Operating room scheduling for non-operating room anesthesia with emergency uncertainty

Jian-Jun Wang[1] · Zongli Dai[1] · Wenxuan Zhang[1] · Jim Junmin Shi[2]

Abstract

How to improve the efficiency of operating rooms (ORs) has always been a challenging problem in the context of healthcare operations management. This paper focuses on the research of operating room scheduling under *non-operating room anesthesia* (NORA) mechanism, in the presence of the uncertainty of emergency arrivals. In particular, we examine the advantages of the NORA mechanism in comparison with traditional surgical anesthesia practice under different operating room settings. Operationally, the process is comprised of two stages: (1) initial scheduling and (2) rescheduling. In the first stage, the initial schedule for elective surgeries under NORA is first performed through our developed model. With experiments, it is shown that for different operating room settings, the NORA mechanism can significantly improve the operating room utilization in comparison with the traditional OR anesthesia process. In the second stage of rescheduling, our experiment results show that the rescheduling model can effectively address the disruptions caused by the random arrival of emergency patients.

Keywords Operating room scheduling · Non-operating room anesthesia · Emergency patients · Heuristic algorithm

1 Introduction

The utilization of operating rooms (ORs) plays a critical role in the operations of hospitals. To wit, optimizing the utilization of ORs is essential to lower the cost of operations, to ease the patient congestion and shorten the waiting time of surgical patients, and to increase the admission rate of patients (Roshanaei et al., 2017). In practice, the service capacity of ORs cannot often meet the needs of patients demand due to the low utilization of ORs. In particular, some hospitals still face the situation that ORs have not been fully utilized which is radically caused by unreasonable or inefficient scheduling (Bandi & Gupta, 2020). Based on the prevailing practice and the related data analyses in some hospitals in China and

✉ Jim Junmin Shi
  jshi@njit.edu

[1] School of Management and Economics, Dalian University of Technology, Dalian 116024, China

[2] Martin Tuchman School of Management, New Jersey Institute of Technology, Newark 07102, USA

the United States, this paper will unveil that one main reason for a low ORs utilization is inefficient scheduling and non-operating occupation (e.g., anesthesia works) in ORs.

Typically, surgery process mainly includes three stages: preoperative, perioperative, and postoperative stages (Batun et al., 2011). Anesthesia, an essential process of surgery, is usually carried out in ORs by anesthesiologists, nurses, or anesthesia-related equipment, during which it makes the scarce resources such as surgeons and surgical equipment being idle (Hopp & Lovejoy, 2013). To address the issue, some hospitals have tried to improve the traditional surgery management process by process reengineering, i.e., setting up an independent anesthesia room to carry out centralized anesthesia for patients (e.g., NORA strategy) (Nagrebetsky et al., 2017). In this setting, the resources involved in the anesthesia process, such as anesthesiologists and nurses, are separated from ORs so that the anesthesia work will no longer occupy costly OR resources. Currently, most studies focus on surgery scheduling under the traditional surgery process (Hamid et al., 2019; Shehadeh & Padman, 2022; Shylo et al., 2013). Hence, existing theoretical research cannot be applied readily to the NORA setting. Furthermore, research on scheduling under non-operating room anesthesia (NORA) strategies is of scarcity. The joint scheduling problem between ORs and anesthesia rooms is more challenging compared with the problem under the traditional operating process.

Emergency arrivals add more wrinkles to surgical scheduling. One big obstacle is how to deal with the disruptions to the already-planned elective patient surgery schedule caused by emergency patient arrivals. This is critical as elective surgery is the main source of revenue to most hospitals. There are two main strategies to address emergency patients. One is to reserve a dedicated OR for emergency patients, and the other is to share ORs between emergency patients and elective patients, i.e., arrived emergency patients will be added and allocated to the upfront of the elective surgery queue. For the former strategy, some literature has mentioned that it is inefficient in the context of low emergency surgery arrival rates, resulting in an undesirable waste of OR resources (Batun et al., 2011; Hallah & Visintin, 2019; Jung et al., 2019). For the latter strategy, some studies set some constraints to make sure that emergency patients can get access to ORs timely within a limited waiting time (Freeman et al., 2016; Miao & Wang, 2021; van Essen et al., 2012). Different from the traditional surgery process, the scheduling of emergency patients under NORA setting needs to consider the occupation of anesthesia rooms and ORs simultaneously. In addition, random emergency arrivals often affect the implementation of the initial scheduling scheme of elective surgery, resulting in surgery delay or cancellation. These may reduce the efficiency of ORs and further cause considerable costs (Samudra et al., 2016). However, the existing OR scheduling studies considering the arrival of emergency patients mainly focus on how to effectively deal with emergency patients, but lack attention on their impact on the existing surgical scheduling scheme (Zhu et al., 2019). Undoubtedly, adding arrived emergency patients to the existing scheduling scheme will introduce additional costs, caused by the reallocation of resources and transfer of equipment. In this case, it is necessary to take a close consideration on the impact of emergency patient arrival on the implementation of elective surgery. Readers are referred to the literature of Dai et al. (2022), Li et al. (2021, 2022), Wang et al. (2021a, 2021b).

In this paper, a surgical scheduling model is formulated for the surgical scheduling problem under the NORA setting. Meanwhile, for the devised model, we propose a heuristic algorithm to solve the problem with efficiency. Besides, a rescheduling model is further proposed to alleviate the impact of uncertain emergency patients on elective patients. The specific process is mainly divided into two stages: initial scheduling and rescheduling. In the initial scheduling stage, the initial scheme under the NORA strategy is first curated through the model. Experiments show that the NORA strategy can significantly improve the utilization

of the ORs compared with the traditional anesthesia process. In the rescheduling stage, the experimental results show that the rescheduling model can effectively address the disruptions of the random arrival of emergency patients to the initial scheduling scheme.

2 Literature review

Our study is mainly related to two streams of literature. The first stream is about how to improve the utilization rate of ORs, and the second one is about how to deal with the arrival of emergency patients.

OR scheduling (also known as *surgery scheduling*) typically involves multiple stages, multiple resources, and multiple participants, and a variety of uncertainties. Some research classifications can be seen in the literature review (Zhu et al., 2019). Among those relevant reference, most studies mainly focus on how to improve the utilization rate of the OR through scheduling models and algorithms (Atighehchian et al., 2020; Rachuba & Werners, 2017; Vancroonenburg et al., 2016). In an effort to improve the efficiency of ORs, some hospitals have isolated and built independent anesthesia rooms to prevent the occupation of non-operating anesthesia in ORs (Nagrebetsky et al., 2017). However, such NORA practice has brought up with new challenges to the traditional scheduling method. Specifically, the separation of the anesthesia unit and the surgical unit requires coordination and compactness between these two units as the surgery must be performed immediately after anesthesia. However, the relevant literature on scheduling pertaining to the NORA mechanism is rather scarce. This study aims to bridge the gap.

In view of the comparison as detailed in Table 1, Tsai et al. (2017) attempted to utilization of anesthesia resources without considering the operating room. Wang et al. (2021a, 2021b) and Liang (2022) only focused on elective surgery without considering the impact on emergency patients. However, the influence of emergency patients' arrival on the utilization of ORs cannot be ignored (Freeman et al., 2016), as it may preempt medical resources and further interfere with elective surgery. Our contribution by this study is to propose a model and algorithm to address the uncertainty of the arrival of emergency patients.

Emergency patients are different from elective patients. Their surgery type, surgery duration, arrival time, and other information cannot be perceived in advance. Accordingly, the scheduling of emergency patients often involves the uncertainty of arrival time, surgery time, and resource demand. How to effectively manage the arrivals of emergency patients is one of the difficulties in OR scheduling (Breuer et al., 2020). To solve this problem, researchers have

Table 1 Comparison of research studies on NORA strategies

	Emergency arrivals	Anesthesia room	Operating room	Offline/online schedule	Contributions
Tsai et al. (2017)		√		Offline	Policy
Wang et al. (2021a, 2021b)		√	√	Offline	Robust model
Liang (2022)		√	√	Offline	MIP model
This paper	√	√	√	Online	MIP model + heuristic algorithm

proposed some strategies such as (1) reserving a separate OR (Wang et al., 2022), (2) leaving a time gap between adjacent surgery processes, or (3) postponing elective surgery (Moosavi & Ebrahimnejad, 2020; Xie et al., 2021; Zhu et al., 2019). The first one belongs to the OR management strategy, while the rest two are of surgery scheduling strategies. In general, there are three common OR management policies for emergency patients: dedicated policy, flexible (or shared) policy, and hybrid policy (Bovim et al., 2020; Breuer et al., 2020). This paper mainly studies the surgery scheduling problem under the shared OR strategy, namely, all ORs can be allocated to emergency patients and elective patients. For this setting, Essen et al. (2012) proposed the concept of break-in-moments (BIMS) to minimize the waiting time of emergency patients by evenly distributing the elective surgery completion time as much as possible. In this paper, we also control the overlapping time between surgery in multiple ORs so as to cope with the random arrivals of emergency patients. In addition, Kroer et al. (2018) considered the arrival of emergency patients and established a model to minimize overtime cost and the number of open ORs. Most of the existing literature focuses on how to effectively deal with emergency patients but barely considers the negative impact pertaining to the insertion of emergency patients (Miao & Wang, 2021). However, in practice, these negative effects may impose the initial elective surgery strategy no longer applicable. To address the issue, this paper aims to consider the impact of emergency patients on the initial scheduling scheme via rescheduling. Rachuba and Werners (2014) considered the objectives from multiple perspectives, including patients, doctors and management objectives. However, they mainly considered the uncertain needs of emergency patients and proposed a preventive scheduling scheme with robustness. In contrast, not only does this paper consider the preventive scheduling scheme for emergency patients, but also the rescheduling scheme after emergency patients are inserted. Concerning that the devised rescheduling model per se requires potential computational time for a solution, we propose a heuristic algorithm accordingly, and show that the algorithm can greatly reduce the solution time of the model (Lu et al., 2020; Wu et al., 2022).

