The operating room case-mix problem under uncertainty and nurses capacity constraints

Zakaria Yahia2 · Amr B. Eltawil1 · Nermine A. Harraz3

Received: 28 February 2015 / Accepted: 10 August 2015 © Springer Science+Business Media New York 2015

Abstract Surgery is one of the key functions in hospitals; it generates significant revenue and admissions to hospitals. In this paper we address the decision of choosing a case-mix for a surgery department. The objective of this study is to generate an optimal case-mix plan of surgery patients with uncertain surgery operations, which includes uncertainty in surgery durations, length of stay, surgery demand and the availability of nurses. In order to obtain an optimal case-mix plan, a stochastic optimization model is proposed and the sample average approximation method is applied. The proposed model is used to determine the number of surgery cases to be weekly served, the amount of operating rooms’ time dedicated to each specialty and the number of ward beds dedicated to each specialty. The optimal case-mix selection criterion is based upon a weighted score taking into account both the waiting list and the historical demand of each patient category. The score aims to maximizing the service level of the operating rooms by increasing the total number of surgery cases that could be served. A computational experiment is presented to demonstrate the performance of the proposed method. The results show that the stochastic model solution outperforms the expected value problem solution. Additional analysis is conducted to study the effect of varying the number of ORs and nurses capacity on the overall ORs’ performance.

Keywords Healthcare management · Operating rooms scheduling · Stochastic case-mix problem · Mixed integer programming · Stochastic programming · Sample average approximation

1 Introduction

Healthcare is becoming one of the largest industries in the developed and developing countries [1]. There are many challenges facing healthcare systems, such as the limited resources, the high cost of medical technology and medication, the high demand, the high customer expectations and the shortage in planning and management decision support tools specially with the increasing complexity of healthcare systems. Consequently, hospitals are more and more aware of the need to use their resources as efficiently as possible, which urges healthcare organizations to increase emphasis on process optimization in order to control and minimize operating costs and improve the provided services levels.

Operating Rooms (ORs) are important assets for hospitals as they generate the largest revenue and, at the same time, represent the largest cost source. According to “Healthcare Financial Management Association” (HFMA), ORs account for a large share of hospital care services and expenditure, at the same time, ORs result in an estimated
40% of hospital revenue [2]. Therefore, small improvements in efficiency could translate into significant savings and benefits to the patient as well as the hospital. For these reasons, scheduling of ORs in order to meet the hospital goals and carry out the operations/surgeries with maximal efficiency is one of the areas that draw considerable attention from the healthcare community. The increase in efficiency of the OR schedule has a bearing on the number of surgery cases executed, the total hospital cost or profit and utilization of resources.

The ORs scheduling decisions can be divided into three different and related problems, namely (i) the Case-Mix Problem (CMP), (ii) the Master Surgery Scheduling Problem (MSSP), and (iii) the Surgery Scheduling Problem (SSP) [3, 4].

The CMP refers to the decision of allocating the ORs’ times to each patient category in order to optimize their performance, in other words, how the available ORs time is distributed among the different specialties. In the MSS problem, the ORs time is allocated to these surgical specialties over the scheduling window (typically, one week) in order to maximize and level resource utilization. Finally, the SSP refers to assigning each surgical case a start time, a day, and an OR with the target of minimizing the waiting time and maximizing resources utilization. In this paper, we are considering the CMP which involves the decision making on surgery cases volumes and capacity requirements of each patient category. For more illustration about the CMP, an example is presented. In this example, an OR is available for eight hours a day, five days a week with total of 40 working hours per week. There are five patient categories. The question that the CMP answers is how can these 40 hours get divided between those five patient categories?. The CMP determines the number of OR hours dedicated to each patient category based on factors like: demand of each patient category, cost and profit for each patient category.

Hospitals can be classified according to the main targets of health care providers into two categories: profit satisfiers and profit maximizers [5]. Profit satisfying hospitals are motivated by professional interests rather than economic returns. A profit satisfying hospital seeks for the preferred case-mix pattern or other preset objectives on condition that it can manage to breakeven without violating the capacity constraints. On the other hand, profit maximizing hospitals are assumed to be positioned in a competitive business environment and thus they are willing to choose the patient cases that will bring the maximum rewards. Most private sector hospitals can be categorized as profit maximizing hospitals, while most government supported hospitals can be classified as profit satisfying ones.

In this paper, a stochastic optimization model for the CMP is proposed. The model generates an optimal case-mix pattern that maximizes the ORs service level. In our study, the ORs service level can be defined as the ability to satisfy the surgery demand and increase the total number of surgery cases that could be served. This model tackles two main challenges in the CMP: nurses availability capacity constraints and the uncertainty in surgery operations.

A worldwide common problem in healthcare systems is nursing staff shortages [6]. Many studies show that the availability of qualified nursing staff continues to be well below the needs of healthcare systems [7]. With nurse shortages, hospital managers are in dire need to optimally utilize current available nurses efficiently and retain them. Capacity shortage of nurses in the surgery department will keep surgery cases from being processed and it will significantly deteriorate OR utilization. Although nursing services account for an important part of a hospitals’ annual operating budget, nurse capacity constraints have received limited interest in CMP literature. To the best of our knowledge, and as indicated by Guerriero and Guido [8], Abdelrasol et al. [9] and most recently by Hof et al. [10], there is a lack in the literature regarding the CMP with uncertainty in demand, uncertain surgery operations and nurse staff capacity constraints.

