase method in which first the sequence, then the planned starting times are calculated. It finds that under exponentially distributed service times, sequencing jobs in order of decreasing means gives the optimal sequence. Reference [12] consider elective case scheduling to maximize the utilization of the operating rooms. Two heuristics for scheduling of surgeries are used: Earliest Start Time Heuristic and Latest Start Time Heuristic. The first heuristic is good at predicting start time, and the second heuristic can eliminate overtime. Reference [25] studies OR scheduling with the aims of maximizing the throughput and minimizing

waiting time and overtime. It is argued that short procedures contain less variability than long procedures. Thus, scheduling short-duration procedures first maximize on-time performance. A two-stage stochastic programming model is presented in [10] to sequence surgeries and schedule surgery durations simultaneously. The paper employs the Sample Average Approximation (SAA) method to find solutions using four different sequencing heuristics. It is concluded that sequencing surgeries using Shortest-Variance-First (SVF) rule outperform other heuristics. Reference [28] prove that the finite scenario SAA method is NP-Complete. They develop a heuristic method based on Bender's decomposition, and the results are compared with the solutions found by the SAA method. It discusses that this heuristic has significantly better results than SVF heuristic when the unit costs are unequal. Reference [2] studies a single machine stochastic scheduling problem and finds that when the job processing times are distributed normally, and earliness and tardiness costs are the same for each job, the optimal sequence is SVF. Reference [15] extends this result and proves that the SVF rule is optimal under the assumption of dilation ordering of processing durations.

When there are multiple available ORs, another dimension of the assignment of surgeries to the ORs is introduced into an already challenging problem. Reference [9] develop a simulation model to handle uncertainties related to the intake process, surgical procedure, and recovery process. A simple simulated annealing method is used in order to improve the patient arrival schedule. Reference [22] develop a stochastic model for OR scheduling with elective and emergent surgeries under the assumption that surgery durations of elective surgeries are known and deterministic. A Monte Carlo optimization is proposed that combines Monte Carlo simulation and mixed-integer program. In this model, elective surgeries are assigned into different periods over the planning horizon in order to minimize the sum of elective patient-related costs and overtime cost of operating rooms. Reference [23] assume that operating rooms are identical, and the SAA method is used as a solution method. They find that the Monte Carlo optimization method converges exponentially to a real optimal solution. Reference [11] develop two methods to minimize the total sum of the fixed cost of opening operating rooms and the variable cost of overtime relative to a fixed length of a day. The first method is a two-stage stochastic linear program with binary decisions in the first stage, and simple recourse in the second stage. The binary decisions in the first stage are the number of operations to be opened, and the assignment of surgeries to operating rooms. The second method tries to minimize the maximum cost associated with the uncertainty of surgery durations. The paper finds that the second method is faster, and it benefits from limiting the worst-case outcome of the recourse problem. Reference [3] employ the L-Shaped method for multiple operating rooms and multiple surgeries problem. In the first stage, the number of operating rooms to be opened, the assignment of surgeries into operating rooms, and the sequence of the surgeries within ORs are determined. Therefore, this stage only includes binary variables. In the second stage, the weighted sum of the waiting time, idle time, and overtime are calculated by using the values of binary variables obtained in the first stage. Finally, [35] develop a method for solving a problem with sequence-dependent setup times of surgeries. The goal is to decide on the number

of operating rooms to open, the assignment of surgeries to operating rooms and the sequence of surgeries within an operating room. It is assumed that surgeries belong to different types, each operating room allows only a set of surgery types and setup time, and surgery durations are deterministic. It is found that the Constraint Programming model gives more efficient results in terms of computation time and solution quality than the Mixed Integer Nonlinear Programming model.