3 Model formulation

This paper studies the OR scheduling problem under the shared OR strategy with independent anesthesia rooms. Typically, elective patients have low uncertainty. In this case, the elective surgery duration can be predicted according to the statistical analysis of hospital historical data or machine learning models (Al-Refaie et al., 2018; Devi et al., 2012; Schneider et al., 2020). Hence, without loss of generality, we simply assume that the elective surgery duration of patients can be determined in advance. Referring to Samudra et al. (2016), emergency patients with a tolerant waiting time are regarded as semi-emergency patients and can be treated within that time interval after arrival. Considering the urgency and health risk of emergency patients, we assume that the anesthesia work of all emergency patients can only be carried out in ORs. In contrast, elective patients are treated via following the NORA process. Under the NORA strategy, elective patients need to be anesthetized separately right prior to the surgery in an independent anesthesia room. The independent anesthesia room generally includes multiple anesthesia beds, anesthesiologists, anesthesia nurses, and anesthesia equipment (Nagrebetsky et al., 2017), which can carry out anesthesia for multiple patients at the same time.

Referring to the extant literature and the prevailing practice in hospitals, we make the following assumptions.

(1) The surgery that has started cannot be interrupted or canceled, and all surgeries must be completed on surgery day (Hopp & Lovejoy, 2013);

(2) In order to ensure the safety of patients and improve the utilization of the ORs, patients after anesthesia need to be transferred to an OR for surgery immediately (Macario, 2010);

(3) The whole surgery process of emergency patients is carried out in an OR (Hopp & Lovejoy, 2013);

(4) The opening time of the anesthesia room is earlier than that of ORs to ensure that patients can be conducted anesthesia before ORs open (Hopp & Lovejoy, 2013);

(5) Each emergency patient has a certain waiting time threshold (Samudra et al., 2016).

In this paper, the surgery scheduling is divided into the initial scheduling stage P_0 and the rescheduling stage P_1, upon the arrival of emergency patients. In the P_0 stage, the starting time and schedule of elective surgeries in the anesthesia rooms and ORs are determined with the objective of maximizing the efficiency of ORs. To ensure that the actual waiting time of emergency patients does not exceed their threshold, our model includes BIM constraints. Since any insertion of emergency patients disrupts the initial scheduling scheme made in P_0 stage, it is necessary to adjust the initial scheduling scheme in P_1 stage with the objective of maximizing OR utilization and minimizing surgical changes.

Before establishing the model, the main variables are summarized in Table 2 based on the above assumptions.

3.1 Initial scheduling model under NORA

At stage P_0, the main objective is to maximize the efficiency of ORs while the random arrival of potential emergency patients is considered by some preventive BIM constraints. We characterize the OR utilization rate through three types of cost (including overtime, idle and operating costs); *cf.* Jung et al. (2019). Decision variables include the OR assignment, the surgery starting time, the anesthesia room assignment, and the start time of the anesthesia for elective patients. Under the NORA setting, a mixed-integer programming model is formulated to generate an initial schedule to improve the efficiency of ORs to treat the potential arrival of emergency patients as follows:

$$\text{Min } Z = \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{t=1}^{T} c_r\, x_{ijt} + \sum_{j=1}^{M}\sum_{t=1}^{T} c_l \left(1 - \sum_{i=1}^{N} x_{ijt}\right) + \sum_{i=1}^{N}\sum_{j=1}^{M}\sum_{t=T}^{T+L} c_o\, x_{ijt}, \quad (1)$$

$$\sum_{t=1}^{T+L}\sum_{j=1}^{M} y_{ijt} = 1, \quad i = 1, 2, \ldots, N, \quad (2)$$

$$\sum_{i=1}^{N} x_{ijt} \leq 1, \quad j = 1, 2, \ldots, M, \ t = 1, 2, \ldots, T+L, \quad (3)$$

$$\sum_{r=t}^{\min(t+p_i-1,\,T+L)} x_{ijr} \geq p_i y_{ijt}, \quad i = 1, 2, \ldots, N, \ j = 1, 2, \ldots, M, \ t = 1, 2, \ldots, T+L, \quad (4)$$

$$\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ijt} = O_t, \quad t = 1, 2, \ldots, T+L, \quad (5)$$

Table 2 Notions of parameters and variables

Indexes

$i = 1, 2, ..., N$: Surgical patients i, N is the total number of surgical patients on surgery day

$j = 1, 2, ..., M$: OR j, M is the number of ORs

$f = 1, 2, ..., F$: Anesthesia room f, F is the number of anesthesia rooms

$t = 1, 2, ..., T + L$: time t, T is the normal opening time of ORs and L is the maximum overtime of each OR

Parameter

p_i: Surgery duration of surgical patient i, $i = 1, 2, ..., N$

s_i: Duration of anesthesia for surgical patients i, $i = 1, 2, ..., N$

c_r: Operating cost per time unit to keep an OR open for elective surgery during regular time T

c_l: Unit idle cost of an OR during normal opening time T

c_o: Unit overtime cost of an OR exceeding normal opening time T

c_t: The change-related cost in rescheduling

c_p: The balance cost in rescheduling

u: The duration of overlap in all ORs cannot exceed u time units

Z: Total cost, including overtime, idle and operating costs

$Beds$: The number of staffed beds in the anesthesia room, i.e., the maximum number of patients who can obtain anesthesia at the same time

H: Index set of time, $H = \{1, 2, \ldots, T + L\}$

Decision variables

x_{ijt}: 0–1 variable, if the operating patient i is using the OR j at time t, the value is 1; otherwise, 0

y_{ijt}: 0–1 variable, if the operating patient i starts the surgery in the OR j at time t, the value is 1; otherwise, 0

y'_{ijt}: 0–1 binary variable, if the operating patient i starts the surgery in the OR j at time t before emergency patients arrive in rescheduling model, the value is 1; otherwise, 0

a_{ift}: 0–1 variable, if the surgical patient i starts anesthesia in the anesthesia room f at time t, the value is 1; Otherwise, 0

O_t: Integer variable, the number of ORs occupied at time t, $O_t \in [0, M]$

$$\sum_{t=k}^{k+u} O_t - \sum_{i=1}^{N} \sum_{t=k+1}^{k+u} \sum_{j=1}^{M} y_{ijt} \le (u+1)M - 1, \quad k = 1, 2, \ldots, T, \tag{6}$$

$$\sum_{t=1}^{T+L} O_t = \sum_{i=1}^{N} p_i, \tag{7}$$

$$\sum_{f=1}^{F} \sum_{t=1}^{T+L} a_{ift} \le 1, \quad i = 1, 2, \ldots, N, \tag{8}$$

$$\sum_{i=1}^{N} a_{ift} \le Beds, \quad f = 1, 2, \ldots, F, \ t = 1, 2, \ldots, T + L, \tag{9}$$

$$\sum_{t=s_i+1}^{T+L} \sum_{j=1}^{M} H_t \cdot y_{ijt} - s_i \sum_{t=s_i+1}^{T+L} \sum_{j=1}^{M} y_{ijt} = (H_t - s_i) \sum_{t=s_i+1}^{T+L} \sum_{f=1}^{F} a_{ift}, \quad i = 1, 2, \ldots, N, \tag{10}$$

$$x_{ijt}, y_{ijt}, a_{ift} \in \{0, 1\},$$
$$i = 1, 2, \ldots, N, \ j = 1, 2, \ldots, M, \ t = 1, 2, \ldots, T + L,$$
$$f = 1, 2, \ldots, F, \tag{11}$$