Uncertainty in demand and in surgery operations result from a variety of reasons. Demand uncertainty means that it is difficult to accurately project surgery demand in the future. Larger than average surgery demands result in shortages in ORs’ resources which are reflected on the ORs’ service level. Demand uncertainty results from a variety of reasons. For instance, seasonality effects, unpredictable diseases and accidents. Furthermore, surgery operations have case-dependent durations and there is often a large variation between scheduled durations and actual durations. Length Of Stay (LOS) is also uncertain as well. Emergency surgeries are another important source that introduces more uncertainty to the problem. Uncertainty in demand and in surgery operations creates a burden on hospitals, this poses a significant challenge because it makes the mission of controlling and managing the ORs is hard.

The objective is to maximize a weighted score under resources capacity constraints. The objective function aims to maximizing the service level of the operating rooms by increasing the total number of surgery cases that could be served. The case-mix selection criterion is based upon a weighted score taking into account both the waiting list and the historical demand of each patient category. The considered objective function is a service level measure of performance rather than a financial one. Thus, it can be applied in both profit satisfying and profit maximizing hospitals planning studies.

The objective of this paper is to solve a stochastic CMP while considering nurses capacity constraints. The contributions of this study are as follows: We formulate the CMP under nurses capacity constraints as a stochastic mixed
integer programming model. Surgery durations, LOS, surgery demand and nurses daily availability are considered as random parameters with known distributions. A sampling based approach (Sample Average Approximation (SAA)) is applied to deal with the uncertainty in surgery operations and nurses capacity and to solve the proposed model.

The remainder of this paper is structured as follows. Section 2 contains a focused literature review on the CMP. The motivation for conducting this study is presented by the end of the literature review section. In Section 3, we present the proposed stochastic optimization model. In Section 4, the solution approach is introduced with description. Numerical experiments are presented in Section 5. Moreover, a numerical experiment is conducted to evaluate how the proposed stochastic model is performing in comparison with the corresponding deterministic model. Finally, concluding remarks are described in Section 6.

2 Literature review

Many review papers discussed the ORs scheduling problems. For recent reviews on this topic we refer to Cardoen et al. [11], Guerriero and Guido [8], Abdelrasol et al. [12] and the most recently Hof et al. [10]. Guerriero and Guido [8] presented a structured literature review on how Operations Research can be applied to the surgical planning and scheduling processes. They classified the research contributions by distinguishing three different decision levels, namely: strategic, tactical and operational. Abdelrasol et al. [12] reviewed the three ORs scheduling problems, they are: the CMP, the MSSP and the SSP. Furthermore, they introduced a research framework for an integrated planning method for the three problems. Most recently, Hof et al. [10] provided the first literature review focusing only on the CMP. They provided a comprehensive, structured literature review classifying CMPs according to modeling approaches, uncertainty of demand and supply, goals of using CMPs, means to achieve desired case mixes, and factors impacting the freedom of choosing case mixes. By the end of the paper, they presented possible directions for future research.

Hof et al. [10], and Guerriero and Guido [8] pointed out many references in the strategic level; however, most of these references discussed the case mix problem on the hospital level (The hospital case mix selection problem (HCMSMP)). In this study, only the literature of the CMP for the surgery patient categories is considered. The CMP literature is discussed based on three variants: (1) The models objective functions; (2) The considered resources capacity constraints; and (3) The models with uncertainty on surgery operations and demand.

2.1 The models objective functions

Hospitals can be classified according to the main targets of health care providers into two categories: profit maximizers and profit satisfiers [5]. Objective functions could be classified based on the main targets into: financial related objective functions; and service level and utilization related ones. The first category is willing to choose the case-mix plan that will bring the maximum rewards; it is most suitable for most private sector hospitals. The second category seeks for the preferred case-mix pattern or other preset objectives on condition that it can manage to break-even without violating the capacity constraints; it is most suitable for non-profit and government supported hospitals. Also, it is applicable for private sector hospital as they may target such service level and utilization oriented measures.

The majority of the publications in the CMP considered financial related objective functions. A variety of financial related performance measures for hospitals are applied such as: Maximizing the hospital revenues, Calichman [13]; maximizing the contribution margins, Ma and Demeulemeester [3] and Dexter et al. [14]; maximizing the profits, Ma et al. [15]; and minimizing the hospital total or variable costs, Dexter et al. [16]. Furthermore, achieving the preferred physicians payments and revenues beside the hospital goals has attracted some researchers such as: Blake and Carter [17]; Mulholland et al. [18]; and Kuo et al. [19]. Under stochastic demand for surgeries: Gupta [20] targeted maximizing the expected revenues depending on the allocated OR time, similarly Fügener [21]; where Dexter et al. [22] considered the mean contribution margin per OR hour in their model.