There are several studies on machine scheduling with sequence-dependent setup times that are relevant to our problem. As in [35], the majority of studies on machine scheduling problems with sequence-dependent setup times focus on deterministic setup times due to the computational burden of stochastic problems. Reference [16] proposes an assignment algorithm, which is an extension of the Hungarian algorithm for scheduling jobs when there exist many identical parallel machines. Reference [26] propose a three-phase heuristic for scheduling of jobs when there exists only one machine. It is assumed that setup times are sequence-dependent, and all jobs are available for processing at time zero. In the first phase, several parameters that characterize the problem are calculated. By using these parameters, the schedule of jobs is found based on a priority rule in the second phase. In the last phase, the schedule obtained in the second phase is improved by using a local improvement procedure. Reference [27] extends the same problem to many parallel and identical machines. The same heuristic is used for the solution of the problem, and the simulated annealing method is applied in the third phase.

In conclusion, there are several studies on the scheduling and sequencing of the surgeries. It is nearly impossible to solve these problems exactly, and different heuristics are developed in order to solve these problems. However, there is no study on the scheduling and sequencing of ORs under time-dependent stochastic surgery durations. The problem of finding the optimal schedule under time-dependent setup times have been studied assuming deterministic setup and surgery durations. We consider stochastic surgery durations, and the computational complexity increases significantly as the number of surgeries and the number of ORs increases.

Problem Formulation

We study a stochastic next-day OR planning problem: Given the list of the surgeries to be performed the next day, one must assign each surgery in this list an OR, a sequence within this OR, and a planned starting time. Let $S = \{1, \ldots, N\}$ be the set of surgeries that need to be scheduled for the next day. There are m identical ORs, and each OR k is assigned with n_k surgeries for $k = \{1, \ldots, m\}$. Trivially we have $\sum_{k=1}^{m} n_k = N$. We assume that the number of surgeries assigned to each OR is determined in advance by hospital management.

All surgeries are assumed to be elective. Surgery durations are independently and non-identically distributed. We measure the surgery duration from wheels-in to wheels-out, i.e., surgery duration includes setup time before the surgery, actual surgery duration, and clean-up time after the surgery. $T_{i,j,k}$ is the random duration of

surgery j, with a cumulative distribution function $F_{i,j}(x)$ and a probability density function $f_{i,j}(x)$, when it is assigned to i-th sequence in k-th OR. We assume that ORs are identical; hence surgery durations do not depend on the assigned OR. The first surgery of the OR starts without any waiting time, but surgeries later in the sequence should wait for the preceding surgery to be completed and the planned start time to arrive.

Our objective is to find a schedule that minimizes the expected weighted cost of waiting time of the patients, idle time of the operating rooms between two successive surgeries, and overtime of the ORs beyond planned opening times. We use weights for these cost components to reflect the fact that their priorities may not be equal. α_1 is defined as the unit cost of the idle time, α_2 is defined as the unit cost of the waiting time, and α_3 is defined as the unit cost of the overtime.

There are two types of decision variables in our model: continuous (planned surgery duration, waiting time, idle time and overtime) and binary (sequence and assignment of surgeries). $D_{i,k}$ is the assigned duration of i-th scheduled surgery in k-th OR. $w_{i,j,k}$ is the waiting time of the patient who is waiting for surgery j, which is assigned to i-th sequence in k-th OR. It takes positive value only if (i-1)-th scheduled surgery in k-th operating room lasts longer than its planned time. $s_{i,j,k}$ is the idle time of the OR right after i-th scheduled surgery in k-th OR. It is the time interval between the completion time of i-th scheduled surgery and the planned starting time of (i + 1)-th scheduled surgery in k-th OR. o_k is the excessive time over the total planned time of all surgeries assigned to k-th OR. The binary variable $x_{i,j,k}$ takes the value of 1 if surgery j is assigned to i-th sequence in k-th operating room.

Our problem could be formulated as a Stochastic Mixed-Integer Problem (SMIP). However, it is nearly impossible to find an exact solution. Instead, we will provide a formulation based on the SAA method, which is one of the most common approaches to solve SMIPs. SAA generates a number of scenarios (samples) for the random variables in the problem and estimates the objective function using the average weighted cost over these samples. In our context, a scenario is one particular sample path realization of all surgeries' random durations. Clearly, the approximation error reduces as the number of scenarios increases.

