



Innovative Applications of O.R.

Master surgery scheduling with consideration of multiple downstream units



Andreas Fügener^a, Erwin W. Hans^b, Rainer Kolisch^{c,*}, Nikky Kortbeek^d, Peter T. Vanberkel^e

^a *Universitäres Zentrum für Gesundheitswissenschaften am Klinikum Augsburg (UNIKA-T), School of Business and Economics, Universität Augsburg, Germany*

^b *Center for Healthcare Operations Improvement and Research (CHOIR), School of Management and Governance, University of Twente, The Netherlands*

^c *TUM School of Management, Technische Universität München, Germany*

^d *Department of Applied Mathematics, University of Twente, The Netherlands*

^e *Department of Industrial Engineering, Dalhousie University, Canada*

ARTICLE INFO

Article history:

Received 5 November 2012

Accepted 8 May 2014

Available online 20 May 2014

Keywords:

OR in health services

Resource allocation

Master surgery scheduling

Ward and ICU occupancy

ABSTRACT

We consider a master surgery scheduling (MSS) problem in which block operating room (OR) time is assigned to different surgical specialties. While many MSS approaches in the literature consider only the impact of the MSS on operating theater and operating staff, we enlarge the scope to downstream resources, such as the intensive care unit (ICU) and the general wards required by the patients once they leave the OR. We first propose a stochastic analytical approach, which calculates for a given MSS the exact demand distribution for the downstream resources. We then discuss measures to define downstream costs resulting from the MSS and propose exact and heuristic algorithms to minimize these costs.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Due to an aging society and technological progress, the demand for health care services is rising in industrialized countries (Hay, 2003; OECD Indicators, 2011). At the same time, cost cuts and human resource shortages lead to increasing pressure on hospital resources. Therefore, the importance of optimizing the usage of scarce resources in hospitals is self-evident. The most expensive resource in most hospitals is the operating room (OR) (Guerriero & Guido, 2011). ORs are clearly connected with other “downstream” resources, for example, the post-anesthesia care unit (PACU), the intensive care unit (ICU), and the general patient wards, hereafter referred to as “wards”. Anderson, Price, Golden, Jank, and Wasil (2011) show that a high level of utilization in hospital wards leads to a higher discharge rate of patients, which might reduce the quality of care. On days with high patient inflow to the ICU the danger of readmissions (Baker, Pronovost, Morlock, Geocadin, & Holzmueller, 2009) and the probability of rejected ICU requests (McManus et al., 2003) strongly increases. Therefore, downstream units should also be considered in surgery planning for medical reasons. When planning the operating rooms and the downstream units, decision makers face a trade-off between the

high complexity of a holistic view and the danger of suboptimal solutions resulting from focusing on isolated units (Vanberkel, Boucherie, Hans, Hurink, & Litvak, 2010).

Many hospitals use a so-called block-booking system when planning surgeries. In this system a medical specialty, e.g. urology, is assigned to blocks denoting a specific amount of time, e.g. a day, in one OR. These blocks can be combined into cyclical master surgery schedules (MSS), where every block is repeated after a fixed cycle, e.g. every two weeks. In planning and scheduling, problems can be categorized according to levels of a decision hierarchy (Hans, van Houdenhoven, & Hulshof, 2011): The strategic, tactical, offline-operational (i.e. planning in advance) and the online-operational (i.e. reacting/monitoring) level. In block-booking systems, decisions are made on all hierarchical levels. At the strategic level the number of blocks assigned to the specialties during a MSS cycle is determined. At the tactical level, OR-days are allocated to specialties in an MSS, such that the strategic allocation is met. At the operational level, patients are scheduled (offline) and rescheduled in case of emergencies or unexpected changes (online). An overview of OR planning may be found in Hans and Vanberkel (2011).

In the paper at hand, we discuss the tactical MSS problem, concentrating on the effect the MSS has on the patient flow to downstream inpatient care units. Surgeries performed in each block of the MSS create a flow of patients through the ICU to the ward, or directly from the OR to the wards, before they

* Corresponding author. Tel.: +49 8928925160.

E-mail addresses: andreas.fuegener@wiwi.uni-augsburg.de (A. Fügener), e.w.hans@utwente.nl (E.W. Hans), rainer.kolisch@tum.de (R. Kolisch), n.kortbeek@utwente.nl (N. Kortbeek), pvanberk@dal.ca (P.T. Vanberkel).

leave the hospital. As the PACU is part of the OR department in many hospitals, we exclude this unit in our tactical problem and denote the ICU and the ward as downstream units. This paper concentrates on the inpatient flow because outpatients leave the hospital the day of surgery and thus require only OR capacities. We define a model to calculate the distributions of recovering patients in the downstream units expected from the MSS. Based on this, we propose an approach for planning the MSS with the objective to minimize downstream costs by leveling bed demand and reducing weekend bed requests.

The remainder of this paper is organized as follows: Section 2 provides a brief overview of the relevant literature. Section 3 presents an algorithm for calculating the distribution of recovering patients in the downstream units – ICU and multiple wards. Section 4 offers a generic model to determine optimal MSSs and a discussion of relevant objective functions to determine downstream costs. In Section 5 we present a branch-and-bound algorithm and different heuristics to minimize these costs. We test the algorithms in Section 6 in an experimental investigation using data from a Dutch hospital. Finally, we discuss managerial implications, limitations, and potential extensions of our study.

2. Literature review

Operating rooms are among the most expensive resources in hospitals and is a focus of a large number of scheduling studies (Cardoen, Demeulemeester, & Beliën, 2010). For recent literature reviews on OR scheduling, see Cardoen et al. (2010) and Guerriero and Guido (2011). Articles about health care models that include both the OR and downstream units are reviewed in Vanberkel et al. (2010). In this section, we focus on articles that combine OR scheduling with the effect on downstream units, such as ICUs or wards.

Adan and Vissers (2002) present a deterministic integer programming approach to schedule patients based on fixed capacities in the OR, the ICU, and the ward. The ICU and ward capacities are the number of beds available for each specialty, while the OR capacity is the total available operating time per day. Additionally, the capacity of the nursing staff is considered. Based on this, a daily admission profile for different specialties that minimizes the deviation from resource utilization targets is obtained. Gartner and Kolisch (2014) propose a binary program which decides for each patient what day the patient is admitted, what day each clinical activity is undertaken and what day the patient is released. The objective is to maximize the sum of the contribution margins of all patients taking into account limited availability of clinical resources. Santibanez, Beliën, and Atkins (2007) discuss various trade-offs in tactical OR planning. They also apply a deterministic mixed-integer program and compare different objectives, e.g. maximizing throughput of patients or leveling the bed requests of downstream units. Their study differentiates between beds and nursing levels as well as between ORs and surgeons. An integer linear program (ILP) model to construct an MSS where patient types are assigned to blocks is formulated by van Oostrum et al. (2008). They seek to minimize the required OR capacity and to level hospital bed requirements. To incorporate the uncertainty of OR durations, they introduce probabilistic constraints. They solve the model in two steps. First, OR capacities are optimized without consideration of hospital-beds using so-called Operating Room Day Schedules (ORDSs), i.e. lists of surgery types that are assigned to one OR day. Then, the ORDSs are assigned to OR days in order to level hospital-bed demand. Therefore, leveling hospital-bed demand is only possible using the pre-computed set of ORDSs. All four aforementioned papers model

multiple downstream units with mainly deterministic approaches, while our study employs a stochastic approach.

Models for creating MSSs with leveled bed occupancy in downstream units are presented in Beliën and Demeulemeester (2007). Contrary to the articles presented above, both the number of patients and the length of stay in the hospital are assumed to be stochastic. A multinomial distribution is used to model the length of stay. The authors aim to minimize the expected bed shortage and employ a mixed-integer programming and a simulated annealing approach. The approach of Beliën and Demeulemeester (2007) differs from our approach in only allowing one downstream resource (ward), while we model the patient flow including the ICU and wards and thus consider multiple downstream units.