On the other hand, Testi et al. [4] solved the CMP with the objective function of maximizing the total benefits. However the total benefits looks like a financial related measure, the calculation of these benefits is more related to the service level of the system. In their study, the calculation of the benefits were depending on both: The historical demand, the number of sessions which have been assigned weekly to each patient category; and the waiting list, the number of sessions necessary to clear the waiting list for each patient category. Based on the explored literature, there is a lack in the literature regarding the CMP with service level and utilization oriented measures.

2.2 The considered resources capacity constraints

The most expensive resource in most hospitals is the OR. Furthermore, ORs are clearly connected with other resources, for example: operating staff, surgeons and nurses; the intensive care unit (ICU); and the general patient wards required by the patients once they leave the OR. The majority of the CMP literature considered the beds in wards
as capacity constraints; where, other resources are rarely considered.

A few papers considered the ICU capacity as a capacity constraint while solving the CMP. Dexter et al. [14] targeted to determine the mix of surgeons OR time allocations that maximize the contribution margin. They believed that different surgeons used differing amounts of hospital ward and ICU time. So, they solved their problem under the capacity constraints of hospital ward and ICU time. They concluded that to achieve substantive improvement in a hospital’s perioperative financial performance despite restrictions on available OR, hospital ward beds, or ICU time should be considered when OR time is allocated; Similarly, Dexter et al. [16]. Mulholland et al. [18] optimized the financial outcomes for both the hospital and physicians under the capacity constraints of general care beds, ICU beds, OR times and recovery room times. Most recently, Fügener [21] combined a strategic and a tactical OR planning problems. The developed model decided how many (strategic) and what (tactical) OR blocks to assign to each medical specialty while considering the impact on the downstream resources ICUs and general patient wards.

Furthermore, a limited publications considered the CMP under nurses capacity constraints; however, most of them are pretty old or discussed the case mix problem on the hospital level (HCMSP): Feldstein [23], Dowling [24], Hughes and Soliman [25], Robbins and Tuntiwongpiboom [26], Rifai and Pecenka [27], and Meyer et al. [28].

2.3 The CMP under uncertainty

In order to provide a more accurate representation of the ORs scheduling problem, the developed models need to incorporate more realistic aspects. In the ORs planning problems, the optimally allocated time must balance the costs of allocating too much time, which may results in idle time for ORs and staff, with the costs of allocating too little time, which may result in overtime charges. To reach this goal, uncertainty on surgery time, LOS, variation on demand and workloads should be taken into account. Most models followed a deterministic approach arguing that strategic problem uncertainties compensate each other or reserving a buffer capacity for variation and uncertainty on surgical demand. A very limited publications considered the CMP under uncertainty.

Dexter et al. [22] proposed a two-stage approach for the allocation of OR time at facilities where the strategic decision had been made to increase the number of ORs. They considered the stochastic nature of patient demand implemented uncertainty of demand in their resource allocation approaches.

Gupta [20] has addressed the surgery capacity allocation and two other different problems arising in the management of surgical suites. For each of them, a general mathematical formulation and potential solution approaches are presented. He formulated the CMP under uncertainty in surgery demand and proposed a multi-stage newsvendor model for the problem of the allocation of OR time to medical specialties. The main limitations of their work are the lack of investigation of the empirical behaviour of the proposed mathematical models and related solution approaches, and the lack of illustration of the applicability of their research thorough testing phase. However, both Gupta [20] and Dexter et al. [22] did not consider any other resources rather than the OR. They excluded other resources such as ward beds and nurses.

Most recently, Fügener [21] used stochastic patient paths to link case mix decisions with the design of an MSS while considering the stochastic downstream resource demand for the ICUs and general patient wards. Demand for these resources is modeled by a stochastic patient path model. They applied an integer programming model to provide the optimal allocation of how many (strategic planning problem) and what (tactical planning problem) OR blocks to assign to each medical specialty. However, he excluded nurses capacity from his proposed model.

As a conclusion from the CMP literature review, there are two main issues. The first one is related to the limited number of capacity constraints that has been considered. Usually such constraints were considered the ORs time, beds, and the ICU. Although nursing services account for an important part of a hospitals’ annual operating budget, nurse capacity constraints have received limited interest in CMP literature. To the best of our knowledge, and as indicated by Guerriero and Guido [8] and most recently by Hof et al. [10], there is a lack in the literature regarding the CMP under the consideration of nurse workforce capacity constraints. This emphasises the need to consider nursing staff workforce capacity constraints while representing the CMP. The second issue is related to uncertainty in demand and in surgery operations. Usually the CMP was solved by following a deterministic approach. To the best of our knowledge, and as indicated by Guerriero and Guido [8] and most recently by Hof et al. [10], there is a lack in the literature regarding the CMP with uncertainty in demand and uncertain surgery operations.

The CMP will be addressed under uncertain surgery operations, which includes uncertainty in surgery durations, length of stay, surgery demand and the availability of nurses. In order to obtain an optimal case-mix plan, a stochastic optimization model is proposed. To sum up, this paper contributes in applying the SAA approach on the CMP with the two aspects of uncertain surgery operations and nurse staff capacity. This research is motivated from the very limited literature on the CMP with those two aspects. Furthermore,
there is a practical need supporting the academic needs for conducting this research.