$$O_t \in \{0, 1, \ldots, M\}, \quad t = 1, 2, \ldots, T + L. \tag{12}$$

In the above model, the objective given in Eq. (1) is to minimize the total cost of overtime, idleness and the operating cost of ORs. The fixed cost of ORs and other costs that are not related to OR efficiency are not considered, as the hospital's main purpose is to increase the utilization rate of ORs. Constraint (2) indicates that each surgery can only be started once. Constraint (3) ensures that at most one surgery can be performed at any moment in each OR. Constraint (4) indicates that the ongoing surgery cannot be interrupted. Constraint (5) represents the number of all ORs occupied at a certain time. Constraint (6) indicates that the overlap duration of all surgery in ORs cannot exceed a certain threshold. Referring to the concept BIMs (van Essen et al., 2012), emergency patients can get timely access to ORs by controlling the start and end time of elective surgeries. Therefore, there may be short idleness between adjacent operations in the initial scheduling plan. This can ensure that emergency patients arriving at any time can be timely inserted into an OR within the threshold time. Constraint (7) indicates the relationship between the occupancy of ORs and the surgery duration. Constraint (8) indicates that the patient is anesthetized only once. Constraint (9) indicates that the number of patients getting anesthesia simultaneously cannot exceed the maximum number of beds. Constraint (10) means that the patient must enter the OR for surgery after anesthesia without waiting. Constraints (10) and (11) limit the range of decision variables.

3.2 Rescheduling model considering emergency patient arrivals under NORA

In the rescheduling model, three types of costs are mainly considered: overtime cost, change-related cost and balance cost. In order to clarify the concept and relationship of the objective, we give a brief example with three ORs, as depicted in Fig. 1. The upper subfigure is the initial scheduling and the lower subfigure is the rescheduling under the same condition, where S is the elective surgery, E is the emergency surgery, the arrival time and surgery duration of the emergency are 7 and 5 respectively, and the normal opening time of the OR is 16. Comparing the rescheduling plan with the initial scheduling plan, it can be seen that the total overtime of the OR is 4, the changed surgeries are $S2$ and $S3$, and the overtime of the three ORs are 0, 0, and 4 respectively. Then the *balance time* is $[(4 - 0) + (4 - 0) + (0 - 0)]$ as shown in Fig. 2. Therefore, the *balance cost* is $[(4 - 0) + (4 - 0) + (0 - 0)] * c_p$, and the total cost is expressed as $Z' = 4c_o + 2c_t + 8c_p$.

As depicted Fig. 1, E-1 to E-6 indicate the insertion time for emergency patients because the elective surgeries that have already started cannot be stopped. The term "*overlapping duration*" refers to the time range in which all ORs are occupied, during which urgent patients had to wait. Emergency patients have a certain waiting time threshold u (Jung et al., 2019). Therefore, it wouldn't influence emergency patients' health as long as the *overlapping duration* $\leq u$. Note that the *overlapping duration* is determined by the scheduling model, while u is given by hospitals. For example, if an emergency patient arrives at time 6, the earliest time he can start surgery is between E-1 and E-3. If the time is E-3 (i.e.,7), then the waiting time is 1. In this case, S2 must be postponed, resulting in OR2 overtime. Further, it may also lead to *overlapping duration* $\geq u$, which means the next emergency patient may wait for more than u. H1 mainly aims to solve the aforementioned problems.

Rescheduling is required after each emergency patient arrives. The time of the emergency patient's arrival and duration of the emergency patient are input. The overlap between ORs
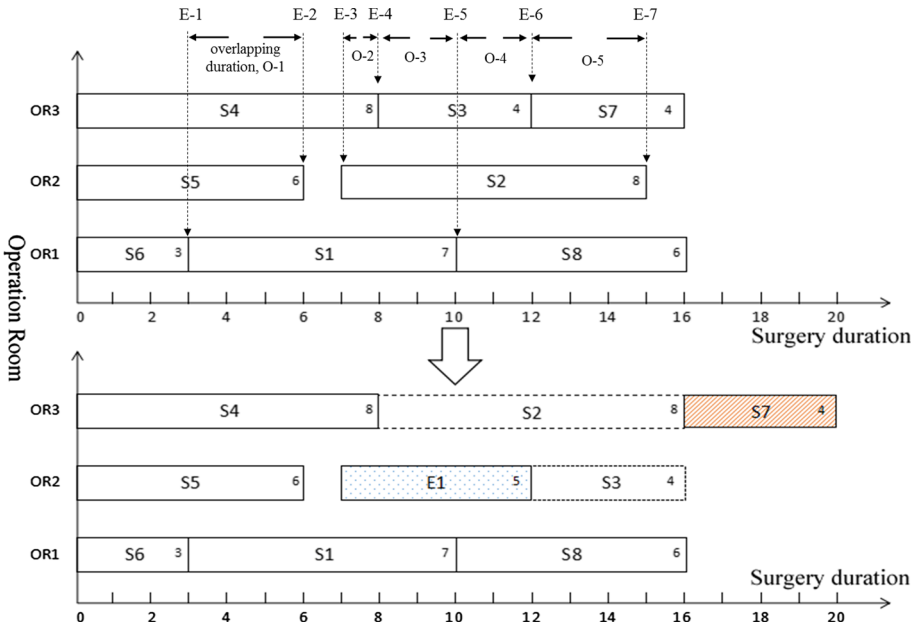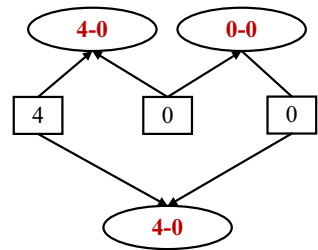
Fig. 1 An example of rescheduling schedule (E1: emergency surgery)

Fig. 2 An example of calculating balance time

needs to be considered in order to meet the subsequent insertion of emergency patients. When adjusting the initial scheduling plan, we try to minimize the adjustment between ORs as much as possible, which is consistent with the rescheduling objective. At the same time, we also consider the unstarted surgery in the OR with overtime. The limit of OR overtime may be determined according to different actual needs.

At the rescheduling stage, changing ORs may cost a lot due to the inconvenience for the implementation of the rescheduling plan, because preoperative preparations, equipment, and other resources may need to be adjusted immediately. Besides, the OR with too much overtime will also incur additional costs (Bandi & Gupta, 2020), and this situation is undesirable. In view of the above analyses, we propose the following rescheduling model:

$$\text{Min } Z' = \sum_{i=1}^{N+1} \sum_{j=1}^{M} \sum_{t=T}^{T+L} c_o x_{ijt} + c_t N' + c_p S, \tag{13}$$

$$N' = \sum_{i=1}^{N} \left(1 - \sum_{t=1}^{T+L} y_{ij't} \right), \quad j' = \sum_{j=1}^{M} \sum_{t=1}^{T+L} j y'_{ijt}, \quad i = 1, 2, \ldots, N, \tag{14}$$

$$S = \sum_{j=1}^{M} \sum_{j'=j}^{M} \left[\left| \sum_{i=1}^{N+1} \sum_{t=T+1}^{T+L} \left(x_{ijt} - x_{ij't} \right) \right| \right] \tag{15}$$

$$\sum_{t=1}^{T+L} \sum_{j=1}^{M} y_{ijt} = 1, \quad i = 1, 2, \ldots, N+1, \tag{16}$$

$$\sum_{i=1}^{N+1} x_{ijt} \leq 1, \quad j = 1, 2, \ldots, M, \ t = 1, 2, \ldots, T+L, \tag{17}$$

$$\sum_{r=t}^{\min(t+p_i-1, T+L)} x_{ijr} \geq p_i y_{ijt}, \quad i = 1, 2, \ldots, N+1, \ j = 1, 2, \ldots, M, \ t = 1, 2, \ldots, T+L, \tag{18}$$

$$\sum_{i=1}^{N+1} \sum_{j=1}^{M} x_{ijt} = O_t, \quad t = 1, 2, \ldots, T+L, \tag{19}$$

$$\sum_{t=k}^{k+u} O_t - \sum_{i=1}^{N} \sum_{t=k+1}^{k+u} \sum_{j=1}^{M} y_{ijt} \leq (u+1)M - 1, \quad k = D+1, \ldots, T+L, \tag{20}$$

$$\sum_{t=1}^{T+L} O_t = \sum_{i=1}^{N+1} p_i, \tag{21}$$

$$\sum_{f=1}^{F} \sum_{t=1}^{T+L} a_{ift} \leq 1, \quad i = 1, 2, \ldots, N+1, \tag{22}$$

$$\sum_{i=1}^{N+1} a_{ift} \leq Beds, \quad f = 1, 2, \ldots, F, \ t = 1, 2, \ldots, T+L, \tag{23}$$

$$\sum_{t=s_i+1}^{T+L} \sum_{j=1}^{M} H_t \cdot y_{ijt} - s_i \sum_{t=s_i+1}^{T+L} \sum_{j=1}^{M} y_{ijt} = (H_t - s_i) \sum_{t=s_i+1}^{T+L} \sum_{f=1}^{F} a_{ift}, \quad i = 1, 2, \ldots, N+1, \tag{24}$$

$$x_{ijt}, y_{ijt}, a_{ift} \in \{0, 1\}, \quad i = 1, 2, \ldots, N+1, \ j = 1, 2, \ldots, M, \ t = 1, 2, \ldots, T+L,$$
$$f = 1, 2, \ldots, F, \tag{25}$$

$$O_t = \{0, 1, \ldots, M\}, \quad t = 1, 2, \ldots, T+L, \tag{26}$$

$$y'_{ijt} : given, \quad i = 1, 2, \ldots, N+1, \ j = 1, 2, \ldots, M, \ t = 1, 2, \ldots, t_u, \tag{27}$$