Min and Yih (2010) propose operational scheduling of elective surgeries that considers both uncertainty and downstream capacity constraints. They formulate a stochastic surgery scheduling problem minimizing the sum of costs directly related to patients and expected overtime costs. The downstream capacities are modeled as constraints. In contrast to their approach, which considers the operational surgery planning level, we focus on the tactical level.

Our study is based on the approach of Vanberkel et al. (2011b) where binomial distributions and discrete convolutions are used to calculate the exact distribution of recovering patients in the ward resulting from a given MSS. Vanberkel et al. (2011b) propose a set of equations to determine the distributions of ward occupancy, patient admissions, patient discharges, and the number of patients on each day of their recovery period. A case study where the algorithm is implemented in a Dutch hospital is presented in Vanberkel et al. (2011a). The authors use their approach to construct several MSSs and to choose one with a favorable ward occupancy pattern. We build upon their study by extending it in the following ways: First, Vanberkel et al. (2011b) only include one ward as a single downstream unit. As the ICU is an important bottleneck in hospitals (Litvak, van Rijsbergen, Boucherie, & Houdenhoven, 2008), we incorporate ICU bed requests as well as multiple wards in our model as a valuable extension. Second, Vanberkel et al. (2011b) do not determine the costs resulting from an MSS. As different downstream costs exist, e.g. costs for providing fixed capacities or costs for weekend staffing, we develop an approach to assign costs to specific MSSs. Third, we introduce several exact and heuristic algorithms to minimize these downstream costs.

To the best of our knowledge, the current study presents the first exact stochastic MSS approach to calculating patient occupancy distributions in the ICU and multiple wards. In addition, we present exact and heuristic algorithms to minimize costs resulting from patients in downstream units.

3. Recovering patients in downstream units

In this section, we describe a model that calculates the exact distribution of post-operative inpatients in the ICU and multiple wards resulting from a given MSS cycle. We do not further distinguish between different ICUs in this study. However, the presented approach can be extended to include several ICUs. We now present the general underlying assumptions regarding the process, the required data, and the detailed model.

After an operation several patient paths exist. In most cases, patients are admitted to a ward. In more severe cases, patients are sent to the ICU. Alternatively, patients might be discharged without being sent to a ward (e.g. due to mortality). Patients in the wards will be transferred to the ICU if their condition becomes unstable. Most patients leave the system only after recovering in a ward, but they might also leave the hospital directly from the ICU

(e.g. in case of death or if transferred to another hospital). The patient paths are outlined in Fig. 1.

In studying data from a large University Hospital in Munich, Germany, we found that more than 98% of inpatients follow one of three paths. The vast majority (92%) follow the path OR → ward → discharge. About 5% follow OR → ICU → ward → discharge. Just above 1% follow the path OR → ICU → discharge, i.e. the previous path with a zero day stay in the ward. It is very rare for patients to return to the ICU after being transferred from the ICU to a ward (just above 1%). Based on this data we simplified the modeled patient pathway as depicted in Fig. 2.

The number of patients sent to the ICU or the wards after one surgery block is modeled by a discrete empirical distribution. This distribution may also include emergency patients who were operated on during this block. A stay in the ICU is denoted by “I”, a stay in a ward of a patient who directly came from the operating room by “WO”, and a stay in a ward of a patient who was transferred from the ICU by “WI”. The lengths of stay (in days) in the ICU or a ward, after being transferred from the OR or from the ICU, are also modeled by discrete empirical distributions. Such distributions are easily obtained from historical records.

The main sets and indices used in the following model are shown in Table 1. The index n is used to determine days after surgery, where 1 denotes the day of surgery. Days after a transfer to a ward from the ICU will be denoted by u . We differentiate between multiple wards $h \in \mathcal{H}$. Each specialty j corresponds to one specific ward, whereas each ward may accommodate more than one specialty. The set of specialties accommodated by ward h is denoted by \mathcal{J}_h . Required historical or estimated data for every specialty $j \in \mathcal{J}$ are as follows:

- $a_j(p)$ represents the probability that $p \in \{0, \dots, P_j\}$ patients are operated on during a surgery block of specialty j .
- b_j represents the probability that a patient of specialty j is admitted to the ICU immediately after surgery. $1 - b_j$ is the probability that the patient is admitted to the ward.
- $c_j^I(n)$ represents the probability that a patient from surgery of specialty j stays $n \in \{1, \dots, N_j^I\}$ days in the ICU after surgery.
- $c_j^{WO}(n)$ represents the probability that a patient from surgery of specialty j stays $n \in \{1, \dots, N_j^{WO}\}$ days in the ward after surgery.
- $c_j^{WI}(u)$ represents the probability that a patient from surgery of specialty j stays $u \in \{0, \dots, N_j^{WI}\}$ days in the ward after being released from the ICU. A stay of zero days implies a direct release from the ICU.

The approach works in three steps (see Fig. 3). First, we calculate for a single surgery block the distributions of recovering patients in the ICU and the specific ward. This step is carried out for each surgical specialty. In the next step we calculate the distributions for a single cyclical block. It is important to note that we assume the MSS to be cyclical. Therefore, each block will be repeated for each new MSS cycle. In the third step we combine all blocks from a cyclical MSS to calculate the occupancy levels for the ICU and each ward $h \in \mathcal{H}$. The first two steps do not depend

on the specific MSS, we only need information about the definition of surgery blocks (e.g. length of a block) and the length of the MSS cycle. Therefore, these steps can be calculated during preprocessing. Due to the structure of the problem, the third step has to be calculated for each MSS we want to evaluate.

3.1. Calculation of the distributions of patients resulting from a single OR block (Step 1)

In the following, we present the algorithm to derive the distributions of number of patients resulting from a single OR block. First, using the probability of an ICU admission and the empirical length of stay distributions, we analyze the pathway of a single patient through the hospital (see Fig. 4). After surgery, a patient of specialty j can be admitted either to the ICU or to the specific ward h . On each day n , a patient in the ICU may either stay or be transferred to this ward. A patient in the ward may either stay or be released from the hospital. We assume that the probability for a patient to be discharged from the ward after being transferred from the ICU only depends on the time since the transfer from the ICU.

Eq. (1) calculates the conditional probability $d_{j,n}^I$ for a patient of specialty j in the ICU to be transferred to the ward on day n , given that he was not released before. Analogously, Eq. (2) calculates the probability $d_{j,n}^{WO}$ that a patient who is in the ward n days after surgery is discharged on that day. Eq. (3) calculates the probability $d_{j,u}^{WI}$ for a discharge u days after the transfer from the ICU to the ward. Patients who leave the hospital after staying in the ICU are modeled to have a stay of zero days in the ward. Calculations (1)–(3) follow the logic of Vanberkel et al. (2011b).

$$d_{j,n}^I = \frac{c_j^I(n)}{\sum_{k=n}^{N_j^I} c_j^I(k)} \quad j \in \mathcal{J}, \quad n \in \{1, \dots, N_j^I\}, \quad (1)$$

$$d_{j,n}^{WO} = \frac{c_j^{WO}(n)}{\sum_{k=n}^{N_j^{WO}} c_j^{WO}(k)} \quad j \in \mathcal{J}, \quad n \in \{1, \dots, N_j^{WO}\}, \quad (2)$$

$$d_{j,u}^{WI} = \frac{c_j^{WI}(u)}{\sum_{k=u}^{N_j^{WI}} c_j^{WI}(k)} \quad j \in \mathcal{J}, \quad u \in \{0, \dots, N_j^{WI}\}. \quad (3)$$

We denote the latest possible day with a positive probability of a patient staying in the ICU and in the ward as N_j^I and $N_j^W = \max(N_j^{WO}, N_j^I + N_j^{WI})$, respectively. Now, we calculate in Eq. (4) for all specialties $j \in \mathcal{J}$ and each day $n \in \{1, \dots, N_j^I\}$ the probabilities $e_{j,n}^I$ that a patient of specialty j who had surgery on day 1 is in the ICU. Accordingly, the same is done for patients staying in the ward ($e_{j,n}^W$ for $n \in \{1, \dots, N_j^W\}$). For the probability that a patient of specialty j is in the ICU on day n we get

$$e_{j,n}^I = \begin{cases} b_j, & n = 1 \\ (1 - d_{j,n-1}^I) e_{j,n-1}^I, & n \in \{2, \dots, N_j^I\} \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

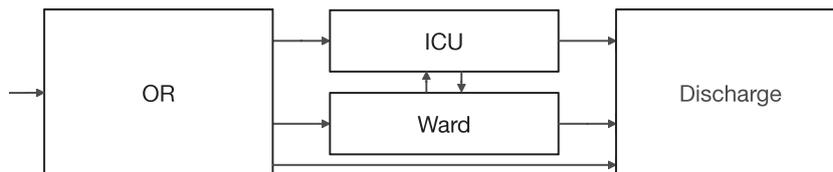


Fig. 1. Patient paths.