3 The proposed mathematical model

In this research, a stochastic mixed integer linear programming model to address the CMP with uncertainty in surgery operations and daily nurse capacity is proposed. As discussed above, the CMP is concerned with the decision of selecting an optimal case-mix under given capacities and other constraints. Even though some other factors, such as the time of day and day of week, might be important in managing the ORs, these factors are usually considered in the more detailed plans. These detailed plans are constructed in the following levels such as the MSSP and the operational SSP. The CMP is mainly a resource allocation problem. Thus, these factors are out of the scope of this paper.

The considered CMP consists of determining how ORs times are assigned to surgeon groups, how many beds are allocated to each ward and how many patients from each category can be taken care of over the course of one week as presented by Testi et al. [4]. Because the solution of the CMP was considered as a constraint for the two lower level problems, the MSSP and the SSP, the same time window was used as one week.

In this paper we use the indices, parameters and decision variables summarized in Table 1.

Three resources have to be allocated, namely, the OR time, the beds and the nurses. Let \( P \) be a set of patient categories, for each patient category \( p \in P \) we associate the integer variable \( (x_p) \) to the number of patients that receive surgery in each patient category, the real variable \( (y_p) \) to the number of OR hours dedicated to each of surgeon group, and the integer variable \( (z_p) \) to the number of beds assigned to each patient category.

The randomness in surgery operations is denoted by a scenario \( \xi \). A scenario defines the vector of outcomes of the parameters: surgery duration, LOS, surgery demand, and nurses capacity. Let \( \Phi(\xi) \) be the corresponding probability of scenario \( \xi \in \Xi \), and \( \sum_{\xi \in \Xi} \Phi(\xi) = 1 \). It should be mentioned that the probability distribution of \( \Phi(\xi) \) is discrete, and the scenario set \( \Xi \) is finite.

<table>
<thead>
<tr>
<th>Table 1 Notation summary</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indices</strong></td>
</tr>
<tr>
<td>( p \in P )</td>
</tr>
<tr>
<td>( \xi \in \Xi )</td>
</tr>
<tr>
<td><strong>Parameters</strong></td>
</tr>
<tr>
<td>( \Phi(\xi) )</td>
</tr>
<tr>
<td>( o_p )</td>
</tr>
<tr>
<td>( C_{yp} )</td>
</tr>
<tr>
<td>( 6WL_p )</td>
</tr>
<tr>
<td>( T_p )</td>
</tr>
<tr>
<td>( A )</td>
</tr>
<tr>
<td>( T_{p}^\xi )</td>
</tr>
<tr>
<td>( TT )</td>
</tr>
<tr>
<td>( L_p^\xi )</td>
</tr>
<tr>
<td>( D )</td>
</tr>
<tr>
<td>( TB )</td>
</tr>
<tr>
<td>( S )</td>
</tr>
<tr>
<td>( ND^\xi )</td>
</tr>
<tr>
<td>( H )</td>
</tr>
<tr>
<td>( BU )</td>
</tr>
<tr>
<td>( NU )</td>
</tr>
<tr>
<td>( UB_p^\xi )</td>
</tr>
<tr>
<td>( LB_p^\xi )</td>
</tr>
<tr>
<td><strong>Decision variables</strong></td>
</tr>
<tr>
<td>( x_p^\xi )</td>
</tr>
<tr>
<td>( y_p )</td>
</tr>
<tr>
<td>( z_p )</td>
</tr>
</tbody>
</table>
For each patient category \( p \in P \) and each scenario \( \xi \in \Xi \), the surgery duration and the LOS are represented by \( T_p^\xi \) and \( L_p^\xi \), respectively. Practically, surgery duration \( T_p^\xi \) depends on a procedure type and surgery team. For explicit representation, the parameter \( T_p^\xi \) can be redefined as \( T_{pq}^\xi \), where \( q \) indexes \( Q \) procedure types or \( Q \) surgery teams. However, this paper uses \( T_p^\xi \) instead of \( T_{pq}^\xi \) for simplification. Because the CMP is mainly a strategic and capacity planning problem, such details could be considered in the more detailed planning levels (MSSP and SSP). It is assumed that \( T_p^\xi \) is identical for the same patient category. \( ND^\xi \) represents the available number of nurses in the day under the scenario \( \xi \). The surgery demand randomness is represented by two parameters called the upper and lower demand limits. For each patient category \( p \in P \) and each scenario \( \xi \in \Xi \), these two parameters are represented by \( UB_p^\xi \) and \( LB_p^\xi \), respectively. Here, for each scenario \( \xi \in \Xi \), the value of the decision variable \( (x_p) \) depends on the randomness and is represented by \( (x_p^\xi) \).