Some of the parameters and decision variables used in this method have an index p representing the scenario number. Let parameter $T_{i,j,k}^{p}$ be the random duration of surgery j in p-th scenario when it is assigned to i-th sequence in k-th OR. Since we assume that surgery durations are time-dependent, this parameter is sampled from the associated distribution with completing surgery j in the i-th sequence. $w_{i,j,k}^{p}$ is the waiting time of the patient who is waiting for surgery j, which is assigned to i-th sequence in k-th OR in p-th scenario. $s_{i,j,k}^{p}$ is the idle time of the operating room right after i-th scheduled surgery in k-th OR in p-th scenario. o_k^{p} is the excessive time over the total planned time of all surgeries assigned to k-th OR in p-th scenario.

One can formulate the problem to be solved with the SAA method as follows:

$$\text{Min} \quad \frac{1}{P} \sum_{p=1}^{P} \left(\sum_{k=1}^{m} \left(\sum_{j=1}^{n} \left(\sum_{i=1}^{n_k} \alpha_1 . s_{i,j,k}^{p} + \sum_{i=2}^{n_k} \alpha_2 . w_{i,j,k}^{p} \right) + \alpha_3 . o_k^{p} \right) \right) \tag{1}$$

s.t.

$$-\sum_{j=1}^{n} w_{i,j,k}^{p} + \sum_{j=1}^{n} w_{i+1,j,k}^{p} - \sum_{j=1}^{n} s_{i,j,k}^{p} + D_{i,k} = \sum_{j=1}^{n} T_{i,j,k}^{p} \cdot x_{i,j,k} \tag{2}$$

$i = 1, \ldots, n_k - 1, k = 1, \ldots, m$ and $p = 1, \ldots, P$

$$-\sum_{j=1}^{n} w_{n_k,j,k}^{p} - \sum_{j=1}^{n} s_{n_k,j,k}^{p} + D_{n_k,k} + o_k^{p} = \sum_{j=1}^{n} T_{n_k,j,k}^{p} \cdot x_{n_k,j,k}. \tag{3}$$

$k = 1, \ldots, m$ and $p = 1, \ldots, P$

$$\sum_{k=1}^{m} \sum_{i=1}^{n_k} x_{i,j,k} = 1 \quad j = 1, \ldots, N \tag{4}$$

$$\sum_{j=1}^{n} x_{i,j,k} = 1 \quad i = 1, \ldots, n_k \text{ and } k = 1, \ldots, m \tag{5}$$

$s_{i,j,k}^{p} \le M_1 \cdot x_{i,j,k} \quad i = 1, \ldots, n_k \text{ and } j = 1, \ldots, N \text{ and } k = 1, \ldots, m \text{ and } p = 1, \ldots, P$ \hfill (6)

$w_{i,j,k}^{p} \le M_2.x_{i,j,k} \quad i = 2, \ldots, n_k \text{ and } j = 1, \ldots, n, k = 1, \ldots, m \text{ and } p = 1, \ldots, P$ \quad (7)

$$\sum_{k=1}^{j} \sum_{i=1}^{n_k} x_{i,j,k} = 1 \quad j = 1, \ldots, m \tag{8}$$

$$\sum_{k=r}^{min\{j,m\}} \sum_{i=1}^{n_k} x_{i,j,k} - \sum_{a=r-1}^{j-1} \sum_{i=1}^{n_k} x_{i,a,r-1} \le 0 \quad j = 2, \ldots, N \text{ and } r = 2, \ldots, min(j,m) \tag{9}$$