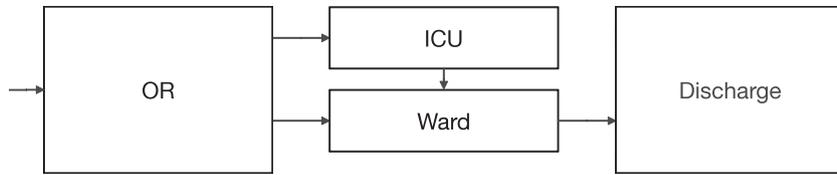


Fig. 2. Simplified patient paths.

Table 1
Sets and indices.

Description	Index \in set
Wards	$h \in \mathcal{H}$
Operating rooms (ORs)	$i \in \mathcal{I}$
Surgery specialties	$j \in \mathcal{J}$
Surgery specialties connected to ward h	$j \in \mathcal{J}_h$
Patients	$p \in \{0, \dots, P_j\}$
Days in the ICU after surgery	$n \in \{1, \dots, N_j^I\}$
Days in a ward after surgery	$n \in \{1, \dots, N_j^{WO}\}$
Days in a ward after ICU	$u \in \{0, \dots, N_j^{WI}\}$
Days in the MSS cycle	$\ell \in \mathcal{L}$
Weekdays in the MSS cycle	$q \in \mathcal{Q}$
Weekend days in the MSS cycle	$\ell \in \mathcal{L} \setminus \mathcal{Q}$

probability that the patient came directly from the OR and is in the ward on day n is denoted by $e_{j,n}^{WO}$, whereas the probability that the patient is in the ward on day n after staying m days in the ICU is $e_{j,m,n}^{WI}$.

$$e_{j,n}^{WO} = \begin{cases} 1 - b_j, & n = 1 \\ (1 - d_{j,n-1}^{WO})e_{j,n-1}^{WO}, & n \in \{2, \dots, N_j^{WO}\} \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

$$e_{j,m,n}^{WI} = \begin{cases} (1 - d_{j,0}^{WI})d_{j,m}^I e_{j,m}^I, & m \in \{1 \dots N_j^I\}, n = m + 1 \\ (1 - d_{j,n-m-1}^{WI})e_{j,m,n-1}^{WI}, & m \in \{1 \dots N_j^I\}, n \in \{m + 2, \dots, m + N_j^{WI}\} \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

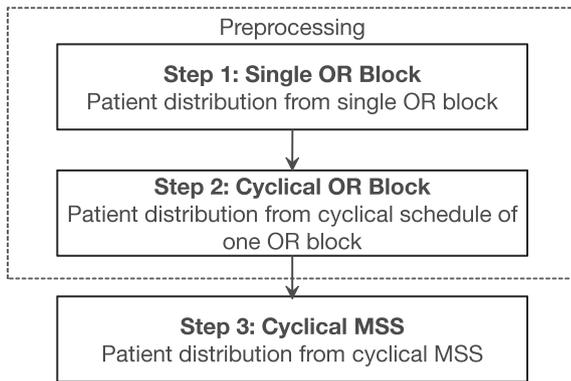


Fig. 3. Process steps.

The calculation of $e_{j,n}^{WO}$ in Eq. (5) is analogous to $e_{j,n}^I$. To calculate $e_{j,m,n}^{WI}$ in Eq. (6), the different transfer times from the ICU are taken into account. After staying m days in the ICU ($n = m + 1$), $e_{j,m,n}^{WI}$ equals the product of (a) the probability $(1 - d_{j,0}^{WI})$ that the patient did not leave the hospital immediately, (b) the probability $d_{j,m}^I$ that he was transferred to the ward that day, and (c) the probability $e_{j,m}^I$ that the patient was in the ICU on day m . Therefore, the probability $e_{j,n}^W$ that a patient is in the ward on day n is calculated in Eq. (7) by adding the probability $e_{j,n}^{WO}$ that he came directly from the OR and the probabilities $e_{j,m,n}^{WI}$ that he stayed m days in the ICU before for all possible number of days $m < n$.

$$e_{j,n}^W = \begin{cases} e_{j,1}^{WO}, & n = 1 \\ e_{j,n}^{WO} + \sum_{m=1}^{n-1} e_{j,m,n}^{WI}, & n \in \{2, \dots, N_j^W\} \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

On day 1, this probability equals b_j , i.e. the probability that the patient is directly transferred to the ICU after surgery. For the following days, the probability decreases as patients might be transferred to the ward. In order to calculate $e_{j,n}^W$, we differentiate between patients who were directly transferred to the ward after leaving the OR and those who were transferred via the ICU. The

Now, we calculate for each day n the probability distribution for the number of patients in the ICU, $f_{j,n}^I(p)$, in Eq. (8) and in the ward, $f_{j,n}^W(p)$, in Eq. (9). The probability that out of k patients who had surgery, p patients are in the ICU or the ward on day n can be determined using a binomial distribution (Vanberkel et al., 2011b).

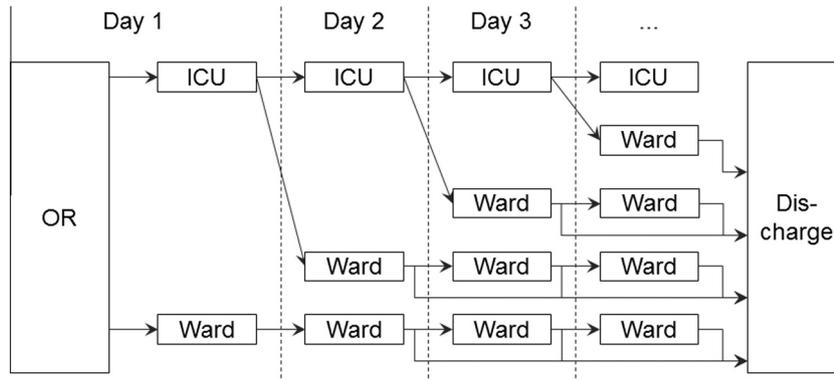


Fig. 4. Patient paths.

Next, we have to sum these probabilities weighted by $a_j(k)$ for all possible k (number of patients that had surgery) that could lead to p patients on day n .

$$f_{j,n}^I(p) = \sum_{k=p}^{P_j} \binom{k}{p} (e_{j,n}^I)^p (1 - e_{j,n}^I)^{k-p} a_j(k) \quad j \in \mathcal{J}, \quad n \in \{1, \dots, N_j^I\}. \quad (8)$$

$$f_{j,n}^W(p) = \sum_{k=p}^{P_j} \binom{k}{p} (e_{j,n}^W)^p (1 - e_{j,n}^W)^{k-p} a_j(k) \quad j \in \mathcal{J}, \quad n \in \{1, \dots, N_j^W\}. \quad (9)$$

3.2. Calculation of the distributions of patients resulting from a cyclical OR block (Step 2)

As the MSS schedule is cyclical, each block will be repeated in every cycle. For example, consider a weekly cycle in which a urological block on Monday takes place on every Monday. As the maximum recovery time of patients usually exceeds the cycle time, patients having their surgery in different cycles might be recovering at the same time. The number of overlapping cycles depends on the cycle length $L = |\mathcal{L}|$ and the maximum length of stay N_j^I for patients in the ICU and N_j^W for patients in the wards. To obtain the distributions of patients on the days of one cycle, we perform discrete convolutions, see (10) and (11), of the patient distributions of all overlapping cycles for the ICU and the specific ward, respectively. We use the symbol $*$ for the discrete convolution. $F_{j,\ell}^I$ ($F_{j,\ell}^W$) represents the distribution on the ℓ th day of a cycle of the number of recovering patients of specialty j in the ICU (ward) which results from a cyclical surgery block on day 1 of all previous cycles including the current cycle.