The objective function in this model aims at maximizing the weighted sum of the number of surgery cases that could be served by each patient category under resources capacity constraints. We consider the current or historical allocated time and the waiting list to calculate the weighting factor for each category. For this purpose, we propose a two-step procedure to define and compute the weight for each category. In the first step, we define the patient category demand \( (O_p) \) as the sum of two components: The first component is the current/historical dedicated hours assigned weekly to patient category \( p \) \( (Cy_p) \). This term is used to eliminate the unnecessary, changes in the case-mix plan and to weight the patient categories based on their current size. This may support the applicability of the proposed plan. The second component is the required hours in order to clear the waiting list of patient category \( p \). This depends on both the average number of patients registered in the waiting list of for patient of category \( p \) \( (WL_p) \) and the average surgery duration for patient of category \( p \) \( (\bar{T}_p) \).

\[
O_p = Cy_p + WL_p \cdot \bar{T}_p \quad \forall p \in P
\]

In the second step, the weighting factor \( (o_p) \) is calculated by dividing the patient category demand \( (O_p) \) by the patient categories total demand.

\[
o_p = \frac{O_p}{\sum_{p \in P} O_p} \quad \forall p \in P
\]

Note that, the choice of calculating \( (o_p) \) has two benefits: First, the formula is intended to provide more weight for the patient categories with longer waiting lists; second, the use of \( \sum_{p \in P} O_p \) as the denominator is aimed at providing more weight to the patient categories with a bigger demand compared to the total demand. The considered weighting procedure is similar to the one proposed by Testi et al. [4]; however, they extended their procedure for each OR’s time block/session.

The following mathematical model (1)–(10) describes the stochastic case-mix problem (SCMP).

Max \[
\sum_{\xi \in \Xi} \sum_{p \in P} \Phi(\xi) \cdot o_p \cdot x_p^\xi
\]

Subject to

\[
T_p^\xi \cdot x_p^\xi \leq y_p \quad \forall p \in P, \forall \xi \in \Xi
\]

\[
\sum_{p \in P} y_p \leq TT
\]

\[
L_p^\xi \cdot x_p^\xi \leq BU \cdot z_p \cdot D \quad \forall p \in P, \forall \xi \in \Xi
\]

\[
\sum_{p \in P} z_p \leq TB
\]

\[
S \cdot \sum_{p \in P} y_p \leq NU \cdot H \cdot D \cdot ND^\xi \quad \forall \xi \in \Xi
\]

\[
x_p^\xi \leq UB_p^\xi \quad \forall p \in P, \forall \xi \in \Xi
\]

\[
x_p^\xi \geq LB_p^\xi \quad \forall p \in P, \forall \xi \in \Xi
\]

\[
x_p^\xi \cdot z_p \in \mathbb{Z}^+ \quad \forall p \in P, \forall \xi \in \Xi
\]

\[
y_p \in \mathbb{R}^+ \quad \forall p \in P
\]

The objective function (1) aims to maximize the weighted sum of the number of surgery cases that could be served by each patient category. As previously mentioned, there are three resources, namely the ORs time, beds and nurses, considered in the model. The constraints are classified into four groups: ORs time capacity constraints (2) and (3), beds capacity constraints (4) and (5), nurses’ capacity constraint (6) and demand bounds constraints (7) and (8). Constraint (2) indicates that the total required surgery time for each patient category cannot exceed the assigned hours to this patient category. Constraint (3) guarantees that the total hours dedicated to all patient categories cannot exceed the ORs’ total available working hours per week. Constraint (4) denotes that the bed occupancy for each of patient categories cannot exceed its allocated capacity. The maximum total capacity that can be allocated to a patient category is \( z_p \cdot D \). However, this is an overestimation because it does not consider the intermediate cleaning and idle time between patients. Therefore, we multiply the maximum total capacity by \( BU \). The value of \( BU \) should be selected carefully in order not to overload beds. Constraint (5) ensures that the total number of beds dedicated to all patient categories cannot exceed the total available beds.
Constraint (6) guarantees that the total required nurses’ capacity in man-hours per week cannot exceed the available number of nurses in man-hours per week. The maximum total available number of nurses in man-hours per week that can be allocated is \( H \cdot D \cdot ND \). However, this is an overestimation, because it does not consider the intermediate idle time between consecutive patients. Therefore, we multiply the maximum total capacity by \( NU \). The value of \( NU \) should be selected carefully in order not to overload nurses. Constraints (7) and (8) ensure that the obtained case-mix will be within the upper and lower expected demand respectively, and finally, constraints (9) and (10) reflect the integer and real properties of the decision variables respectively.

Under certain conditions, the proposed model can be formulated in a closed form and hence an exact solution can be obtained. These conditions include the guarantee of: small problem size, and the stochastic parameters should be following the normal or uniform distributions. If the problem size gets larger, the closed form model is intractable. Furthermore, it is well known that the distribution of surgery duration is close to a lognormal distribution [29]. While there are computational benefits, due to the limitations of the closed form expression, this paper employs the SAA approach as a solution method.

### 4 Solution method using sample average approximation

SAA [30] is a sampling based approach that can be applied to solve the SCMP (i.e. model (1)–(10)). Since we cannot directly optimize (\( \sum_{\xi \in \Xi} \sum_{p \in P} \Phi(\xi) \cdot o_p \cdot x_p^\xi \)), we instead maximize the expected value which could be written as: \( \mathbb{E}_{\xi \in \Xi} [\sum_{p \in P} o_p \cdot x_p^\xi] \). While directly computing the expected value is not possible for most problems, it can be approximated through Monte Carlo sampling in some situations. The expected value, with sample size \( N \), can be approximated by the average of the realizations:

\[
\mathbb{E}_{\xi \in \Xi} \left[ \sum_{p \in P} o_p \cdot x_p^\xi \right] \approx \frac{1}{N} \sum_{n=1}^{N} \sum_{p \in P} o_p \cdot x_p^n
\]

Based on this approximation, the objective function is deterministic, and deterministic optimization methods can be used to solve the SCMP.