Some of the constraints in the rescheduling model are similar to the initial scheduling model, and we will explain the meaning of new constraints. Constraint (13) indicates that the model's objective is composed of three parts, where c_o, c_t, and c_p are respectively the overtime, change and balance cost per unit time. The overtime cost is the same as the initial scheduling model. The change-related cost is the cost of an elective surgery switching from one OR to another. In order to prevent a certain OR from being overloaded, the balanced cost is set as the penalty caused by the excessive overtime of the surgery. Constraint (14)

is the number of swaps of elective patients compared with the initial scheduling plan after emergency patients arrive. Constraint (15) represents the *balance time* of all ORs. In constraint (20), D is the minimum value of the emergency patient's surgery start time t_r and the earliest completion time of the currently undergoing elective surgery. This constraint ensures that future emergency patients can also get timely treatment after inserting one emergency patient. In constraint (27), t_u is the actual arrival time of the emergency patient. This constraint means that we take the solution of the initial scheduling model as the input of our rescheduling model.

The surgical scheduling problem with random arrivals of emergency patients is NP-hard when the size of the ORs exceeds 2 (Lamiri et al., 2008). For the initial scheduling model, when the size of the OR exceeds 20, the traditional solution method cannot obtain optimal solutions in a reasonable time, and the rescheduling model is more difficult to solve. In order to solve the problem effectively in a short time, it is reasonable to develop approximate algorithms. In what follows, we designed a heuristic algorithm to solve the large-scale problem and verified the performance of the algorithm in numerical experiments.

4 Solution methodologies

Two heuristic algorithms are proposed for the initial scheduling problem P_0 and the rescheduling problem P_1. The purpose of developing this heuristic is to provide a simple, effective and easy-to-implement method that does not rely on complex software. The algorithm is based on the rule of $LPT - SPT$ proposed by Jung et al. (2019). In the two heuristic algorithms, the subset of surgeries is fixed, and the number of ORs $M \geq 2$.

First of all, the heuristic algorithm H_0 is leveraged to solve the initial scheduling problem. The algorithm is delineated as follows.

In the heuristic algorithm H_0, we take the duration of elective surgery, the number of ORs and the number of anesthesia rooms as input, and the output is an initial scheduling plan. We schedule all surgeries with shorter duration in the same OR. These surgeries can be started and ended quickly. In other ORs surgeries with longer duration are scheduled to deal with the random arrival of emergency patients.

The cost of the OR overtime cost and idle cost is the minimal (i.e., the schedule is optimal) if either of following two conditions is satisfied. (1) If in schedule S, $\tau_j \leq T$, $j = 1, 2, \ldots, M$. (2) If in schedule S, $\tau_j \geq T$, $j = 1, 2, \ldots, M$ and only surgeries with surgery duration less than two hours are assigned to OR_1.

Heuristic Algorithm H_0

Step 1 The elective surgeries are arranged in descending order according to their surgery duration, then we get an ordered set of elective surgeries S ;

Step 2 Schedue OR_1 by SPT rule and $OR_2 \sim OR_M$ by LPT rule.

Step 3 *IF* $S \neq \varnothing$, find the earliest available OR $\overline{OR}$,   *IF* $\overline{OR} = OR_1$, schedule the last surgery in the set S to OR_1 , and remove this surgery from the set S .

IF $\overline{OR} \neq OR_1$, schedule the first surgery in the set S to $\overline{OR}$, and remove this surgery from the set S .   *IF* $S = \varnothing$ ，*End* (To ensure that the occupancy of different ORs is as even as possible. In the case with a certain number of surgeries, the smaller the utilization difference between different ORs, the lower the cost of overtime and idle time, in which sense, it is closer to the optimal).

Step 4 Estimate whether all ORs can accommodate an emergency surgery within duration u after time t_i . Specifically, if the surgery for the i th patient is to schedule at time t_i , then check, starting from time t_i , whether overlapping duration $\leq u$. If so, then an emergency surgery can be accommodated within duration u after time t_i . If they can, then schedule current surgery, else restart estimate from time $t_i + 1$. The steps will be repeated until the surgery is scheduled or to the moment $t_i + u$. This step can ensure that emergency patients who arrive randomly can be inserted into an OR (To minimize the OR overtime cost and idle cost by minimizing the delay of the operating time t_i , which can help reduce the overtime of the OR).

Step 5 Determine the time t_i for the surgery to be scheduled, and check the availability of the anesthesia room. If the anesthesia room is idle at time $t_i - s_i$, patients can be anesthetized, or else the start time of the current surgery needs to be adjusted based on Step 3.

Remark For (1), the fact that no ORs have overtime means that all surgeries have been scheduled. At this point, Schedule S^* is optimal as the sum of OR idle cost and surgery operating cost is the minimum. For (2), only the surgeries with surgery duration less than two hours are assigned to OR_1; it means that OR_1 has enough insert points to deal with the emergency patients. OR_j, $j = 2, 3, \ldots, M$ can be scheduled closely, with no idle cost and the lowest overtime cost. For other conditions, to ensure BIM constraint, in Step 4, we adjust the starting time of surgery from t_i to $t_i + u$, which helps ease the delay of the starting time obtained in Step 3.

Next, the heuristic algorithm H_1 for rescheduling is introduced. The steps are shown as follows.

Heuristic Algorithm H_1

Step 1 Input initial scheduling scheme obtained by algorithm H_0;

Step 2 Define set U including the emergency surgeries, which is an ordered set that represents the arrival order of emergency patients.

Step 3 IF $U \neq \varnothing$, find all insertable time points for the first emergency patient and define the time points set as T_{IN}. If the T_{IN} set has only one insertion point, insert the emergency patient and remove it from the set. If the T_{IN} set has multiple insertion points, insert emergency patients into the OR with more free time and remove them from the set. Then adjust the time of the elective surgery that has not started in the OR, postpone the start time of the elective surgery. Next check whether there is overtime in that OR. If an OR's overtime exceeds the threshold, adjust the last surgery in that OR to another idle OR.

Step 4 This step is similar to Step 4 of the algorithm H_0.

Step 5 This step is similar to Step 5 of the algorithm H_0.

---

5 Computational studies and experiments

5.1 Experiment setup

In order to verify and analyze the effectiveness of the OR scheduling model under the independent anesthesia room and the algorithm, we designed three sets of experiments. Experiment 1, Experiment 2, and Experiment 3 respectively verify the effectiveness of the NORA strategy, the effectiveness of the proposed algorithm, and the performance of rescheduling.

We set experimental parameters by referring to literature, discussing with some practitioners, and analyzing some data as reported in Table 3. The parameter setting is as follows: the normal working time of the OR is generally no more than 8 hours (h) per day, and the maximum overtime is no more than 4 h. In the numerical experiment, we set half an hour as one time unit. The normal opening time of all ORs is set to $T = 16$, which is 8 h. The maximum OR overtime is $L = 8$ or 4 h. The surgery duration is between 1.5–4 h. The surgery duration follows the uniform distribution from the interval $p_i \in [3, 8]$ and must be an integer. In order to simplify the experiment, the duration of anesthesia for the surgery is 0.5 h, that is, $s_i = 1$. The number of patients that the anesthesia room can serve at the same time is 4. The maximum time that emergency patients can tolerate is 4. Referring to Zhang et al. (2014), the unit open cost c_r, overtime cost c_l and idle cost c_o of the initial scheduling model are set to 1/unit time, 1.5/unit time and 2/unit time. At the same time, the change-related cost, overtime cost and balance cost of the rescheduling model are set to 4/unit time, 2/unit time and 0.5/unit time.