$$F_{j,\ell}^I = f_{j,\ell}^I * f_{j,\ell+L}^I * \dots * f_{j,\ell+\lfloor (N_j^I - \ell)/L \rfloor}^I \quad j \in \mathcal{J}, \quad \ell \in \mathcal{L} \quad (10)$$

$$F_{j,\ell}^W = f_{j,\ell}^W * f_{j,\ell+L}^W * \dots * f_{j,\ell+\lfloor (N_j^W - \ell)/L \rfloor}^W \quad j \in \mathcal{J}, \quad \ell \in \mathcal{L} \quad (11)$$

3.3. Calculation of the distributions of patients resulting from a cyclical MSS (Step 3)

To calculate a cyclical MSS, we obtain the patient distributions coming from each block (i, q) , where i denotes the operating room and q the day of the cycle. We assume that surgery blocks are only provided on weekdays. For a given MSS, x is set and each $x_{i,q,j}$ has a value of 1 if specialty j is assigned to block (i, q) and a value of 0 otherwise. $\bar{F}_{i,q,\ell}^I$ in (12) ($\bar{F}_{h,i,q,\ell}^W$ in (13)) is the distribution of the number of recovering patients in the ICU (ward h) on day ℓ of the MSS cycle coming from surgery in OR i on day q of the MSS cycle.

$$\bar{F}_{i,q,\ell}^I = \begin{cases} \sum_{j \in \mathcal{J}} F_{j,\ell-q+1}^I x_{i,q,j}, & \ell \geq q \\ \sum_{j \in \mathcal{J}} F_{j,\ell-q+1+L}^I x_{i,q,j}, & \text{otherwise.} \end{cases} \quad i \in \mathcal{I}, \quad q \in \mathcal{Q}, \quad \ell \in \mathcal{L}, \quad (12)$$

$$\bar{F}_{h,i,q,\ell}^W = \begin{cases} \sum_{j \in \mathcal{J}_h} F_{j,\ell-q+1}^W x_{i,q,j}, & \ell \geq q \\ \sum_{j \in \mathcal{J}_h} F_{j,\ell-q+1+L}^W x_{i,q,j}, & \text{otherwise.} \end{cases} \quad h \in \mathcal{H}, \quad i \in \mathcal{I}, \quad q \in \mathcal{Q}, \quad \ell \in \mathcal{L} \quad (13)$$

Now we have to convolve the distributions of all blocks to obtain the patient distribution resulting from the MSS. F_ℓ^I in (14) ($F_{h,\ell}^W$ in (15)) denotes the distribution of recovering patients in the ICU (ward h) on day ℓ of the MSS cycle. $\text{sup}\{\mathcal{I}\}$ denotes the last operating room, $\text{sup}\{\mathcal{Q}\}$ the last weekday with an active surgery slot.

$$F_\ell^I = \bar{F}_{1,1,\ell}^I * \bar{F}_{1,2,\ell}^I * \dots * \bar{F}_{\text{sup}\{\mathcal{I}\}, \text{sup}\{\mathcal{Q}\}, \ell}^I \quad \ell \in \mathcal{L} \quad (14)$$

$$F_{h,\ell}^W = \bar{F}_{h,1,1,\ell}^W * \bar{F}_{h,1,2,\ell}^W * \dots * \bar{F}_{h, \text{sup}\{\mathcal{I}\}, \text{sup}\{\mathcal{Q}\}, \ell}^W \quad h \in \mathcal{H}, \ell \in \mathcal{L} \quad (15)$$

The steps presented in this section calculate for a given MSS the distribution of patients for every day ℓ in the MSS cycle for the ICU and every ward $h \in \mathcal{H}$. Note that $x_{i,q,j}$ is assumed to be set for now but will become a variable when we are searching for a good MSS. In the next sections we present methods to minimize downstream costs of an MSS using these distributions.

4. Generic model and discussion of objectives

In Section 4.1 we present a generic model that minimizes downstream costs using a general assignment problem. We then discuss different downstream cost functions for this model in Section 4.2.

4.1. Generic model

We define a generic assignment problem that minimizes the downstream costs $c(x)$. $c(x)$ is a function of the distribution of patients in the downstream units calculated in Steps 1–3 in the previous section resulting from the MSS x , i.e. the assignment of all blocks (i, q) to a specialty j .

$$\text{Min } c(x) \quad (16)$$

s.t.

$$\sum_{j \in \mathcal{J}} x_{i,q,j} \leq 1 \quad i \in \mathcal{I}, \quad q \in \mathcal{Q} \quad (17)$$

$$\sum_{i \in \mathcal{I}} \sum_{q \in \mathcal{Q}} x_{i,q,j} \geq d_j \quad j \in \mathcal{J} \quad (18)$$

$$\sum_{i \in \mathcal{I}} x_{i,q,j} \leq s_{qj} \quad q \in \mathcal{Q}, \quad j \in \mathcal{J} \quad (19)$$

$$x_{i,q,j} \in \{0, 1\} \quad i \in \mathcal{I}, q \in \mathcal{Q}, \quad j \in \mathcal{J}. \quad (20)$$

Eqs. (17) and (18) are the assignment problem constraints. Eq. (17) ensures that at most one specialty is assigned to each block, while (18) ensures that the number of blocks assigned to each specialty at least equals the number of required blocks d_j obtained from strategic planning. The maximum number of blocks s_{qj} assigned to each specialty per day is modeled in (19). Eq. (16) represents a generic objective function. This model can easily be adjusted to deal with specific constraints, e.g. some specialties have to operate in specific ORs.

4.2. Discussion of downstream cost functions

Discussions with operating room managers indicated that there are four cost components that drive downstream costs: fixed costs, overcapacity costs, staffing costs, and additional weekend staffing costs.

Fixed costs. We consider the costs for creating and maintaining fixed capacities. We define c_h^I and c_h^W as the costs for creating and maintaining the capacity for one patient in the ICU and ward h per cycle, respectively. An example of fixed costs is the costs associated with an ICU bed. The model determines the required capacity of these resources to ensure certain service levels α^I and α_h^W . We denote $Q_\ell^I(\alpha^I)$ as the α^I -quantile of the distribution F_ℓ^I of the number of patients in the ICU on day ℓ . $Q_{h,\ell}^W(\alpha_h^W)$ is the α_h^W -quantile for the distribution $F_{h,\ell}^W$ of patients in ward h . For example, $Q(.99)$ denotes the capacity that will not be exceeded with a probability of 99%. The number of beds we need to provide in the ICU and in each ward h are therefore $\text{cap}^I(\alpha^I) = \max_{\ell \in \mathcal{L}} (Q_\ell^I(\alpha^I))$ and

$cap_h^W(\alpha_h^W) = \max_{\ell \in \mathcal{L}} (Q_{h,\ell}^W(\alpha_h^W))$, respectively. We obtain total fixed costs of

$$cost^f = c^{f,I} cap^I(\alpha^I) + \sum_{h \in \mathcal{H}} c_h^{f,W} cap_h^W(\alpha_h^W). \quad (21)$$

Overcapacity costs. Overcapacity costs are costs that incur due to requiring capacity beyond cap^I and cap_h^W . This situation occurs, for example, when patients must be transferred to ICUs or wards in other hospitals, or to a “wrong” ICU or ward, as capacity limits (depending on the service levels α^I and α_h^W) are reached. We assign costs of $c^{o,I}$ and $c_h^{o,W}$ for each patient above existing capacities per day. The expected number of these patients per day is $exc^I(\alpha^I) = \sum_{\ell=1}^Q \sum_{p=cap^I(\alpha^I)+1}^{UB^I} p F_{\ell}^I(p)$ in the ICU and $exc_h^W(\alpha_h^W) = \sum_{\ell=1}^Q \sum_{p=cap_h^W(\alpha_h^W)+1}^{UB_h^W} p F_{h,\ell}^W(p)$ in ward h . UB^I and UB_h^W is an upper bound on the number of patients that request a bed in the ICU and ward h , respectively. A simple upper bound is the product of the number of overlapping cycles and the maximum number of patients per cycle. We obtain overcapacity costs of

$$cost^o = c^{o,I} exc^I(\alpha^I) + \sum_{h \in \mathcal{H}} c_h^{o,W} exc_h^W(\alpha_h^W). \quad (22)$$

Both, the fixed costs and the overcapacity costs depend on the service levels α^I and α_h^W . The higher the service level is, the higher the fixed costs and the lower the overcapacity costs are. Setting the appropriate service levels α^I and α_h^W should be done on a strategic level and is therefore outside the scope of this study.