SAA is applied in order to solve the SCMP model. According to Shapiro et al. [30], the optimal solution of the SAA problem provides an exact optimal solution of the true SCMP (i.e. model (1)–(10)) with probability of one for a sample size \( N \) that is large enough. Moreover, Shapiro and Homem-de-Mello [31] showed that the probability of providing an exact optimal solution of the true problem approaches one exponentially fast as \( N \) tends to infinity. The following mathematical model describes the SAA problem of the SCMP with sample size \( N \).

\[
\text{Max} \frac{1}{N} \sum_{n=1}^{N} \sum_{p \in P} o_p \cdot x_p^n \tag{11}
\]

Subject to

\[
T_p^n \cdot x_p^n \leq y_p \quad \forall p \in P, n = 1, ..., N \tag{12}
\]

\[
\sum_{p \in P} y_p \leq TT \tag{13}
\]

\[
L_p^n \cdot x_p^n \leq BU \cdot z_p \cdot D \quad \forall p \in P, n = 1, ..., N \tag{14}
\]

\[
\sum_{p \in P} z_p \leq TB \tag{15}
\]

\[
S \cdot \sum_{p \in P} y_p \leq NU \cdot H \cdot D \cdot ND^n \quad n = 1, ..., N \tag{16}
\]

\[
x_p^n \leq UB_p^n \quad \forall p \in P, n = 1, ..., N \tag{17}
\]

\[
x_p^n \geq LB_p^n \quad \forall p \in P, n = 1, ..., N \tag{18}
\]

\[
x_p^n, z_p \in \mathbb{Z}^+ \quad \forall p \in P, n = 1, ..., N \tag{19}
\]

\[
y_p \in \mathbb{R}^+ \quad \forall p \in P \tag{20}
\]

### 5 Implementation of the proposed method

#### 5.1 Case study data

In this section, an application of the proposed method based on real data collected at the surgery department in Karmoze Hospital, a non-profit hospital located in Alexandria, Egypt, is presented. The data used in this paper is collected based on database records and information derived through interviews with hospital’s surgery suites managers and staff. There are nine patient categories (|\( P | = 9 |): General surgery-A (GS-A), General surgery-B (GS-B), General surgery-C (GS-C), Gynecology, Maxillofacial, Orthopedics, Otolaryngology, Urology, and Vascular. The surgery department in Karmoze hospital consists of three surgery suites, each of them consists of three ORs, with a total of nine ORs (identical ORs). The hospital includes many wards with 302 beds but only 178 beds are dedicated to the surgery department. The operating rooms are available Saturday through Thursday for 5 hours daily (from 9:00 AM to 2:00 PM). In total, the working hours for the ORs (\( TT \)) are 270 hours per week. In order to obtain the discrete distribution of surgery duration, \( LOS \) and surgery demand the actual data of surgery procedures is compiled and analyzed from ORs database from 2010 to 2012.
In order to obtain the discrete distribution of surgery duration, the actual data of surgery procedures is analyzed. Most distributions we derived from the data analysis are consistent with empirical studies by May et al. [29] and Spangler et al. [32] which conclude that a lognormal distribution fits in actual surgery durations. We assume that all patients in the same patient category follow identical distribution of surgery duration.

The average LOS in General surgery-A (GS-A), General surgery-B (GS-B), General surgery-C (GS-C), Gynecology, Maxillofacial, Orthopedics, Otolaryngology, Urology, and Vascular are 2.5 days, 2.5 days, 1.8 days, 1.5 days, 2.5 days, 2.25 days, 1.75 days, 3.0 days and 4.8 days, respectively. In order to obtain the discrete distribution of LOS, the actual data of surgery cases is analyzed. The distributions we derive from the data analysis are either Normal or Log-normal based on patient category. However the LOS may depend on the surgery duration, in this paper we assume that the relation between both of them is an uncontrollable part on the data.

Due to the random absence and vacations of the nurses, there is a significant variation in the daily number of available nurses. According to the ORs records, the available number of nurses per day ($ND^n$) ranges between 14 and 19 according to a uniform distribution.

The historical data for each patient category was used to develop the distribution of surgery demand. Examination of the histogram in Fig. 1 and the box-plots in Fig. 2 illustrates the variations in the number of surgery cases per week. The histogram shown in Fig. 1 illustrates the variations in

![Fig. 1 Histogram for the total weekly demand](image1)

![Fig. 2 Box-plots for the weekly demand for each patient category](image2)
the total number of surgery cases for all patient categories that are handled during the week. The empirical probability distribution for the demand comes with a mean value of 252 surgery cases per week, with a standard deviation of 28 and with a lower and upper bounds of 172 and 325, respectively. This analysis emphasizes the stochastic nature of surgery demand. The Box-plots shown in Fig. 2 illustrate the variation in the number of surgery cases for each patient category. The demand on the vascular category has the lowest median and the tightest range, followed by the maxillofacial category with slightly higher median and slightly wider range. On the other hand, the orthopaedics category has the highest median and the loosest range.