$s_{i,j,k}^{p}, w_{i,j,k}^{p} \ge 0 \quad i = 2, \ldots, n_k \text{ and } j = 1, \ldots, N \text{ and } k = 1, \ldots, m \text{ and } p = 1, \ldots, P$ \hfill (10)

$$o_k^{p} \ge 0 \quad k = 1, \ldots, m \text{ and } p = 1, \ldots, P \tag{11}$$

$$D_{i,k} \ge 0 \quad i = 1, \ldots, n_k \text{ and } k = 1, \ldots, m \tag{12}$$

$$x_{i,j,k} \in \{0, 1\} \quad i = 1, \ldots, n_k \text{ and } j = 1, \ldots, N \text{ and } k = 1, \ldots, m \tag{13}$$

The objective function (1) is the weighted average cost of the waiting time of the patients, the idle time of the ORs, and the overtime of the surgeons. Constraints (2) and (3) provide the relationship between the waiting time of the patients, the idle time of the ORs, and the overtime of the surgeons. Constraints (4) and (5), respectively, ensure that surgery is assigned to exactly one sequence in exactly one OR, and each

sequence in each OR is assigned to exactly one surgery. Constraints (6) and (7) guarantee that the waiting time and the idle time variables are equal to 0 unless j-th surgery is assigned to i-th sequence in k-th OR. In these constraints, M indicates a large number. Constraints (10), (11), and (12) define the non-negativity restrictions of the waiting time, the idle time, the overtime, and the assigned duration variables, respectively. Constraint (13) is the integrality constraint.

Constraints (8) and (9) are symmetry breaking constraints: In terms of sequencing of surgeries, there are N! possible allocation of surgeries. However, it is possible to decrease the number of allocations that should be considered, since we assume that ORs are identical. This constraint set prevents obtaining the same solution again. Constraint (8) helps assigning the surgeries to operating rooms in the lexicographical order, since assigning surgeries lexicographically will give a feasible solution to the problem (see, for example, [30]). According to this constraint, surgery 1 should be assigned to OR 1. Surgery 2 can be assigned to the OR 1 or OR 2. By following this pattern, surgery $(m - 1)$ can be assigned to the first $(m - 1)$ ORs, and surgery m can be assigned to any OR. Constraint (9) prevents assigning surgeries to ORs that have a larger OR number index than the smallest OR number index of empty ORs (see, for example, [11]). For example, surgery 4 cannot be assigned to the third OR if there exists no surgery in the second OR.

Solution Methodology

As the number of operating rooms and the number of surgeries increase, the problem size increases drastically. For these cases, the SAA method cannot provide a solution within a reasonable computation time. As an alternative, we propose to solve our problem using the L-Shaped method (see, for example, [3], and [5, 24, 34]).

The L-shaped method computes a solution in two stages, defined as a master problem and a subproblem. In the master problem, the assignment of surgeries to ORs and their sequences within the OR are determined by solving a mixed-integer problem. The objective of this problem is to minimize the value of a variable, denoted by Θ, which is the expected recourse function. Θ is the expected value of the objective function of the subproblem given the sequences of surgeries in their assigned ORs. In the second stage, the total weighted sum of the waiting time of the patients for the surgeries, the idle time of the operating rooms, and the overtime of the ORs is minimized. A slightly modified version of the SAA model defined in the previous section is used to solve the subproblem. Since the values of the binary variables are assigned in the first stage, the second stage of the L-Shaped method is solving a simple linear programming problem. Figure 1 presents the overall algorithm of the L-Shaped method.

As an input, the algorithm requires $\alpha_1, \alpha_2, \alpha_3$ weights, the number of operating rooms (m), and the total number of surgeries (N). It is assumed that each operating room will have an equal number of surgeries. Furthermore, parameters of time-dependent surgery duration distributions for each surgery should be provided.

Require $\alpha_1, \alpha_2, \alpha_3$, m and N;
Require Parameters of distributions of surgery durations for i=1,2,…,N/m and j=1,2,…,N;
$w \Leftarrow 0$;
$\Theta \Leftarrow -\infty$;
while $w > \Theta$ **do**
    solve the master problem to obtain the values of x decision variables and the corresponding objective value Θ;
    substitute x into the subproblem;
    generate scenarios and obtain T matrix;
    solve the subproblem;
    $w \Leftarrow \pi^T * (h - T * x)$;
    add $\Theta \geq \pi^T * (h - T * x)$ to the master problem;
end while
w is the optimal value of the problem.