Staffing costs. Staffing costs are dependent on the number of patients, i.e. occupied beds. The staffing decision for every bed is made in advance. Therefore, we assume a service level of β^I and β_h^W for staffing beds. For example, a hospital might staff the .75 quantile of demand to be understaffed no more than 25% of the time. For simplicity, we assume the costs for staffing one bed per day are constant with $c^{s,I}$ for the ICU and $c_h^{s,W}$ for the wards. The total number of beds to be staffed during one cycle in the ICU and in ward h are therefore $sta^I(\beta^I) = \sum_{\ell \in \mathcal{L}} Q_{\ell}^I(\beta^I)$ and $sta_h^W(\beta_h^W) = \sum_{\ell \in \mathcal{L}} Q_{h,\ell}^W(\beta_h^W)$, respectively. The staffing costs, $cost^s$, assuming constant wages, are

$$cost^s = c^{s,I} sta^I(\beta^I) + \sum_{h \in \mathcal{H}} c_h^{s,W} sta_h^W(\beta_h^W). \quad (23)$$

Weekend staffing costs. Usually, there are additional costs for staffing beds on weekends. The additional costs for one bed per day are $c^{we,I}$ and $c_h^{we,W}$. The total number of beds to be staffed on the weekends of one cycle in the ICU and in ward h are therefore $sta^{we,I}(\beta^I) = \sum_{\ell \in \mathcal{L} \setminus \mathcal{Q}} Q_{\ell}^I(\beta^I)$ and $sta_h^{we,W}(\beta_h^W) = \sum_{\ell \in \mathcal{L} \setminus \mathcal{Q}} Q_{h,\ell}^W(\beta_h^W)$, respectively. The additional costs on weekends are

$$cost^{we} = c^{we,I} sta^{we,I}(\beta^I) + \sum_{h \in \mathcal{H}} c_h^{we,W} sta_h^{we,W}(\beta_h^W). \quad (24)$$

Many combinations of downstream costs are possible. In our case study, we employ downstream costs of $c(x) = cost^f + cost^{we}$. The resulting objective function is

$$\text{Min } c^{f,I} cap^I + \sum_{h \in \mathcal{H}} c_h^{f,W} cap_h^W + c^{we,I} sta^{we,I} + \sum_{h \in \mathcal{H}} c_h^{we,W} sta_h^{we,W}. \quad (25)$$

5. Solution approaches

The generic model presented in the previous section is a classical assignment problem. Although the generalized assignment problem is well-known to be NP-hard, there are efficient procedures, such as branch-and-bound (Ross & Soland, 1975), to solve even large instances to optimality. However, the calculation of

the objective function value is, due to the convolution of distributions, quite extensive. Hence, in addition to an optimal branch-and-bound procedure, we discuss the following two heuristic strategies to solve the master surgery scheduling problem:

1. Exact objective function and heuristic solution method.
2. Approximated objective function and exact solution method.

For strategy 1, we apply an incremental improvement heuristic, a 2-Opt heuristic, and simulated annealing. For strategy 2, we consider two approximated objective functions: the first uses expected values only, while the second employs a combination of expected values and variances. The last two approaches show some similarities to Beliën and Demeulemeester (2007), who minimize expected shortage of ward beds by linearization of their model.

5.1. Straightforward branch-and-bound

The straightforward branch-and-bound (SBB) algorithm is based on complete enumeration but avoids redundant solutions. These redundant solutions could be caused by having different combinations of the same specialties on the same day in different ORs. The algorithm fills up block after block of the MSS using a depth-first search. It assigns all blocks, i.e. combinations of days q and operating rooms i , to specialties j starting with the specialty with the lowest index. After each block of a day is assigned to a specialty, the next day is started. To avoid redundant solutions, remaining blocks on the same day will only be filled with specialties with the same or a higher index. An example for the solutions is presented in Fig. 5. Here, we show for 5 blocks (1 day with 1 OR, 2 days with 2 ORs) and 3 specialties (specialty 1 and 2 with 2 blocks each, specialty 3 with one block) all 11 possible non-redundant solutions (compared to a total of $\frac{5!}{2!2!} = 30$ solutions to assign these 3 specialties to 5 blocks). As an example, the solution 3 assigns the two required blocks of specialty 1 to OR 1 on days 1 and 3. The two required blocks of specialty 2 are assigned to operating rooms 1 and 2 on day 2, and the block of specialty 3 is assigned to operating room 2 on day 3.

After assigning a specialty to a block, the algorithm updates the distributions of patients in the ICU and the wards. A lower bound of the objective function is calculated by considering all blocks already planned leaving all blocks not yet planned empty. An upper bound is the best feasible solution obtained so far. A good first upper bound may be obtained by simulated annealing, as detailed in the following section. A partial solution is fathomed as soon as its lower bound is not strictly smaller than the upper bound. If a new feasible solution is obtained that is below the current upper bound, the upper bound is updated. We present the example of Fig. 5 with upper and lower bounds, fathoming of non-optimal solutions and the optimal solution (dark nodes with white numbers) in Fig. 6. While this method is exact, it may only be applied to small problem instances due to high computation times.

5.2. Exact objective function and heuristic solution method

Incremental improvement heuristic. The incremental improvement heuristic (IIH) is motivated by the way MSSs are altered in practice. Usually, an MSS already exists and the hospital is not willing to allow many changes, since the MSS affects many departments such as outpatient clinics. The proposed heuristic will seek the best option if only one swap of two blocks is allowed. Therefore, it will realize the swap with the maximum incremental improvement. This method may be used to show improvements for a defined maximum number of swaps.

2-Opt heuristic. We repeat the incremental improvement heuristic until no further improvement of the objective function

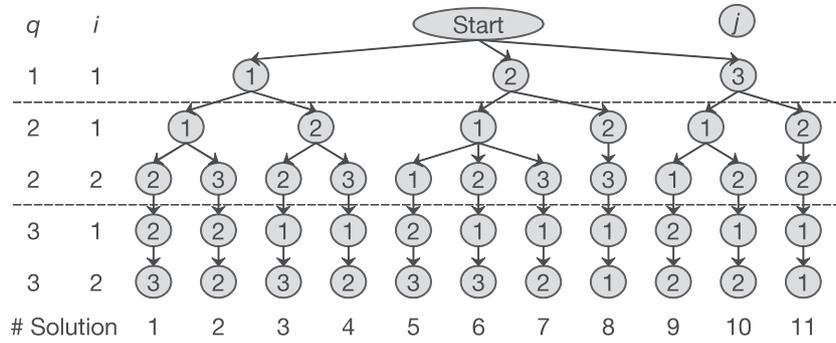


Fig. 5. All non-redundant solutions for an example with 5 blocks and 3 specialties.

