Omega 97 (2020) 102086

Contents lists available at ScienceDirect

Omega

journal homepage: www.elsevier.com/locate/omega

Assessing the impact of uncertainty and the level of aggregation in case mix planning☆

Sebastian McRae*, Jens O. Brunner

University Center of Health Sciences at Klinikum Augsburg (UNIKA-T), Faculty of Business and Economics, University of Augsburg, Universitätsstraße 16, Augsburg 86159, Germany

ARTICLE INFO

Article history:
Received 27 June 2018
Accepted 5 July 2019
Available online 10 July 2019

Keywords:
Healthcare management
Stochastic case mix planning
Resource allocation
Operating room scheduling
Framework
Integer programming
Simulation

ABSTRACT

In this article, we present a framework for evaluating the impact of uncertainty and the use of different aggregation levels in case mix planning on the quality of strategic decisions regarding the case mix of a hospital. In particular, we analyze the effect of modeling (i) demand, (ii) resource use, and (iii) resource availability as stochastic input parameters on the performance of case mix planning models. In addition, the consequences of taking the weekly structure with inactive days without surgeries into account are assessed (iv). The purpose of this paper is to provide a guideline for the decision-maker planning the case mix on the consideration of stochastic aspects and different aggregation levels. We formulate a mixed integer programming model for case mix planning along with different stochastic and deterministic extensions. The value of the different extensions is analyzed using a factorial design. The resulting stochastic models are solved using sample average approximation. Simulation is used to evaluate the strategies derived by the different models using real-world data from a large German hospital. We find that highly aggregated basic case mix planning models can overestimate the objective value by up to 10% and potentially lead to biased results. Refining the problem decreased the gap between projected case mix planning results and simulated results considerably and led to improved solutions.

1. Introduction

Hospital care in Western countries must balance adequate patient care and the economic provision of such services. An efficient utilization of resources is necessary to fulfill both goals simultaneously. In recent years, many Western countries have reacted to this challenge by introducing case-related reimbursement systems for hospitals based on diagnosis related groups (DRGs). In such systems, hospitals have an incentive to treat patients efficiently, since the reimbursement for patients only depends on the diagnosis but not on individually generated costs [1]. In addition, hospital providers have an incentive to focus on services generating positive contribution margins. May et al. [2] emphasize that attracting desirable patient groups should be carefully balanced with the availability of services for persons in need.

The problem of determining the optimal mix of patients is referred to as the case mix planning problem. Case mix planning is highly relevant for a sustainable economic performance of hospitals that are reimbursed according to DRGs. Case mix planning is most often seen as the first phase of a sequential approach to periodic hospital resource planning and scheduling [3,4]. The target volumes derived by case mix planning, along with corresponding resource allocation schemes, are used to feed downstream problems such as master surgery scheduling and admission planning problems, that set the frame for scheduling individual patients.

Different approaches to case mix planning addressing different aspects have been proposed in the literature. Potentially, the consideration of stochastic influences and the incorporation of downstream planning problems can improve model outcomes. However, it is still unclear, which aspects add value to modeling case mix planning problems and which only increase the complexity of the problem.

The main contribution of this paper is (i) to identify the kinds of uncertainties to consider when planning the case mix and (ii) to consider whether case mix planning should be solved simultaneously with the downstream problem of assigning surgical resources to different days. We present a framework for evaluating these aspects, consisting of an optimization phase and a phase of operational evaluation. In the first phase, case mix planning approaches are formulated as stochastic mixed integer programs which are solved using sample average approximation (SAA). The case mix strategies derived by these models are evaluated in the second

☆ This manuscript was processed by Associate Editor O. Prokopyev.

* Corresponding author.
E-mail address: sebastian.mcrae@unikat.uni-augsburg.de (S. McRae).

https://doi.org/10.1016/j.omega.2019.07.002

phase, on an operational level, using simulation. The simulation is an integral part of the framework since it helps to assess the impact of different case mix planning approaches which would otherwise have to be analyzed using costly field studies or pilot projects. In addition, we introduce a common notation scheme suitable for comparing current case mix planning models. The framework is applied using data of a large German hospital.

The remainder of this paper is organized as follows. Relevant literature is discussed in Section 2. In Section 3, a framework providing a blueprint for modeling case mix planning problems is presented. It is used to discuss general modeling aspects and to classify different modeling approaches to case mix planning problems described in the literature. The framework is deployed for solving the problem of determining the optimal case mix of a large German hospital in Section 4. In Section 5, findings are summarized and suggestions for future research are given.