Fig. 1 L-Shaped method algorithm

The algorithm starts by initializing w and Θ variables: w is set to 0, and Θ is set to $-\infty$ in this step. w is the objective value of the subproblem, and Θ is the objective value of the master problem. At each step of this algorithm, Θ and w values are obtained, respectively. The algorithm continues to iterate as long as the value of w is larger than the value of Θ. At each iteration, an optimality cut is added to the master problem, and these cuts help Θ converge to the optimal value of the problem.

In the algorithm, the value of w is calculated as $\pi^T \times (h - T \times x)$, π is the matrix of optimal simplex multipliers, and h is the vector of constants of each constraint. In our problem, we take $h = 0$ since there exist no constants in all constraints. T is the coefficient matrix of the values of the $x_{i,j,k}$ binary variables. This matrix consists of the surgery duration values generated in different scenarios. Once the subproblem is solved, if the value of w is larger than Θ, an optimality cut, $\Theta \geq \pi^T \times (h - T \times x)$, is generated and added to the master problem.

The master problem formulation starts with the objective function, which is simply the minimization of the variable Θ. The constraints in the first iteration are constraints (4), (5), (8), (9), and (13) of the SAA problem formulation, as defined in the previous section. In the subsequent iterations, additional optimality cuts are added to the master problem, as explained before.

Once the assignment of surgeries to ORs and their sequences within the ORs are determined in the master problem, the subproblem aims to find the planned starting times of the surgeries. The objective function is exactly that of the SAA problem formulation as defined in the previous section, which is the average total weighted cost of idle time, waiting time, and overtime across the scenarios. The constraints of the subproblem are Eqs. (2–3), (6–7), and (10–12) from the SAA problem formulation.

Numerical Results

To understand the computational performance of the L-shaped method and quantify the value of considering time-dependent surgery durations, we conduct a computational study. We investigate the relationship between different parameter settings and the penalty of ignoring time-dependent surgery durations, which is calculated as the per-centage difference between the optimum objective value of the problem with and without the assumption of time-dependent surgery durations, as shown below:

$$PI = \frac{Obj^i - Obj}{Obj} \times 100\%$$

In this equation, Obj^i is the objective value obtained when time-dependent surgery durations are ignored, and Obj is the optimum objective value obtained when time-dependent surgery durations are incorporated.

In order to decrease the computation time of the L-shaped method further, the SVF rule is used to sequence the surgeries. Particularly, we include the following constraint to the master problem in order to sequence the surgeries according to the SVF rule:

$$\sum_{j=1}^{N} x_{i-1,j,k} \times V_{i-1,j} - \sum_{j=1}^{N} x_{i,j,k} \times V_{i,j} \leq 0 \quad i = 2, \ldots, n_k \text{ and } k = 1, \ldots, m$$

In this constraint $V_{i,j}$ is the variance of the distribution of the surgery duration of surgery j when it is assigned to i-th sequence. Finally, we use mean surgery durations as the planned surgery durations as a heuristic in the initial iteration of the L-shaped method to give an initial lower bound to the objective function of the master problem. The weighted cost obtained from this heuristic is used as a lower bound for the objective function of the master problem.

Experimental Setting

In our computational study, we vary a few parameters to study the extent of PI and the relationship between PI and our problem parameters. The number of operating rooms is chosen from the set $\{2, 3, 4\}$, and the number of surgeries for each operating room is taken as 3. Following the literature on surgery duration estimation, we use the log-normal distribution family to model surgery duration distributions. We take the time unit as an hour. We consider two different modes for the surgery duration distributions: short duration and long duration. For each mode, two different values for mean and coefficient of variation (CV) of surgeries durations are generated. The mean of surgery durations is generated randomly from U[0.5, 1] to model short duration surgeries and from U[1, 4] to model long duration surgeries. CV of surgery

durations is generated randomly from U[0.4, 0.8] for short duration surgeries, and from U[0.8, 1.2] for long-duration surgeries.