Table 3 Parameters setting

Experimental parameters	Values	Support
The normal working time of the OR	8 h	Practitioners
The maximum overtime	4 h	Practitioners
Time unit	0.5 h	Jung et al. (2019)
The surgery duration	1.5–4 h	Data
The duration of anesthesia	0.5 h	Data
The maximum time that emergency patients can tolerate	$u = 2$ hours	Zhang et al. (2014)
Unit operating cost c_r, overtime cost c_l and idle cost c_o	$c_r = 1$/unit time, $c_l = 1.5$/unit time, $c_o = 2$/unit time	Jung et al. (2019)
The change-related cost, overtime cost and balance cost	$c_t = 4$/unit time, $c_o = 2$/unit time, $c_p = 0.5$/unit time	Practitioners

We conduct experiments under different OR scales, and each group of experiments generates surgery duration p_i based on the total surgery duration $\overline{P}$ of N surgery, $\overline{P} = \sum_{i=1}^{N} p_i$. The $\overline{P}$ can be used to describe a utilization situation of ORs. All codes and numerical experiments are conducted in Matlab R2014a with Gurobi 8.1.0.

5.2 Performance comparison: operating room anesthesia (ORA) versus NORA

Experiment 1 was conducted for the initial scheduling stage P_0. Under the premise of considering emergency patients, we compared the non-OR anesthesia strategy with the traditional OR anesthesia strategy and verified the advantages of the NORA strategy under different OR scales. All solutions in Experiment 1 are obtained by Gurobi. The detailed experimental results are reported in Table 4 as below.

In view of Table 4, the objective value of the NORA strategy is better than the traditional operating room anesthesia (ORA) strategy in which a long overtime is incurred. The NORA strategy improves the turnover rate of ORs which saves a lot of resources. With the increase in the number of ORs and surgeries, NORA strategies can save more OR resources and allocate more surgery. In addition, according to related literature (Marjamaa et al., 2009), the open cost of an anesthesia room unit is far less than that of an OR resource. In addition, the NORA strategy can help reasonably allocate related resources such as anesthesiologists and anesthesia nurses, enabling a medical team to serve multiple anesthetized patients, which effectively improves resource efficiency and reduces OR opening costs.

5.3 Performance analysis of proposed heuristic algorithm

For the initial scheduling stage, we verified the effectiveness of the proposed heuristic algorithm H_0. In this part, the scale of the OR is set to be small, medium and large, and there are 3 experimental examples under each scale. In order to show the effectiveness of the algorithm, we compared the H_0 with the solver and genetic algorithm (GA), which has been successfully applied to the surgical scheduling model; cf. Guo et al. (2016).

The experimental results are summarized in Table 5. In a small-scale case, the solver can obtain the optimal solution in a short time. At a medium scale, the solver can also get the

Table 4 Experimental comparison of ORA and NORA strategy (OB: objective value; OT: OR overtime; IT: OR idle time.)

Parameters	Total surgery duration $\overline{P}$	NORA			ORA			OB(ORA)-OB(NORA)
		OB	OT	IT	OB	OT	IT	
N=8, M=3, F=1	48	50.5	1	1	66.5	9	1	16
	46	49	0	2	60	6	0	11
	52	56	4	0	74.5	13	1	18.5
N=10, M=4, F=1	62	65	0	2	80	8	0	15
	64	64	0	0	86.5	11	1	22.5
	59	66.5	0	5	74	5	0	7.5
N=14, M=5, F=2	81	82	1	0	110	15	0	28
	78	81	0	2	104	12	0	23
	80	82.5	1	1	110.5	15	1	28
N=17, M=6, F=2	84	102	0	12	108.5	6	1	6.5
	85	101.5	0	11	110.5	7	1	9
	87	100.5	0	9	114.5	9	1	14

Table 5 Experimental comparison between the proposed algorithm and genetic algorithm

Scale	Instances	Parameters	Solver		GA		H_0	
			OB	Time/s	OB	Time/s	OB	Time/s
Small	1	N=8, M=3, F=1	50.50	12.17	50.5	370.30	55.50	< 0.01
	2	N=10, M=4, F=1	65.00	14.57	67	388.09	67.50	< 0.01
	3	N=14, M=5, F=2	82.00	28.23	86	383.72	89.50	< 0.01
Medium	4	N=17, M=6, F=2	97.50	35.98	103	405.34	105.00	< 0.01
	5	N=20, M=7, F=2	118.50	111.91	126	399.23	128.50	< 0.01
	6	N=23, M=8, F=2	129.00	71.64	147	407.41	141.50	< 0.01
Large	7	N=26, M=9, F=3	146.50	174.79	163.5	406.34	161.50	< 0.01
	8	N=29,M=10,F=3	–	–	188	596.43	181.50	< 0.01
	9	N=35,M=12,F=3	–	–	228	414.21	210.50	< 0.01

"–" indicates that the optimal solution cannot be obtained within 8 h. Small scale: the number of surgeries is from 0 to 16; Medium scale: the number of surgeries is from 17 to 25; Large scale: the number of surgeries is from 26 to 35

optimal solution in about 2 min. But when the problem scale increases, the optimal solution cannot be obtained within 8 h. In comparison, the GA algorithm is faster, but as the problem scale increases, it still takes about 5 min in order to solve the problem. This means that all patients who have not yet started surgery need to wait for several minutes to be rescheduled, which will result in a waste of time in ORs. In contrast, we found that the proposed heuristic algorithm has a very short computational time, less than 0.01 s on average, which is better than the GA algorithm. This shows that H_0 is effective to solve the proposed model.

To further verify the effectiveness of the algorithm, we use a common evaluation indicator defined as follows:

$$GAP = \frac{(Z - Z_{best})}{Z_{best}} \times 100$$

where Z represents the objective value obtained by the current algorithm, and Z_{best} represents the optimal value of the problem.

Considering that the optimal objective value may not be obtained in large-scale cases, referring to Jung et al. (2019), we replace Z_{best} with the optimal lower bound LB of the solution. LB is the optimal objective value that can be reached under a certain scale in the ideal state where some constraints are ignored.

$$LB = c_r \cdot \min(\overline{P}, M \cdot T) + c_o \cdot \max(\overline{P} - M \cdot T, 0) + c_l \max(M \cdot T - \overline{P}, 0);$$

$$\overline{P} = \sum_{i=1}^{N} p_i,$$

where M and T indicate the number of open ORs and the normal open time of ORs, respectively.

The experimental results are reported in Fig. 3. Compared with H_0, GA algorithm has a compelling advantage in solving accuracy at a small scale. However, that advantage diminishes as the scale gets bigger. On the other hand, compared with GA algorithm, H_0 shows more obvious advantage in terms of the solving speed. Especially for large-scale problems, the algorithm still leads to a short solution time. It implies that scheduling decisions can be obtained faster in response to emergent emergency patients, which is more applicable in practice. Undoubtedly, the combination of H_0 and GA will produce a better solution accuracy without increasing the algorithm time.
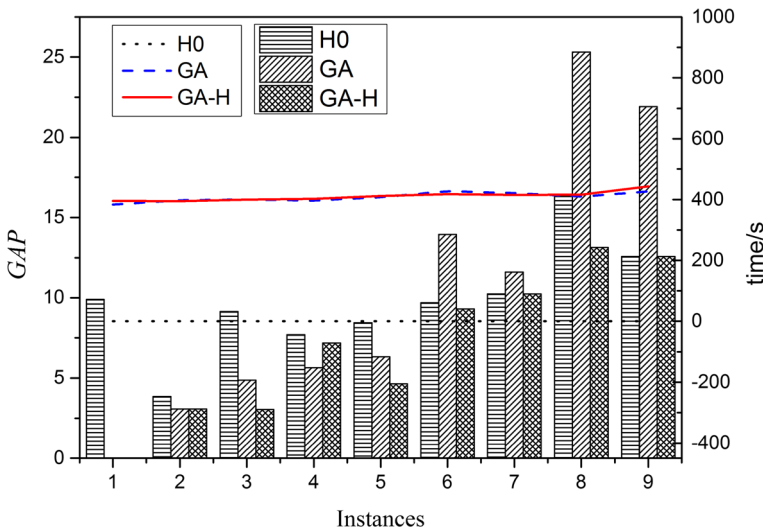

Fig. 3 Comparison between the proposed heuristic algorithm and intelligent algorithm (H_0: objective function values/running time of proposed heuristic algorithm; GA: objective function values /running time of GA algorithm; GA-H: objective function values/running time of Hybrid algorithm based on GA algorithm and H_0)

5.4 Performance analysis for emergency patient arrival

In the experiment for the rescheduling stage, considering the different needs of emergency patients, we generate emergency patient arrival time and surgery duration when the number of emergency patients is 1, 2, and 3 respectively. We set a traditional complete rescheduling (CR) strategy for comparison, which is a commonly-used rescheduling scheme whose objective is to minimize the overtime cost. The proposed rescheduling is incomplete rescheduling (ICR) whose objective is to minimize changes in ORs. In the experiment, we analyze the results of the MIP model and the heuristic algorithm H_1.