5.2 Experimental results

The mathematical model (11)–(20) describes the SAA problem of the SCMP is solved optimally with the commercial ILP solver LINGO 12.0 (LINDO Systems Inc.). All tests were run on an Intel Core i5 (2.6 GHz) with 4 GB of RAM, running under Windows 7.

In order to discuss the stability and robustness of the proposed stochastic model with respect to uncertainty, a sensitivity analysis for the number of scenarios is conducted. In this analysis the trade-off between sample size \( N \) and the quality of the solution is considered. First, the SCMP is solved at different number of sample sizes, \( N = 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 150, \) and \( 200 \). At each sample size five replications are conducted \( (M = 5) \).

Figure 3 demonstrates how the objective function value of the SAA optimal solution changes as the sample size \( N \) increases. The results show the convergence of the SAA solutions with exponential rates. According to Fig. 3, a good solution can be obtained with relatively small sample size. Analysis based on the sample variance shows that the sample variance decreases abruptly at the beginning but later it levels off reaching a plateau at around 100 scenarios and declined to around zero at 150. Based on 150 scenarios and 5 replications per sample, the 95% confidence interval for the average optimal solution is less than \( \pm 1.5\% \). It is observed that 5 replications are enough to obtain a reasonable confidence interval of the average. If the variance is too large, then the value of \( M \) should be increased until satisfying an underline criterion. Based on these observations, the number of scenarios is set to 150 as an appropriate number of scenarios for the computational experiments. Evaluating more scenarios would increase the computational time with very little gain in terms of the quality of the solution.

The dotted line in Fig. 3 indicates the solution obtained by the Expected Value Parameters (EVP) of the corresponding stochastic problem. The solution of EVP is obtained by replacing the random parameters \( T^p_n \), \( L^p_n \), \( UB^p_n \), \( LB^p_n \) and \( ND^p \) by their means and then solving the resulting deterministic problem [33]. The result shows that, compared

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Effects of ORs capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORs</td>
<td>NSCs</td>
</tr>
<tr>
<td>9 ORs</td>
<td>288</td>
</tr>
<tr>
<td>8 ORs</td>
<td>288</td>
</tr>
<tr>
<td>7 ORs</td>
<td>259</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Effects of ORs capacity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurses</td>
<td>NSCs</td>
</tr>
<tr>
<td>ND</td>
<td>288</td>
</tr>
<tr>
<td>ND+2</td>
<td>303</td>
</tr>
<tr>
<td>ND+4</td>
<td>313</td>
</tr>
</tbody>
</table>
to the SAA problem, the deterministic problem produced a poor solution. Numerical experiments are performed to evaluate the value of the stochastic optimization compared to the deterministic EVP. A detailed result is shown in the next section.

As ORs’ capacity and nurses’ capacity are two of the most critical surgical resources at any operating theatre, the proposed model is solved at different levels of ORs and nurses available capacities. The purpose of this analysis is to investigate the effect of varying the availability of both ORs and nurses on the overall performance. The comparisons are conducted in terms of the Number of Surgery Cases (NSCs), the ORs’ activity index and the model objective function score.

The ORs’ activity can be defined as, the total NSCs in the optimal case-mix relative to the target NSCs over all patient categories. The target NSCs may be called the target activity and it is predefined by the hospital managers. Basically, the considered index is inspired from the product mix problem literature [34] and adapted to the CMP. This index is used in the comparison in order to show how far that the proposed case-mix plans from the targets of the hospital managers. The ORs’ activity index is calculated based on the following equation:

\[ \text{ORsActivity} = \frac{1}{A} \frac{1}{N} \sum_{n=1}^{N} \sum_{p \in P} x_{pn}^n. \]

\( A \) is the target number of surgery cases over all patient categories (surgery cases/week). According to the demand analysis and discussions with ORs managers, the denominator term of the objective function which reflects the target activity (the target number of surgery cases overall patient categories), \( A \) is assumed to be 300 surgery cases per week.

First, the effect of varying the operating rooms’ capacity is discussed. The target of this analysis is to investigate the possibility of reducing the working/open ORs, hence to reduce the cost. In regards to ORs cost rates, Macario [35] suggested for ORs administrators to use a ballpark number such as $15 to $20 per OR minute for a basic surgical procedure. Weinbroum et al. [36] estimated the cost of one OR to the health consumer or insurance carrier by approximately $10 to $20 per minute or $600 per hour. Similarly, Brodsky [37] estimated the cost of one OR by approximately $20 per minute. For the target of minimizing the operating cost and maximizing resources utilization, the system performance with lower number of working ORs is examined.