We consider the time-dependency of surgery durations in our computational study as follows. The mean and standard deviation of the surgery durations of surgeries assigned to the first sequence in each operating room and surgeries assigned to the last sequence in each operating room is adjusted by multipliers. The value of the coefficient of variation will remain the same since both mean and standard deviation will be multiplied by the same value. The multiplier of surgeries assigned to the first sequence in each operating room is chosen from the set $a_{first} = \{0.9, 0.95, 0.983\}$, and it is called the first-surgery multiplier. The multiplier of surgeries assigned to the last sequence in each operating room is chosen from the set $a_{last} = \{0.8, 0.9\}$, and it is called the last-surgery multiplier. These parameter values are chosen following the empirical study of [19].

Furthermore, there will be four different combinations of $\{\alpha_1, \alpha_2, \alpha_3\}$: $\{0.2, 0.2, 0.6\}$ is used to observe the effect of high overtime weight, $\{0.2, 0.6, 0.2\}$ is used to observe the effect of high waiting time weight, $\{0.6, 0.2, 0.2\}$ is used to observe the effect of high idle time weight, and $\{0.33, 0.33, 0.33\}$ is used to model equal weights settings.

There are three options for the number of ORs, three options for the first-surgery multiplier, three options for the last-surgery multiplier, four different combinations for $\{\alpha_1, \alpha_2, \alpha_3\}$, two options for the mean of surgery durations, and two options for coefficient of variation of surgery durations. In total we consider $3 \times 3 \times 2 \times 4 \times 2 \times 2 = 288$ different parameter combinations.

Analysis of the Results

We summarize the minimum, average, and maximum percentage penalty of ignoring time dependency, PI, as we change problem parameters one at a time in Table 1.

The average penalty decreases as the first-surgery multiplier increases, as expected. The slope of the decrease is smaller as one gets closer to 1. Similarly, the average penalty decreases as the last-surgery multiplier increases from 0.8 to 0.9. Remember that when the last surgery duration is overestimated, only the idle time cost due to the last surgery being completed before planned time increases. None of the preceding surgeries are affected. Hence the penalty could be limited. However, it is possible to modify OR assignment and sequence in order to take advantage of time-dependency. These results support our insights.

Next, we investigate the effect of having long surgeries or surgeries with higher duration variability on the penalty of ignorance. When the list includes relatively long surgeries, the average penalty is higher. On the other hand, when the list includes surgeries with relatively more variable durations, the average penalty decreases. We believe that when there is sufficient variability in the surgery durations, the advantage of considering time-dependency disappears.

Table 1 Statistics of PI as problem parameters vary

Parameter	Value	Min (%)	Avg (%)	Max (%)
a_{first}	0.9	0.06	1.96	6.52
	0.95	0.04	1.56	6.78
	0.983	0	1.59	4.88
a_{last}	0.8	0.06	2.04	6.52
	0.9	0	1.36	6.78
Mean mode	Small	0	1.36	5.70
	Large	0.05	1.99	6.78
CV mode	Small	0.09	2.15	6.78
	Large	0	1.12	3.27
Unit cost comb	{0.2, 0.2, 0.6}	0.12	2.01	6.52
	{0.2, 0.6, 0.2}	0.09	1.52	6.78
	{0.33, 0.33, 0.33}	0.06	1.7	5.44
	{0.6, 0.2, 0.2}	0	1.54	4.74

Keeping the unit cost of idle time as 0.2, we observe that the average penalty of ignoring time-dependent surgery durations decreases when the unit cost of waiting time increases, and the unit cost of overtime decreases. Moreover, when the unit cost of waiting time is kept as 0.2, we find that the average penalty of ignoring time-dependent surgery durations decreases as the unit cost of idle time increases, and the unit cost of overtime decreases. Finally, we observe that the average penalty is high when the overtime cost has significant weight or all unit costs are equal. The former observation is due to the effect of the last-surgery multiplier.