The specific experimental results are reported in Tables 6, 7 and 8 under different settings, where t_u is the arrival time of emergency patients, and p_e is the surgery duration of emergency patients. For example, $t_u = [2, 5]$, $p_e = [7, 4]$ indicate that the arrival time of the first emergency patient and the second emergency patient are 2 and 5 respectively, and the surgery duration of the two patients are 7 and 4 respectively. In the third experiment, the main performance indicators considered are the total cost of ORs, overtime cost, change-related cost, and the actual start time of emergency patient surgery.

It can be observed from Table 6 that under the same cases, the objective value obtained by solver is better than the result of the heuristic algorithm H_1, but when the size of the ORs and the number of patients become larger, the solver can no longer effectively cope with the decision making in a reasonable time. In addition, under the NORA strategy, the total cost of the solver and the heuristic algorithm is better than that under the CR strategy.

Table 6 Performance of the reschedule models and algorithm with $t_u = 7$, $p_e = 5$

Scale	Total surgery duration	ICR								CR			
		Solver				H_1				H_1			
		t_r	Z'	co	ct	t_r	Z'	co	ct	t_r	Z'	co	ct
$N = 17$	94	8	18	8	0	8	20	12	0	8	30.5	6	20
$M = 6$	96	8	26	12	0	8	28	16	4	8	48	12	24
$F = 2$	98	9	30.5	14	0	8	32.5	18	4	9	44	14	20
$N = 23$	126	8	27.5	10	0	8	32	16	4	8	49.5	10	24
$M = 8$	128	8	27.5	10	0	8	29	14	0	8	57.5	10	36
$F = 2$	130	8	36.5	14	0	8	38	20	4	8	53	14	32
$N = 29$	158	8	22.5	6	4	8	35	16	4	8	70	6	52
$M = 10$	160	–	–	–	–	8	37	18	4	8	74	10	52
$F = 3$	162	–	–	–	–	8	44.5	22	4	8	86	14	56

t_r is insert time for emergency patient; Z' is the total cost; co is overtime cost; ct is variable cost; "–" indicates that no result has been calculated in 3600 s

Table 7 Performance of the reschedule models and algorithm with $t_u = [2, 5]$, $p_e = [7, 4]$

| Scale | Total surgery duration | ICR | | | | | | | | | CR | | | |
| | | Solver | | | | H_1 | | | | | H_1 | | | |
		t_r	Z'	co	ct	t_r	Z'	co	ct		t_r	Z'	co	ct
$N = 17$	94	[5,8]	28.5	14	0	[4,9]	30	16	4		[5,8]	47.5	14	20
$M = 6$	96	[5,8]	35	18	0	[4,8]	35.5	18	8		[5,8]	65.5	18	32
$F = 2$	98	[5,8]	31	22	0	[4,8]	40.5	22	8		[5,8]	65.5	22	24
$N = 23$	126	[5,10]	33	18	8	[4,8]	41	20	4		[5,9]	83	16	44
$M = 8$	128	[5,8]	27	18	0	[4,8]	33.5	20	4		[5,8]	89.5	18	48
$F = 2$	130	[5,8]	37	22	0	[4,9]	37.5	22	4		[5,8]	86.5	22	32
$N = 29$	158	[4,8]	36	14	0	[4,8]	41	20	4		[5,8]	96.5	14	60
$M = 10$	160	–	–	–	–	[4,8]	39.5	22	4		[4,8]	91	18	56
$F = 3$	162	–	–	–	–	[4,8]	49.5	26	4		[5,8]	107.5	22	60

Table 8 Performance of the reschedule models and algorithm with $t_u = [4, 8, 9]$, $p_e = [6, 6, 4]$

Scale	Total surgery duration	ICR								CR			
		Solver				H_1				H_1			
		t_r	Z'	co	ct	t_r	Z'	co	ct	t_r	Z'	co	ct
$N = 17$	94	[6,8,11]	40	28	4	[8,9,12]	43	28	4	[6,8,11]	83	28	40
$M = 6$	96	[5,9,10]	47.5	32	4	[8,9,12]	50	38	4	[5,9,11]	78.5	32	32
$F = 2$	98	[5,10,12]	61.5	38	8	[8,9,10]	66.5	42	8	[5,9,11]	92	36	40
$N = 23$	126	[5,9,10]	49	30	4	[8,9,11]	50	32	4	[5,9,10]	103.5	30	40
$M = 8$	128	[5,9,10]	53	32	4	[8,9,12]	57	34	8	[5,9,11]	110	32	52
$F = 2$	130	[5,9,10]	57	36	4	[8,9,12]	63	38	8	[5,9,10]	121	36	52
$N = 29$	158	[5,9,11]	57	28	4	[8,9,11]	59	28	8	[5,9,11]	133	28	68
$M = 10$	160	–	–	–	–	[8,11,13]	84	36	8	[5,9,11]	138	32	64
$F = 3$	162	–	–	–	–	[7,9,13]	76	40	8	[5,9,11]	144	36	68

5.5 Sensitivity analysis

Given that the unit operating cost, OR overtime cost and idle cost of the initial scheduling model, and the emergency patient arrival time, the changed-related cost, overtime cost and balance cost of the rescheduling phase may affect the effectiveness of the model and algorithm, we conduct a sensitivity analysis.

As shown in Tables 9, 10 and 11, we first tested the impact of different unit costs on the effectiveness of the model and heuristic algorithm. It can be found that within a certain parameter value range, the change of unit cost only affects the value of the objective function, and has no effect on the utilization rate of ORs. This also shows that the initial scheduling model has good robustness. At the same time, experiments on heuristic algorithms also show

Table 9 Performance of the initial scheduling model at different unit cost levels

Unit cost			Objective values	OR overtime	OR idle time
$c_r = 1.5$	$c_l = 1.5$	$c_o = 2$	174	3	1
$c_r = 1.3$	$c_l = 1.5$	$c_o = 2$	151.8	3	1
$c_r = 0.9$	$c_l = 1.5$	$c_o = 2$	107.4	3	1
$c_r = 0.5$	$c_l = 1.5$	$c_o = 2$	63	3	1
$c_r = 1$	$c_l = 2$	$c_o = 2$	119	3	1
$c_r = 1$	$c_l = 1.8$	$c_o = 2$	118.8	3	1
$c_r = 1$	$c_l = 1.4$	$c_o = 2$	118.4	3	1
$c_r = 1$	$c_l = 1.0$	$c_o = 2$	118	3	1
$c_r = 1$	$c_l = 1.5$	$c_o = 4$	124.5	3	1
$c_r = 1$	$c_l = 1.5$	$c_o = 3.5$	123	3	1
$c_r = 1$	$c_l = 1.5$	$c_o = 2.5$	120	3	1
$c_r = 1$	$c_l = 1.5$	$c_o = 1.5$	117	3	1

Table 10 Performance of the algorithm at different unit cost levels

Unit cost			Objective values	OR overtime	OR idle time
$c_r = 1.5$	$c_l = 1.5$	$c_o = 2$	182	7	5
$c_r = 1.3$	$c_l = 1.5$	$c_o = 2$	160.6	7	5
$c_r = 0.9$	$c_l = 1.5$	$c_o = 2$	117.8	7	5
$c_r = 0.5$	$c_l = 1.5$	$c_o = 2$	75	7	5
$c_r = 1$	$c_l = 2$	$c_o = 2$	131	7	5
$c_r = 1$	$c_l = 1.8$	$c_o = 2$	130	7	5
$c_r = 1$	$c_l = 1.4$	$c_o = 2$	128	7	5
$c_r = 1$	$c_l = 1.0$	$c_o = 2$	126	7	5
$c_r = 1$	$c_l = 1.5$	$c_o = 4$	142.5	7	5
$c_r = 1$	$c_l = 1.5$	$c_o = 3.5$	139	7	5
$c_r = 1$	$c_l = 1.5$	$c_o = 2.5$	132	7	5
$c_r = 1$	$c_l = 1.5$	$c_o = 1.5$	125	7	5

Table 11 Performance of rescheduling model at different unit cost levels

Unit cost			Objective values	OR overtime	Changes of surgery	OR balance
$c_o = 8$	$c_t = 4$	$c_p = 0.5$	38	2	4	12
$c_o = 4$	$c_t = 4$	$c_p = 0.5$	30	2	4	12
$c_o = 2$	$c_t = 4$	$c_p = 0.5$	26	2	4	12
$c_o = 1$	$c_t = 4$	$c_p = 0.5$	24	3	3	18
$c_o = 2$	$c_t = 8$	$c_p = 0.5$	39	5	2	26
$c_o = 2$	$c_t = 4$	$c_p = 0.5$	26	2	4	12
$c_o = 2$	$c_t = 2$	$c_p = 0.5$	18	2	4	12
$c_o = 2$	$c_t = 1$	$c_p = 0.5$	14	2	4	12
$c_o = 2$	$c_t = 4$	$c_p = 4$	40	14	3	0
$c_o = 2$	$c_t = 4$	$c_p = 3$	40	14	3	0
$c_o = 2$	$c_t = 4$	$c_p = 2$	40	14	3	0
$c_o = 2$	$c_t = 4$	$c_p = 1$	32	2	4	12

that the algorithm has better robustness. Table 11 shows the experimental results of the impact of different unit cost parameters on the rescheduling model. It can be found that within a certain range, different unit costs have little impact on the rescheduling model. But if some parameter values are set to be large enough, it will certainly have an impact on the model. For example, when the unit balance cost is much greater than the unit overtime cost, the model will try to ensure the balance of the ORs, resulting in a longer OR overtime. Nevertheless, these situations are rare because they do not meet the requirements of the hospitals. Therefore, in most cases, the rescheduling model is also relatively robust.