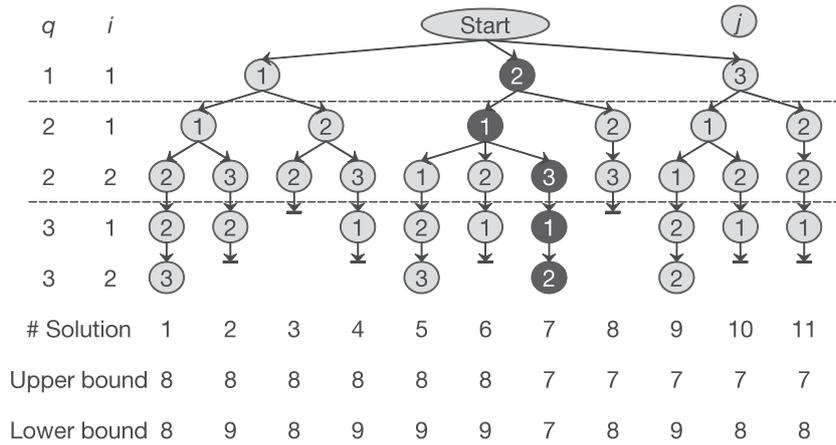


Fig. 6. Optimal solution for an example with 5 blocks and 3 specialties.

is observed. In this case it is equivalent to the 2-Opt (2OH) approach known from the traveling salesman literature (Lin & Kernighan, 1973).

Simulated annealing. IIH and 2OH presented above are likely to get stuck in a local optimum and the quality of the obtained solution is sensitive to the starting solution. To overcome this weakness we propose a simulated annealing approach (SA) (Aarts, Korst, & Michiels, 2005) with the same neighborhood as IIH and 2OH. In contrast to IIH, proximity to the given initial solution cannot be controlled with SA. SA will accept every move, i.e. swap of two blocks, that improves the objective function. A swap causing an increase in the objective function will be accepted with a probability which decreases over time. We implement the SA using a geometric cooling schedule with $t_k = cf \cdot t_{k-1}$, where cf denotes the cooling factor and t_k the temperature level in round k . The lower the cooling factor is, the faster the cool down occurs and thus the faster the SA terminates. More details of the algorithm are given in Section 6.

5.3. Approximated objective function and exact solution method

The following two solution approaches approximate the objective function, such that it can be solved with standard optimization software.

Approximated objective function based on expected values. We approximate the quantiles Q_ℓ^l and $Q_{h,\ell}^w$ used in the exact objective function by their expected values, $E(F_{j,\ell}^l)$ for the ICU and $E(F_{h,j,\ell}^w)$ for ward h , multiplied with parameters a^l and a_h^w for fixed capacities and b^l and b_h^w for weekend staffing, respectively. We

calculate these parameters as the average quotient of the quantiles $Q(\cdot)$ and the expected values $E(\cdot)$ of the given initial solution. Table 2 states the approximated quantiles. We denote the heuristic using expected values as EV.

As defined in the previous section, the objective function is

$$\text{Min } c^l \text{cap}^l + \sum_{h \in \mathcal{H}} c_h^w \text{cap}_h^w + c^{we,l} \text{sta}^{we,l} + \sum_{h \in \mathcal{H}} c_h^{we,w} \text{sta}_h^{we,w}. \quad (26)$$

where cap^l and cap_h^w denote the capacity levels, $\text{sta}^{we,l}$ and $\text{sta}_h^{we,w}$ the cumulated beds to be staffed on weekends during one MSS cycle for the ICU and ward h , respectively. For the approximation of the objective function, Constraints (27)–(29) need to be added to the generic model presented in Section 4.1. We only show the constraints for the ICU, the ones for each ward h are formulated analogously.

$$E(F_\ell^l) = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{q=1}^{\ell} E(F_{j,\ell-q+1}^l) x_{i,q,j} + \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} \sum_{q=\ell+1}^L E(F_{j,\ell-q+1+L}^l) x_{i,q,j} \quad \ell \in \mathcal{L} \quad (27)$$

$$a^l E(F_\ell^l) \leq \text{cap}^l \quad \ell \in \mathcal{L} \quad (28)$$

$$\sum_{\ell \in \mathcal{L} \cup \mathcal{Q}} b^l E(F_\ell^l) = \text{sta}^{we,l} \quad (29)$$

In (27), the values for the expected number of patients in the ICU are determined for each day. In (28) the required capacity for the ICU is calculated. Finally, in (29) the number of patients per weekend day relevant for staffing is determined.

Approximated objective based on expected values and variances. The algorithm EV neglects the distribution of patients as it only considers the expected values. With the approximated objective function based on expected values and variances (EVV), we assume the distributions of patients to be normally distributed

Table 2
Approximations for EV heuristic.

Quantile	Exact model	EV heuristic
Patients in the ICU relevant for fixed capacities	$Q_{\ell}^I(\alpha^I)$	$a^I E(F_{\ell}^I)$
Patients in the ward h relevant for fixed capacities	$Q_{h,\ell}^W(\alpha_h^W)$	$a_h^W E(F_{h,\ell}^W)$
Patients in the ICU relevant for weekend staffing	$Q_{\ell}^I(\beta^I)$	$b^I E(F_{\ell}^I)$
Patients in the ward h relevant for weekend staffing	$Q_{h,\ell}^W(\beta_h^W)$	$b_h^W E(F_{h,\ell}^W)$

Table 3
Approximations for EVV heuristic.

Quantile	Exact model	EVV heuristic
Patients in the ICU relevant for fixed capacities	$Q_{\ell}^I(\alpha^I)$	$E(F_{\ell}^I) + z^{cap,I} SD(F_{\ell}^I)$
Patients in the ward h relevant for fixed capacities	$Q_{h,\ell}^W(\alpha_h^W)$	$E(F_{h,\ell}^W) + z_h^{cap,W} SD(F_{h,\ell}^W)$
Patients in the ICU relevant for weekend staffing	$Q_{\ell}^I(\beta^I)$	$E(F_{\ell}^I) + z^{sta,I} SD(F_{\ell}^I)$
Patients in the ward h relevant for weekend staffing	$Q_{h,\ell}^W(\beta_h^W)$	$E(F_{h,\ell}^W) + z_h^{sta,W} SD(F_{h,\ell}^W)$

and approximate the quantiles using the expected value and the approximated standard deviation. To avoid a square root function, the standard deviations $SD(F_{\ell}^I)$ and $SD(F_{h,\ell}^W)$ are approximated by a linear function of the variance $V(F_{j,\ell}^I)$ for the ICU and $V(F_{j,\ell}^W)$ for the ward. We employ one linear factor for the ICU, sr^I in Constraint (31), and one for each ward h , sr_h^W . We use the factors that minimize the squared errors for the variances of the given initial solutions. The z -values for the quantiles are $z^{cap,I}$ and $z_h^{cap,W}$ for the capacity levels and $z^{sta,I}$ and $z_h^{sta,W}$ for the weekend staffing levels. Table 3 states the approximated quantiles.

The objective function (26), the assignment problem constraints (17)–(20), and the constraints to determine the expected values (27) stay unchanged. Constraints to determine the variances (30) and the approximated standard deviations (31) need to be added. The constraints determining the capacities (32) and the beds to be staffed at weekends (33) have to be changed. The approximation of the square root function to determine the standard deviation in (31) can be carried out in many ways. The most simple one is to use a linear function. To account for differences in the variances, a piecewise linear function as described in van Essen, Bosch, and van der Veen (2011) may also be applied. Again, we only present the constraints for the ICU (30)–(33), the constraints for each ward h are formulated analogously.

$$V(F_{\ell}^I) = \sum_{i \in I} \sum_{j \in J} \sum_{q=1}^{\ell} V(F_{j,\ell-q+1}^I) x_{i,q,j} + \sum_{i \in I} \sum_{j \in J} \sum_{q=\ell+1}^L V(F_{j,\ell-q+1+L}^I) x_{i,q,j} \quad \ell \in \mathcal{L} \quad (30)$$

$$sr^I V(F_{\ell}^I) = SD(F_{\ell}^I) \quad \ell \in \mathcal{L} \quad (31)$$

$$E(F_{\ell}^I) + z^{cap,I} SD(F_{\ell}^I) \leq cap^I \quad \ell \in \mathcal{L} \quad (32)$$

$$\sum_{\ell \in \mathcal{L} \cup \mathcal{Q}} E(F_{\ell}^I) + z^{sta,I} SD(F_{\ell}^I) = sta^{we,I} \quad (33)$$

6. Numerical study

We tested all solution approaches for three scenarios with an MSS cycle of two weeks, seven specialties, one ICU and two wards.