2. Literature

Case mix planning involves long-term planning of patient volumes, typically for a year or longer. Simultaneously including all downstream decisions and stochastic influences when planning the case mix is not a realistic option, due to the highly complex nature of processes within the hospital. Therefore, case mix planning models provide an aggregated view on long-term patient and resource planning and only address a small subset of complicating aspects at one time. A comprehensive literature review on case mix planning is given by Hof et al. [5]. Recent research addresses the problem of assessing the impact of economies of scale and scope on the optimal case mix of a hospital [6].

One stream of literature with increasing importance during the last years focuses on the incorporation of stochastic aspects. The case mix planning problem with stochastic demand can be interpreted as a multi-product newsvendor problem with limited capacity. Solution procedures for such problems are discussed in several papers [7–11]. Analytical results for the multi-product newsvendor problem with a single side constraint are presented by Erlebacher [12]. A review of multi-product newsvendor problems is given by Turken et al. [13]. Khouja [14] and Qin et al. [15] provide further reviews on newsvendor problems. Olivares et al. [16] present a general approach to estimate overage and underage costs in newsvendor-type decisions and apply it to the problem of reserving operating room time for surgical cases. Dexter et al. [17] model a case mix problem with stochastic demand assuming a uniform distribution of weekly demand arguing that this is a convenient representation of uncertain demand with no further knowledge. They state that it is sufficient to have a vague idea of the distribution for annual resource planning since resource allocations can be adjusted on the operational level. Based on the previous paper, Gupta [18] presents a non-linear programming solution approach to identify the optimal case mix for any kind of demand distribution. His model formulation adds the consideration of salvage values. Additionally to the allocation of operating room capacities, Choi and Wilhelm [19] incorporate decisions regarding the number of surgeries per day to be scheduled for each operating room. They model surgery duration as a stochastic parameter which adds another dimension of stochastic influences. Further, they include a penalty for the number of patients that cannot be accommodated due to capacity shortages. They discuss different adaptations of the model along with corresponding solution methods since the original model is too complicated to be solved with standard software. Yahia et al. [20] simultaneously analyze the allocation of operating room and bed capacity. They include variability in the demand, length of stay, duration of surgery, and availability of nurses. They present a scenario approach using SAA to solve the stochastic problem. They use historical data to estimate the distribution of weekly demand. Fügener [21] uses stochastic patient paths to model different pathways of treatment to connect the case mix planning problem with the master surgery schedule (MSS). Uncertainty of objective function coefficients is discussed by Dexter et al. [17] in a preprocessing step to exclude surgeons with a very high variability of contribution margins from the allocation of additional operating room time. Leeftink and Hans [22] propose a case mix classification scheme which focuses on the surgery duration relative to the length of the operating room block time and the coefficient of variation of the surgery duration arguing that these are the two predominant parameters characterizing the case mix when scheduling surgeries.

Freeman et al. [23] use a solution pool approach to address different kinds of uncertainties in case mix planning. A solution of the pool is assessed as follows. In a preprocessing step, nonelective patient demand is simulated to determine buffer capacity. In an iterative procedure, elective demand is drawn from a random distribution and a hierarchical goal program is used to assign operating room blocks to different specialties on the different days of the planning horizon. In the last step of each iteration, the solution is evaluated on an operational level simulating patient arrivals, surgery duration, and length of stay. The decision maker can then select a solution from the solution pool based on different performance criteria. In a supplementary experiment, they evaluate the value of using simulated operating room times instead of expected values as input for the mathematical program. They find that this modification can potentially diversify the set of solutions in the solution pool. Randomly generating elective demand or operating room times can be compared to scenario generation procedures used in SAA approaches. The major benefit of using a large pool of potential solutions is the possibility of evaluating them on an operational level, taking further complicating aspects into account. Decision makers can thus make decisions on the case mix based on different solutions that meet certain performance criteria. However, this comes at the cost of potentially using suboptimal case mix strategies.

Testi et al. [24] also use a hierarchical approach for operating room planning. In the first step, the case mix planning phase, they assign operating room blocks to different wards. The solution is then translated into an MSS that determines the assignment of operating room blocks to wards on specific days of the planning horizon. In the third step, patient arrivals, the surgery duration, and the length of stay of individual patients are simulated to compare the current MSS with the MSS resulting from the optimization models. This simulation is further used to identify promising admission rules for elective patients.