As depicted in Fig. 4, PS indicates that the arrival of emergency patients follows a Poisson process, and NH follows a non-homogeneous process. It can be observed that at different
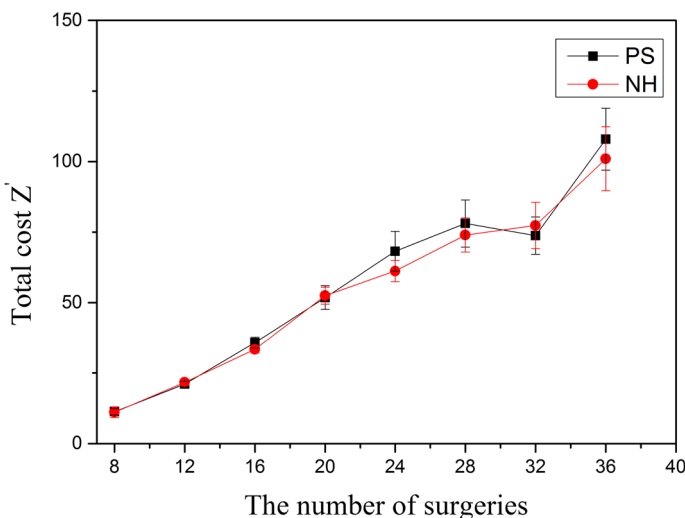

Fig. 4 Comparison between the Poisson (PS) process and the non-homogeneous (NH) process

scales, the total cost shows no significant difference in the arrival time between PS and NH, which indicates that the model can cope with the emergency patient arrivals that follow different processes as aforementioned.

6 Concluding remarks

Given the low utilization rate of operating room resources, we examine the advantages of NORA strategy and explore how to improve operating room efficiency with the implementation of NORA strategy. Then, we further consider the impact of emergency patients on elective surgeries. The research is divided into two stages. First, an initial elective surgery scheduling model under the NORA strategy is established, while a prescheduling plan for emergency patients is generated based on the BIM constraints. Then, the rescheduling model is proposed to adjust the initial scheduling plan after emergency patient arrives. In this stage, the objective is to minimize the surgical changes of the elective patients and maximize the efficiency of the operating rooms. Considering the complexity of the model, we designed two heuristic algorithms to solve the initial scheduling and rescheduling models respectively. Numerical results show that the NORA strategy can save operating room resources so that more patients can be scheduled, which becomes more obvious as the size of the operating room increases. In addition, compared to the traditional procedure for surgical scheduling, the NORA strategy greatly reduces surgery-related costs. In terms of algorithms, although the optimal solution can be obtained by applying commercial solvers, the solution time is not desirable in large-scale instances. In contrast, the proposed heuristic algorithm requires less solution time on the premise of obtaining a satisfactory solution, so that the hospital administrator can respond more quickly. The comparative analysis demonstrates that the rescheduling model, taking into account the efficiency of the operating rooms and different costs, can provide a more realistic scheduling plan.

This article verifies the advantages of the NORA strategy and enriches the research of operating room scheduling under the NORA strategy. In addition, in response to the problem of random arrival of emergency patients, a more effective rescheduling plan is proposed, which provides a decision reference for hospital managers to improve the scheduling efficiency of the ORs. In practice, hospitals must make a prompt respond to their random arrival of emergency patients for surgery. To this end, hospitals usually reserve several ORs for the sake of emergency patients. However, these ORs are often idle when emergency patients do not arrive, which is an obvious waste of OR resources. To reduce the waste, we propose a model that can both cope with the uncertainty of emergency patient arrivals and improve operating room utilization.

We shall acknowledge that further studies can be done by considering more uncertainty and the availability of downstream resources: First of all, the surgery duration of elective patients may not be able to be accurately estimated. In addition, other downstream resources may also affect the surgical process. These resources may also become a bottleneck that blocks the efficient operations of operating rooms. Readers are referred to the following references for additional details, Chang et al. (2019, 2021, 2022), Tian et al. (2022), Shi et al. (2014), Gao et al. (2020), Katehakis et al. (2015, 2016), among many others.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Al-Refaie, A., Judeh, M., & Chen, T. (2018). Optimal multiple-period scheduling and sequencing of operating room and intensive care unit. *Operational Research, 18*(3), 645–670. https://doi.org/10.1007/s12351-016-0287-0

Atighehchian, A., Sepehri, M. M., Shadpour, P., & Kianfar, K. (2020). A two-step stochastic approach for operating rooms scheduling in multi-resource environment. *Annals of Operations Research, 292*(1), 191–214. https://doi.org/10.1007/s10479-019-03353-5

Bandi, C., & Gupta, D. (2020). Operating room staffing and scheduling. *Manufacturing & Service Operations Management, 22*(5), 958–974. https://doi.org/10.1287/msom.2019.0781

Batun, S., Denton, B. T., Huschka, T. R., & Schaefer, A. J. (2011). Operating room pooling and parallel surgery processing under uncertainty. *INFORMS Journal on Computing, 23*(2), 220–237. https://doi.org/10.1287/ijoc.1100.0396

Bovim, T. R., Christiansen, M., Gullhav, A. N., Range, T. M., & Hellemo, L. (2020). Stochastic master surgery scheduling. *European Journal of Operational Research, 285*(2), 695–711. https://doi.org/10.1016/j.ejor.2020.02.001

Breuer, D. J., Lahrichi, N., Clark, D. E., & Benneyan, J. C. (2020). Robust combined operating room planning and personnel scheduling under uncertainty. *Operations Research for Health Care, 27*, 100276. https://doi.org/10.1016/j.orhc.2020.100276

Chang, A. J., El-Rayes, N., & Shi, J. (2022). Blockchain technology for supply chain management: A comprehensive review. *FinTech, 1*(2), 191–205.

Chang, J. A., Katehakis, M. N., Shi, J. J., & Yan, Z. (2021). Blockchain-empowered Newsvendor optimization. *International Journal of Production Economics, 238*, 108144.

Chang, A. C., Lu, H., & Shi, J. J. (2019). Stockout risk of production-inventory systems with compound Poisson demands. *Omega, 83*, 181–198.

Dai, Z., Wang, J. J., & Shi, J. J. (2022). How does the hospital make a safe and stable elective surgery plan during COVID-19 pandemic? *Computers & Industrial Engineering, 169*, 108210.

Devi, S. P., Rao, K. S., & Sangeetha, S. S. (2012). Prediction of surgery times and scheduling of operation theaters in optholmology department. *Journal of Medical Systems, 36*(2), 415–430. https://doi.org/10.1007/s10916-010-9486-z

Freeman, N. K., Melouk, S. H., & Mittenthal, J. (2016). A scenario-based approach for operating theater scheduling under uncertainty. *Manufacturing and Service Operations Management, 18*(2), 245–261. https://doi.org/10.1287/msom.2015.0557

Gao, L., Shi, J., Gorman, M. F., & Luo, T. (2020). Business analytics for intermodal capacity management. *Manufacturing & Service Operations Management, 22*(2), 310–329.

Guo, M., Wu, S., Li, B., Song, J., & Rong, Y. (2016). Integrated scheduling of elective surgeries and surgical nurses for operating room suites. *Flexible Services and Manufacturing Journal, 28*(1–2), 166–181. https://doi.org/10.1007/s10696-014-9199-6

Hallah, R. M., & Visintin, F. (2019). A stochastic model for scheduling elective surgeries in a cyclic Master Surgical Schedule. *Computers & Industrial Engineering, 129*, 156–168. https://doi.org/10.1016/j.cie.2019.01.030

Hamid, M., Nasiri, M. M., Werner, F., Sheikhahmadi, F., & Zhalechian, M. (2019). Operating room scheduling by considering the decision-making styles of surgical team members: A comprehensive approach. *Computers and Operations Research, 108*, 166–181. https://doi.org/10.1016/j.cor.2019.04.010

Hopp, W., & Lovejoy, W. (2013). *Hospital operations: Principles of high efficiency health care*. Pearson Education.