First, the model is solved at the baseline case of 9 ORs, and then it is resolved at 8 ORs and 7 ORs. Table 2 compares test results from the three different ORs’ capacities. The comparisons are conducted in terms of the NSCs, the ORs’ activity index and the model objective function score. Results show that closing one OR (8 ORs) has no impact on the NSCs that can be handled and the ORs’ activity index as well as the model objective function value. However, if we closed two OR (7 ORs), less NSCs could be processed with lower ORs’ activity index and model objective function value. From Table 2 it can be observed that ORs’ capacity limits the overall performance if more than one OR are closed.
In order to study the effect of varying the availability of nurses, the available number of nurses per day \((ND^p)\) is increased by two and four nurses. Comparisons between the current case (range of 14–19 and average of 16 nurses per day) and other cases are shown in Table 3. The comparisons are conducted in terms of the NSCs, the ORs’ activity index and the model objective function score. Results show that the overall performance is improved with the increases on the number of nurses. However, increasing the number of daily nurses by more than four is not effective.

In summary, the number of ORs can be reduced by one without any negative effect. Also, it can be seen that increasing the daily number of nurses by two and up to four enhances the overall ORs performance.

### 5.3 Evaluation of stochastic solutions

Numerical experiments are performed to compare two solutions which are obtained from the stochastic optimization problem (Stochastic) and the deterministic expected value problem (EVP). The comparison aims at evaluating the benefit of the stochastic model over its computational efforts. From the results, we claim that the stochastic model provides a better solution than EVP in terms of the robustness to the randomness in surgery operations.

The test problem presented in Section 5.1 is used for this study. The summary of the evaluation procedure is given in Fig. 4. First, samples for surgery demand \((RGD)\) for each patient category are randomly generated. Arbitrarily, the number of samples is determined as 100 \((i=1...100)\) for each patient category. Second, the amount of OR’s time required to handle the sample demand for each of patient category is estimated. Thus the demand transferred to OR’s time required capacity \((RGD_p^o \times T_p^o)\). Third, for both the stochastic model and EVP solutions, the obtained case-mix plans \((y_p)\) are defined as benchmarks. Then, the demand samples are compared against the benchmarks to check the feasibility of serving these demand samples. If the random demand sample is lower than or equal to the case-mix plan, it is feasible to serve this demand sample. This holds for each demand sample and for each patient category. Otherwise, there is a capacity shortage preventing serving this demand sample. These capacity shortages are accumulated for all patient categories in each demand sample. We denote this summation as “overcapacity cases”. The higher the overcapacity cases the worse is the case-mix plan. Finally, by the end of these comparisons, two performance measures are evaluated: the number of overcapacity occurrences and the average number of overcapacity cases. The number of overcapacity occurrences determines how many times, out of the 100 samples, shortages have occurred (the generated demand is higher than the case-mix plan for any of patient categories). The average number of overcapacity cases determines the average for the overcapacity cases in term of surgery cases (on average, by how many cases the generated demand is higher than the case-mix plan over-all demand samples). In Fig. 5, the number of overcapacity cases for 100 demand samples are presented. The dotted line and the solid line represent the number of overcapacity cases of the EVP and the stochastic model solutions, respectively. The EVP has always a higher number of overcapacity cases than the stochastic model. The two performance measures, which are presented in Table 4, indicate that the stochastic model solution outperforms the EVP. The stochastic case-mix plan reduces the number of overcapacity occurrences significantly. Additionally, reduction of the average number of overcapacity cases is identified. Upon statistical analysis of the results, it was found that the stochastic model performs better than the EVP. The mean of EVP is significantly greater than the mean of stochastic model with \(p\)-value equals 0.001 \((p\text{-value } < 0.05)\).

### 6 Conclusions and recommendations

This study proposed a stochastic optimization model for the case-mix problem with considering nurses capacity constraints. To the best of our knowledge, this is the first time to attempt this problem. The sample average approximation algorithm is employed to solve the problem. Numerical experiments demonstrated the convergence of statistical bounds with moderate sample size for a given test problem. Finally, numerical experiments were conducted to show that the stochastic model solution outperforms the Expected Value Problem solution.

The optimal case-mix selection criterion is based upon a weighted score taking into account both the waiting list and the historical demand of each patient category. The score aims to maximizing the service level of the ORs by increasing the total NSCs that could be served. The optimal amount of ORs’ time and and the optimal number of ward beds dedicated to each patient category were determined. The

<table>
<thead>
<tr>
<th>Number of overcapacity occurrences (times)</th>
<th>Average number of overcapacity per scenario (cases)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stochastic</td>
<td>34</td>
</tr>
<tr>
<td>EVP</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>21</td>
</tr>
</tbody>
</table>

---

In summary, the number of ORs can be reduced by one without any negative effect. Also, it can be seen that increasing the daily number of nurses by two and up to four enhances the overall ORs performance.
proposed model was solved under three resource capacity constraints: ORs’ time, beds and nurses. Future work will be done on building a new time table based on the proposed case-mix plan through solving the MSSP. Then testing these proposed plans in the operational level by solving the SSP. Finally trying to integrate these three models together in order to obtain a more tuned solution.

Acknowledgments This research project is sponsored by the Egyptian ministry of higher education grant and the Japanese International Cooperation Agency (JICA) in the scope of the Egypt Japan University of Science and Technology (E-JUST). The authors wish to thank the three anonymous referees for their valuable comments which have helped in improving the quality of this paper.

References