The results of these numerical experiments are described in this section. The starting point for the data collection was the data employed in Vanberkel et al. (2011b), which considers the OR and a single ward only. In order to acquire the missing data for the ICU and a second ward, we proceeded as follows. We interviewed a Dutch hospital manager responsible for patient logistics. The required values for ICU probability and the length of stay distributions for the ICU and the wards after a stay in the ICU were derived from data locally available in the ICUs and wards, respectively. We cross-checked the values with data from a German hospital with similar specialties. A summary of the data for each specialty can be found in Table 4. The downstream costs to be minimized are the fixed costs $cost^I$ and the additional weekend staffing costs $cost^{we}$. As a result of discussions with the Dutch hospital manager, we set the values of the cost parameters as presented in Table 5.

The downstream units in the case study are one ICU and two wards. The specialties Hyperthermic Intraoperative Chemotherapy (HIPEC), General Surgery and Breast Surgery share one ward, the remaining specialties share the other ward. There is no limit on the number of blocks of any specialty on any given day other than the number of ORs. To test the performance of our solution approaches on instances of varying size, we build three scenarios. In building these scenarios, we scale the MSS the hospital currently uses. By doing this, we keep the percentage of required blocks per specialty approximately constant for all three scenarios. We denote these MSSs as “hospital MSS”. We distinguish the following three examples in our case study:

- A small MSS with only one OR on every weekday of the two weeks and a second OR on Wednesdays. The number of OR blocks is therefore 12.
- A medium MSS with three ORs available during weekdays (30 blocks).
- A large MSS with nine ORs available during weekdays (90 blocks).

As a starting solution, we use the hospital MSS for all scenarios, Fig. 7 shows the hospital MSS for the medium scenario.

As we cannot compute the optimal solution for the medium and the large scenario, we use the hospital MSS as a reference point. We compare the simple branch-and-bound (SBB), the incremental improvement heuristic (IIH), the 2-Opt heuristic (2OH), simulated annealing (SA), the approximated objective function based on expected values (EV), and the approximated objective based on expected values and variances (EVV). For the IIH we continue swapping blocks until a maximum of one third of all blocks are swapped. For the simulated annealing (SA), a cooling factor of $cf = 0.9$ is chosen as proposed by Aarts et al. (2005). The number of iterations for each temperature level is 5 times the number of OR blocks for each case. The initial temperature is $t_0 = 9,000$, and the SA stops when the temperature falls below $t = 1,000$. EV and EVV are solved using CPLEX solver employing a branch-and-cut algorithm using IBM ILOG CPLEX Optimization Studio Version 12.2. All remaining heuristics and SBB were solved with MATLAB R2013a. All computations were run on a Windows-based Intel(R) Core 2 Duo CPU with 3.16 gigahertz. For each scenario we compare computation time, total cost, relative improvement of the starting solution, and the percentage of changed blocks. In order to analyze the sensitivity of computation times with respect to the cost parameter settings, we set up a ceteris paribus experimental design to evaluate variations in cost parameters. As a result we found computation times varying only slightly for realistic changes in the parameter settings.

A summary of the results of the 12 block example is provided in Table 6. In this example the IIH consists of swapping four blocks

Table 4

Specialty j , expected number of patients per OR day $E_{a_j}(p)$, probability for ICU b_j , expected LoS ICU $E_{c_j^i}(n)$, ward (after OR) $E_{c_j^{wo}}(n)$ and ward (after ICU) $E_{c_j^{wi}}(n)$.

Specialty j	$E_{a_j}(p)$	$b_j(\%)$	$E_{c_j^i}(n)$	$E_{c_j^{wo}}(n)$	$E_{c_j^{wi}}(n)$
Urology	3.2	5	5.4	3.7	6.4
HIPEC	1.0	100	20.3	11.2	21.3
Gynecology	4.3	3	3.5	2.7	4.4
General surgery	4.3	18	8.0	5.0	9.0
Breast surgery	5.5	3	2.5	2.2	3.5
Orthopedics	4.1	5	6.1	4.0	7.1
Plastic surgery	4.2	2	3.4	2.7	4.4

Table 5

Input parameters.

Description	Notation	Value
Service level for fixed capacities	α^I, α_h^W	0.99
Fixed costs ICU bed/two weeks	$c^{f,I}$	€5,000
Fixed costs ward bed/two weeks	$c_h^{f,W}$	€500
Service level for staffing	β^I, β_h^W	0.75
Additional costs staffing ICU bed at weekends/day	$c^{we,I}$	€700
Additional costs staffing ward bed at weekends/day	$c_h^{we,W}$	€120

and results in an improvement of 2.9% compared to the hospital MSS. 2OH swaps half the blocks and achieves an improvement of 3.8%. The SA takes nearly twice the computing time, changes nearly every block, but obtains a much better solution than the 2OH (8.2% instead of 3.8% in cost savings are gained). The two heuristics using an approximated objective function are very fast (around 1 second), but achieve quite different results. EV suffers from a poor approximation and only yields improvements of about 2%, while EVV achieves the highest improvement of more than 9%. EV uses 242 variables and 119 constraints, EVV 328 variables and 205 constraints. The optimal SBB demonstrates that a maximum improvement of 10.3% is possible.

Table 7 provides the results for the example with 30 blocks where all heuristics achieve comparable improvements between 6% and 7%. We stopped the optimal SBB after 25,000 seconds and achieved a comparably low improvement of 3.6%. EV and EVV are much faster than the other heuristics; EV uses 340 variables and 133 constraints, EVV 426 variables and 219 constraints. However, they change 90% and 67% of the blocks of the initial MSS.

The results of the large example with 90 blocks (see Table 8) show large computing times. EV, which employs an optimal

Table 6

Results small scenario.

Algorithm	Computation time (seconds)	Total costs	Improvement (%)	Changed blocks (%)
Hospital MSS		68,420		
SBB	19,816	61,400	10.3	92
IIH	8	66,420	2.9	33
2OH	15	65,800	3.8	50
SA	26	62,780	8.2	92
EV	1	67,400	1.5	83
EVV	1	61,780	9.7	92

Table 7

Results medium scenario.

Algorithm	Computation time (seconds)	Total costs	Improvement (%)	Changed blocks (%)
Hospital MSS		130,860		
SBB	25,000	126,120	3.6	70
IIH	187	123,000	6.0	33
2OH	336	121,800	6.9	50
SA	299	122,420	6.4	80
EV	7	122,600	6.3	90
EVV	5	122,360	6.5	67

Table 8

Results large scenario.

Algorithm	Computation time (seconds)	Total costs	Improvement (%)	Changed blocks (%)
Hospital MSS		341,020		
SBB	25,000	341,020	0.0	0
IIH	10,490	322,780	5.3	18
2OH	10,490	322,780	5.3	18
SA	2,827	315,740	7.4	81
EV	12,242	309,480	9.2	84
EVV	2,987	313,260	8.1	86

solution approach, shows the highest computational times. However, at the same time it achieves the greatest improvements. IIH and 2OH suffer from the large problem size as well, since for each swap all possibilities have to be calculated. In fact, in this

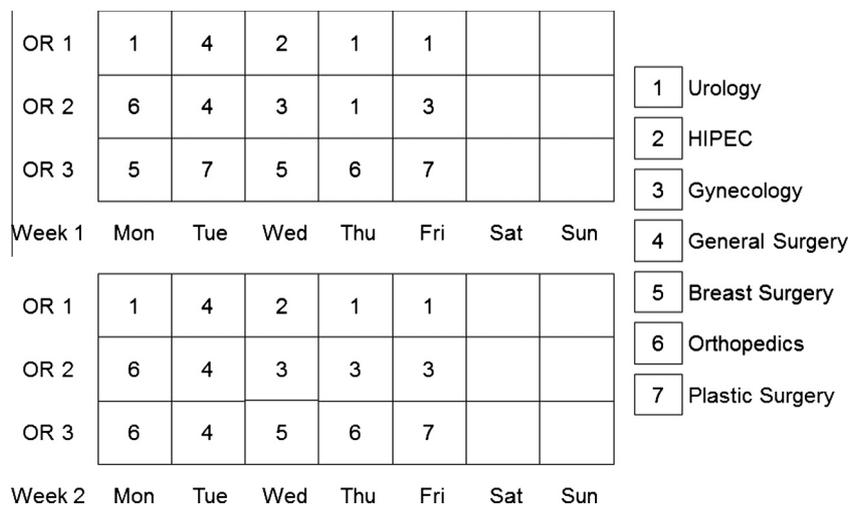


Fig. 7. Starting solution medium MSS.

example for each iteration more than 4,000 swaps have to be evaluated. Still, IIH and 2OH show the lowest improvement of only 5.3%. SA and EVV show similar computation times of less than one hour, with improvements of 7.5% and 8.1%, respectively. EV uses 929 variables and 217 constraints, EVV 1,014 variables and 303 constraints. SBB was unable to find a solution that improves the starting solution within the time limit of 25,000 seconds.