Jung, K. S., Pinedo, M., Sriskandarajah, C., & Tiwari, V. (2019). Scheduling elective surgeries with emergency patients at shared operating rooms. *Production and Operations Management, 28*(6), 1407–1430. https://doi.org/10.1111/poms.12993

Katehakis, M. N., Melamed, B., & Shi, J. J. (2015). Optimal replenishment rate for inventory systems with compound Poisson demands and lost sales: A direct treatment of time-average cost. *Annals of Operations Research*, 1–27.

Katehakis, M. N., Melamed, B., & Shi, J. (2016). Cash-flow based dynamic inventory management. *Production and Operations Management, 25*(9), 1558–1575.

Kroer, L. R., Foverskov, K., Vilhelmsen, C., Hansen, A. S., & Larsen, J. (2018). Planning and scheduling operating rooms for elective and emergency surgeries with uncertain duration. *Operations Research for Health Care, 19*, 107–119. https://doi.org/10.1016/j.orhc.2018.03.006

Lamiri, M., Xie, X., & Dolgui, A. (2008). A stochastic model for operating room planning with elective and emergency demand for surgery. *European Journal of Operational Research, 185*, 1026–1037. https://doi.org/10.1016/j.ejor.2006.02.057

Li, Z. P., Wang, J. J., Chang, A. C., & Shi, J. (2021). Capacity reallocation via sinking high-quality resource in a hierarchical healthcare system. *Annals of Operations Research, 300*(1), 97–135.

Li, Z. P., Wang, J. J., Perera, S., & Shi, J. J. (2022). Coordination of a supply chain with Nash bargaining fairness concerns. *Transportation Research Part E: Logistics and Transportation Review, 159*, 102627.

Liang, Y. (2022). *Implications of non-operating room anesthesia policy for operating room efficiency*. Brock University.

Lu, H., Chen, X., Shi, J., Vaidya, J., Atluri, V., Hong, Y., & Huang, W. (2020). Algorithms and applications to weighted rank-one binary matrix factorization. *ACM Transactions on Management Information Systems (TMIS), 11*(2), 1–33.

Macario, A. (2010). What does one minute of operating room time cost? *Journal of Clinical Anesthesia, 22*(4), 233–236. https://doi.org/10.1016/j.jclinane.2010.02.003

Marjamaa, R. A., Torkki, P. M., Hirvensalo, E. J., & Kirvelä, O. A. (2009). What is the best workflow for an operating room? A simulation study of five scenarios. *Health Care Management Science, 12*(2), 142–146. https://doi.org/10.1007/s10729-008-9073-8

Miao, H., & Wang, J. J. (2021). Scheduling elective and emergency surgeries at shared operating rooms with emergency uncertainty and waiting time limit. *Computers and Industrial Engineering, 160*(June), 107551. https://doi.org/10.1016/j.cie.2021.107551

Moosavi, A., & Ebrahimnejad, S. (2020). Robust operating room planning considering upstream and downstream units: A new two-stage heuristic algorithm. *Computers & Industrial Engineering, 143*, 106387. https://doi.org/10.1016/j.cie.2020.106387

Nagrebetsky, A., Gabriel, R., Dutton, R., & Urman, R. (2017). Growth of nonoperating room anesthesia care in the United States: A contemporary trends analysis. *Anesthesia & Analgesia, 124*(4), 1261–1267. https://doi.org/10.1213/ANE.0000000000001734

Rachuba, S., & Werners, B. (2014). A robust approach for scheduling in hospitals using multiple objectives. *Journal of the Operational Research Society, 65*(4), 546–556. https://doi.org/10.1057/jors.2013.112

Rachuba, S., & Werners, B. (2017). A fuzzy multi-criteria approach for robust operating room schedules. *Annals of Operations Research, 251*(1), 325–350. https://doi.org/10.1007/s10479-015-1926-1

Roshanaei, V., Luong, C., Aleman, D. M., & Urbach, D. (2017). Propagating logic-based Benders' decomposition approaches for distributed operating room scheduling. *European Journal of Operational Research, 257*(2), 439–455. https://doi.org/10.1016/j.ejor.2016.08.024

Samudra, M., Van Riet, C., Demeulemeester, E., Cardoen, B., Vansteenkiste, N., & Rademakers, F. E. (2016). Scheduling operating rooms: Achievements, challenges and pitfalls. *Journal of Scheduling, 19*(5), 493–525. https://doi.org/10.1007/s10951-016-0489-6

Schneider, A. J., Theresia van Essen, J., Carlier, M., & Hans, E. W. (2020). Scheduling surgery groups considering multiple downstream resources. *European Journal of Operational Research, 282*(2), 741–752. https://doi.org/10.1016/j.ejor.2019.09.029

Shehadeh, K. S., & Padman, R. (2022). Stochastic optimization approaches for elective surgery scheduling with downstream capacity constraints: Models, challenges, and opportunities. *Computers and Operations Research, 137*(July 2021), 105523. https://doi.org/10.1016/j.cor.2021.105523

Shi, J. (2016). Optimal continuous production-inventory systems subject to stockout risk. *Annals of Operations Research*, 1–28.

Shi, J., Yue, X., & Zhao, Y. (2014). Operations sequencing for a multi-stage production inventory system. *Naval Research Logistics (NRL), 61*(2), 144–154.

Shi, J., Zhao, Y., Kiwanuka, R. B. K., & Chang, J. (2019). Optimal selling policies for farmer cooperatives. *Production and Operations Management, 28*(12), 3060–3080.

Shylo, O. V., Prokopyev, O. A., & Schaefer, A. J. (2013). Stochastic operating room scheduling for high-volume specialties under block booking. *INFORMS Journal on Computing, 25*(4), 682–692. https://doi.org/10.1287/ijoc.1120.0530

Tian, X., Shi, J., & Qi, X. (2022). Stochastic sequential allocations for creative crowdsourcing. *Production and Operations Management, 31*(2), 697–714.

Tsai, M. H., Cipri, L. A., O'Donnell, S. E., Matthew Fisher, J., & Andritsos, D. A. (2017). Scheduling non-operating room anesthesia cases in endoscopy: Using the sandbox analogy. *Journal of Clinical Anesthesia, 40*, 1–6. https://doi.org/10.1016/j.jclinane.2017.03.010

van Essen, J. T., Hans, E. W., Hurink, J. L., & Oversberg, A. (2012). Minimizing the waiting time for emergency surgery. *Operations Research for Health Care, 1*(2–3), 34–44. https://doi.org/10.1016/j.orhc.2012.05.002

Vancroonenburg, W., De Causmaecker, P., & Vanden Berghe, G. (2016). A study of decision support models for online patient-to-room assignment planning. *Annals of Operations Research, 239*(1), 253–271. https://doi.org/10.1007/s10479-013-1478-1

Wang, J., Dai, Z., Jim, A. C., & Shi, J. (2022). Surgical scheduling by Fuzzy model considering inpatient beds shortage under uncertain surgery durations. *Annals of Operations Research*. https://doi.org/10.1007/s10479-022-04645-z

Wang, J. J., Li, Z. P., & Shi, J. J. (2021a). Hospital referral and capacity strategies in the two-tier healthcare systems. *Omega, 100*, 102229.

Wang, J.-J., Liu, H., & Shi, J. (2021b). *Robust surgical scheduling for non-operating room anesthesia (NORA) under surgical duration uncertainty*. Working paper, Dalian University of Technology.

Wu, J., Zhang, P., Wang, Y., & Shi, J. J. (2022). Integrated aviation model and metaheuristic algorithm for hub-and-spoke network design and airline fleet planning. *Transportation Research Part E, 164*, 102755.

Xie, J., Zhuang, W., Ang, M., Chou, M. C., Luo, L., & Yao, D. D. (2021). Analytics for hospital resource planning—Two case studies. *Production and Operations Management, 30*(6), 1863–1885. https://doi.org/10.1111/poms.13158

Zhang, Z., Xie, X., & Geng, N. (2014). Dynamic surgery assignment of multiple operating rooms with planned surgeon arrival times. *IEEE Transactions on Automation Science and Engineering, 11*(3), 680–691. https://doi.org/10.1109/TASE.2013.2267273

Zhu, S., Fan, W., Yang, S., Pei, J., & Pardalos, P. M. (2019). Operating room planning and surgical case scheduling: A review of literature. *Journal of Combinatorial Optimization, 37*(3), 757–805. https://doi.org/10.1007/s10878-018-0322-6

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.