Computational performance of our L-shaped method deteriorates as the number of ORs increases due to the increasing number of binary variables: 36 binary variables are used when there are 2 ORs. This number is 81 with 3 ORs and 144 with 4 ORs. Table 2 shows the number of iterations in the L-shaped algorithm as well as the total solution time in seconds. There is a superlinear increase in the average number of iterations as the number of ORs increases. However, the increase in the average solution time is more superlinear. This observation highlights the fact that the master problem requires significantly more time as the number of ORs increases. To check the effect of using SVF heuristic in the master problem, we computed the solution

Table 2 Solution time and the number of iterations as the number of ORs varies

	# of iterations			Time (seconds)		
# of ORs	Min	Avg	Max	Min	Avg	Max
2	2	2.61	14	12.27	25.16	99.52
3	2	4.90	104	16.63	174.70	1471.56
4	2	8.77	315	24.52	1315.37	8853.20

with 2 ORs without this heuristic. We find that the optimum values are very close, yet the average (maximum) solution time increases from 25.16 (99.52) seconds to 167.57 (2650.69) seconds. Therefore, we conclude that using SVF heuristic in the master problem is quite effective.

Conclusions

ORs are responsible for the majority of a hospital's revenues, and also the costliest part of healthcare services. In this paper, we consider the next-day OR scheduling problem with multiple ORs. We study this problem under time-dependent and stochastic surgery durations. We formulate the problem as a Stochastic Mixed Integer Program and propose an L-shaped algorithm to solve this problem within a reasonable time limit. Through a computational study, we quantify the penalty of ignoring time-dependency in creating OR schedules and investigate the effects of problem parameters on this penalty. Results show that penalty of ignoring time-dependent surgery durations increases as the unit cost of idle time or the unit cost of waiting time decreases. Furthermore, the penalty of ignoring time-dependent surgery durations is directly proportional to surgery durations, whereas inversely proportional to surgery duration variability.

Further research is needed to develop heuristics to solve problems with larger numbers of surgeries and ORs within reasonable computational times. Another avenue for research is to incorporate other factors that effect surgery duration, such as surgery team and other resources. Finally, integration of pre and postoperative processes and their resource usage could create additional insights for the OR managers.

References

1. Abdelrasol ZY, Harraz N, Eltawil A (2013) A proposed solution framework for the operating room scheduling problems. Proc World Congr Eng Computs Sci 2:23–25
2. Baker KR (2014) Minimizing earliness and tardiness costs in stochastic scheduling. Eur J Oper Res 236(2):445–452
3. Batun S, Denton BT, Huschka TR, Schaefer AJ (2011) Operating room pooling and parallel surgery processing under uncertainty. INFORMS J Comput 23(2):220–237
4. Begen MA, Queyranne M (2011) Appointment scheduling with discrete random durations. Math Oper Res 36(2):240–257
5. Birge JR, Louveaux F (2011) Introduction to stochastic programming. Springer Science & Business Media
6. Cardoen B, Demeulemeester E, Beliën J (2010) Operating room planning and scheduling: a literature review. Eur J Oper Res 201(3):921–932