With regard to these results three conclusions can be drawn. First, the heuristics achieve improvements of around 6–9% in the three examples. Second, while IIH and 2OH achieve moderate improvements by maintaining most parts of the current schedule, they are outperformed by the SA, the EV and the EVV heuristics in most cases. Third, a majority of the approaches discussed in this paper require long computation times for the large example. Only the SA and the EVV require computation times of less than one hour. EV and EVV could be run using a time limit. Even if the optimal solution of the approximated objective function might not be found within the time limit, good results can be achieved after a relatively short time period. However, as we are discussing a tactical problem, relatively high computation times of up to a few hours can be reasonably tolerated in practical settings.

7. Conclusion

In this paper we presented an algorithm for calculating the exact distributions of patients both in the ICU and the wards resulting from a given cyclical MSS. We further discussed measures as fixed capacities and staffing levels to estimate the downstream costs of an MSS and proposed algorithms to find an MSS with the objective to minimize costs. We considered several heuristics. Two simple heuristics that swap MSS blocks, a simulated annealing algorithm that finds good solutions in a reasonable period of time, and a simple branch-and-bound, that can only be applied for small problems. We further tested solution methods approximating the objective function and solving the resulting model to optimality by using off-the-shelf solver. These showed excellent results for medium and large instances, but required long computation times for large instances.

For large instances there is further room for research on heuristics. For example, one could investigate on a combination of the approximated objective function, such as EV or EVV, with a non-optimal solution method to gain satisfactory results within short time. Furthermore, there are many possible modeling extensions. Upstream units like the outpatient clinic can be incorporated as surgeons work there too, so scheduling in both departments could be coordinated. Effects on the post-anesthesia care unit could also be incorporated. Operations on weekends (for emergency patients only) as well as pre-operative stays in ICUs and wards or patients with no surgery could be included. Moreover, for practice, relevant constraints such as differently equipped ORs and minimum time between blocks of the same specialty may be considered. The algorithms and heuristics proposed in this paper can be adapted to these extensions.

Summarizing our findings, we conclude there is significant potential in cost savings and quality improvements in considering downstream units when designing tactical operating room

schedules. Accounting for weekend staffing and leveling bed requests may further contribute to employee satisfaction and decrease negative medical effects (Baker et al., 2009).

References

- Aarts, E., Korst, J., & Michiels, W. (2005). Simulated annealing. In E. K. Burke & G. Kendall (Eds.), *Search methodologies: Introductory tutorials in optimization and decision support techniques* (pp. 187–211). New York: Springer.
- Adan, I. J. B. F., & Vissers, J. M. H. (2002). Patient mix optimisation in hospital admission planning: A case study. *International Journal of Operations and Production Management*, 22(4), 445–461.
- Anderson, D., Price, C., Golden, B., Jank, W., & Wasil, E. (2011). Examining the discharge practices of surgeons at a large medical center. *Health Care Management Science*, 14(4), 338–347.
- Baker, D. R., Pronovost, P. J., Morlock, L. L., Geocadin, R. G., & Holzmueller, C. G. (2009). Patient flow variability and unplanned readmissions to an intensive care unit. *Critical Care Medicine*, 37(11), 2882–2887.
- Beliën, J., & Demeulemeester, E. (2007). Building cyclic master surgery schedules with leveled resulting bed occupancy. *European Journal of Operational Research*, 176(2), 1185–1204.
- Cardoen, B., Demeulemeester, E., & Beliën, J. (2010). Operating room planning and scheduling: A literature review. *European Journal of Operational Research*, 201(3), 921–932.
- van Essen, J. T., Bosch, J., & van der Veen, E. (2011). *Approximating concave functions with piecewise linear functions*. Working Paper, Twente University, Netherlands.
- Gartner, D., & Kolisch, R. (2014). Scheduling the hospital-wide flow of elective patients. *European Journal of Operational Research*, 233(3), 689–699.
- Guerrero, F., & Guido, R. (2011). Operational research in the management of the operating theatre: A survey. *Health Care Management Science*, 14(1), 89–114.
- Hans, E. W., & Vanberkel, P. T. (2011). Operating theatre planning and scheduling. In R. Hall (Ed.), *Handbook of health care systems scheduling. Springer international series in operations research & management science* (Vol. 168, pp. 105–130). New York: Springer.
- Hans, E. W., van Houdenhoven, M., & Hulshof, P. J. H. (2011). A framework for health care planning and control. In R. Hall (Ed.), *Handbook of health care systems scheduling. Springer international series in operations research & management science* (Vol. 168, pp. 303–320). New York: Springer.
- Hay, J. W. (2003). Hospital cost drivers: An evaluation of 1998–2001 state-level data. *American Journal of Managed Care*, 9(1), 13–24.
- Lin, S., & Kernighan, B. (1973). An effective heuristic algorithm for the travelling salesman problem. *Operations Research*, 21(2), 489–516.
- Litvak, N., van Rijsbergen, M., Boucherie, R. J., & Houdenhoven, M. V. (2008). Managing the overflow of intensive care patients. *European Journal of Operational Research*, 185(3), 998–1010.
- McManus, M. L., Long, M. C., Cooper, A., Mandell, J., Berwick, D. M., Pagano, M., et al. (2003). Variability in surgical caseload and access to intensive care services. *Anesthesiology*, 98(6), 1491–1496.
- Min, D., & Yih, Y. (2010). Scheduling elective surgery under uncertainty and downstream capacity constraints. *European Journal of Operational Research*, 206(3), 642–652.
- OECD Indicators (2011). *Health at a glance 2011*.
- Ross, G. T., & Soland, R. M. (1975). A branch and bound algorithm for the generalized assignment problem. *Mathematical Programming*, 8(1), 91–103.
- Santibanez, P., Beliën, J., & Atkins, D. (2007). Surgical block scheduling in a system of hospitals: An application to resource and wait list management in a british columbia health authority. *Health Care Management Science*, 10(3), 269–282.
- Vanberkel, P. T., Boucherie, R. J., Hans, E. W., Hurink, J. L., & Litvak, N. (2010). A survey of health care models that encompass multiple departments. *International Journal of Health Management and Information*, 1(1), 37–69.
- Vanberkel, P. T., Boucherie, R. J., Hans, E. W., Hurink, J. L., van Lent, W. A. M., & van Harten, W. H. (2011a). Accounting for inpatient wards when developing master surgical schedules. *Anesthesia & Analgesia*, 112(6), 1472–1479.
- Vanberkel, P. T., Boucherie, R. J., Hans, E. W., Hurink, J. L., van Lent, W. A. M., & van Harten, W. H. (2011b). An exact approach for relating recovering surgical patient workload to the master surgery schedule. *Journal of the Operational Research Society*, 62(10), 1851–1860.
- van Oostrum, J. M., van Houdenhoven, M., Hurink, J. L., Hans, E. W., Wullink, G., & Kazemier, G. (2008). A master surgical scheduling approach for cyclic scheduling in operating room departments. *OR Spectrum*, 30(2), 355–374.