7. Centers for Medicare and Medicaid Services (2015) NHE summary including share of GDP, CY 1960–2015. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/NHEGDP15.zip. Accessed at March 2016
8. Denton B, Gupta D (2003) A sequential bounding approach for optimal appointment scheduling. IIE Trans 35(11):1003–1016
9. Denton BT, Rahman AS, Nelson H, Bailey AC (2006) Simulation of a multiple operating room surgical suite. In: Simulation conference, 2006. WSC 06. Proceedings of the winter, IEEE, pp 414–424
10. Denton B, Viapiano J, Vogl A (2007) Optimization of surgery sequencing and scheduling decisions under uncertainty. Health Care Manag Sci 10(1):13–24
11. Denton BT, Miller AJ, Balasubramanian HJ, Huschka TR (2010) Optimal allocation of surgery blocks to operating rooms under uncertainty. Oper Res 58(4):802–816
12. Dexter F, Traub RD (2002) How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. Anesth Analg 94(4):933–942
13. Erdogan SA, Denton BT (2011) Surgery planning and scheduling. Wiley encyclopedia of operations research and management science
14. Erixon F, Van der Marel E (2011) What is driving the rise in health care expenditures?: an inquiry into the nature and causes of the cost disease. ECIPE working papers 5
15. Guda H, Dawande M, Janakiraman G, Jung KS (2016) Optimal policy for a stochastic scheduling problem with applications to surgical scheduling. Prod Oper Manag 25(7):1194–1202
16. Guinet A (1993) Scheduling sequence-dependent jobs on identical parallel machines to minimize completion time criteria. Int J Prod Res 31(7):1579–1594
17. Hans EW, Van Houdenhoven M, Hulshof PJ (2012) A framework for healthcare planning and control. Handbook of healthcare system scheduling. Springer, Berlin pp 303–320
18. Healthcare Financial Management Association (2003) Achieving operating room efficiency through process integration. Healthc Financial Manage J Healthc Financial Manage Assoc 57(3):1–8
19. Kayış E, Khaniyev TT, Suermondt J, Sylvester K (2015) A robust estimation model for surgery durations with temporal, operational, and surgery team effects. Health Care Manag Sci 18(3):222–233
20. Kc DS (2013) Does multitasking improve performance? Evidence from the emergency department. Manuf Serv Oper Manag 16(2):168–183
21. Khaniyev T, Kayış E, Güllü R (2020) Next-day operating room scheduling with uncertain surgery durations: exact analysis and heuristics. Eur J Oper Res, forthcoming
22. Lamiri M, Xie X, Dolgui A, Grimaud F (2008) A stochastic model for operating room planning with elective and emergency demand for surgery. Eur J Oper Res 185(3):1026–1037
23. Lamiri M, Grimaud F, Xie X (2009) Optimization methods for a stochastic surgery planning problem. Int J Prod Econ 120(2):400–410
24. Laporte G, Louveaux FV (1993) The integer L-shaped method for stochastic integer programs with complete recourse. Oper Res Lett 13(3):133–142
25. Lebowitz P (2003) Schedule the short procedure first to improve OR efficiency. AORN J 78(4):651–659
26. Lee YH, Bhaskaran K, Pinedo M (1997) A heuristic to minimize the total weighted tardiness with sequence-dependent setups. IIE Trans 29(1):45–52
27. Lee YH, Pinedo M (1997) Scheduling jobs on parallel machines with sequence-dependent setup times. Eur J Oper Res 100(3):464–474
28. Mancilla C, Storer R (2012) A sample average approximation approach to stochastic appointment sequencing and scheduling. IIE Trans 44(8):655–670
29. Robinson LW, Chen RR (2003) Scheduling doctors' appointments: optimal and empirically-based heuristic policies. IIE Trans 35(3):295–307
30. Sherali HD, Smith JC (2001) Improving discrete model representations via symmetry considerations. Manage Sci 47(10):1396–1407

31. Wang PP (1993) Static and dynamic scheduling of customer arrivals to a single-server system. Nav Res Logist 40(3):345–360
32. Wang PP (1999) Sequencing and scheduling N customers for a stochastic server. Eur J Oper Res 119(3):729–738
33. Weiss EN (1990) Models for determining estimated start times and case orderings in hospital operating rooms. IIE Trans 22(2):143–150
34. Van Slyke RM, Wets R (1969) L-shaped linear programs with applications to optimal control and stochastic programming. SIAM J Appl Math 17(4):638–663
35. Zhao Z, Li X (2014) Scheduling elective surgeries with sequence-dependent setup times to multiple operating rooms using constraint programming. Oper Res for Health Care 3(3):160–167