A similar hierarchical approach is formulated by Ma and Demeulemeester [25]. At the first phase of case mix planning, they roughly determine the assignment of operating room blocks to surgeon groups and beds to wards on different days of the planning horizon. The resulting solution is then translated into a finer MSS minimizing the expected total bed shortage to level the resulting bed occupancy. They simulate the operating theater using stochastic patient arrivals and stochastic length of stay. They assume that a surgery is not canceled once the patient is scheduled arguing that patients are placed in alternative wards if there are no available beds in the appropriate ward. They use deterministic surgery durations in their simulation arguing that the surgery duration does not influence patient admissions and that they focus on the evaluation of misplaced assignments of patients to wards instead of assessing underutilization or overutilization of operating rooms. Both of the previous approaches do not explicitly model any stochastic influence in the phase of case mix planning.

Table 1 summarizes the approaches that address at least one of the following aspects: (i) stochastic demand, (ii) stochastic resource requirements, (iii) stochastic supply, (iv) time-phased

Table 1
Classification of complicating problem characteristics considered at the level of case mix planning (CMP).

Reference	Aspects considered at the level of CMP				Solution approach	
	Stochastic demand	Stochastic resource requirements	Stochastic supply	Time-phased allocation	CMP solution approach	Operational evaluation
Choi and Wilhelm [19]	✓	✓	–	–	Heuristics/SAA	–
Dexter et al. [17]	✓	–	–	–	Non-linear programming	–
Freeman et al. [23]	✓	✓	–	✓	Solution pool	✓
Fügener [21]	–	✓	–	✓	Piecewise linear approx.	–
Gupta [18]	✓	–	–	–	Non-linear programming	–
Ma and Demeulemeester [25]	–	–	–	✓	Integer programming	✓
Testi et al. [24]	–	–	–	–	Integer programming	✓
Yahia et al. [20]	✓	✓	✓	–	SAA	–
Our approach	✓/ –	✓/ –	✓/ –	✓/ –	SAA	✓

resource allocation, and (v) operational evaluation. It complements the summaries of literature on case mix planning presented in Hof et al. [5] and Freeman et al. [23]. The literature classified in this table serves as the basis for the framework discussed in Section 3. There exist frameworks for resource planning in healthcare [26,27]. However, these frameworks provide a more general perspective or focus on topics other than case mix planning.

As can be seen in the table, all different aspects are considered in a subset of papers. There is no setting described in the literature considering all aspects simultaneously. Note that a check mark in the table columns concerning the different complicating aspects relates to an explicit modeling in the phase of case mix planning. For example, Testi et al. [24] address the problem of time-phased planning in a subsequent phase of their three-phase approach but not during the phase of case mix planning. The various combinations of complicating modeling aspects used in the literature raise the question on the necessity and relevance of incorporating the single aspects. Simplified models potentially increase the degree of transparency and hence acceptance. There is a trade-off between added value and increased complexity of the problem with respect to data requirements and computational complexity; this is the focus of our work. In this paper, we describe a framework where each complicating aspect can be integrated if it provides value.

Summarizing, case mix planning approaches can be used to identify targets for patient volumes along with corresponding allocation schemes of resources to organizational units within a hospital, such as departments, wards, or similar. Major differences exist between the different case mix planning approaches in terms of sources of uncertainty and the temporal granularity. Nothing in the literature answers the question of which types of complicating aspects add value to case mix planning and how the choice of complicating modeling aspects should be prioritized.

3. Methodological framework

Case mix planning is the problem of identifying the optimal composition and volume of patients in a hospital [5]. It can be seen as the first phase of a multiple phase resource planning approach for hospitals [24]. In this regard, it can be used to formulate case mix goals and to allocate resources to organizational units within a hospital. The relationship between resource allocation schemes and the number of patients that can be served is very complex due to the stochastic nature of demand and supply and the complexity of the processes. Case mix planning is similar to aggregated planning approaches in manufacturing such as rough cut capacity planning. Major differences between models for resource planning in hospitals and for resource planning in manufacturing include: (i) Perishability: services in hospitals are created and consumed simultaneously, while it is in general possible to stock goods in manufacturing. (ii) Heterogeneity: resource requirements in hospitals vary from patient to patient necessitating the use of models that are different from those applied in highly standardized mass production.

An open challenge in modeling case mix planning problems concerns the choice of an appropriate level of detail. Capacity planning on a more aggregated level increases the robustness of the results. This observation is noted by Wortman et al. [28] when regarding solution approaches for the rough cut capacity planning problem. Computational issues are another major criterion to opt for aggregated model formulations due to the high complexity of the processes. Potential advantages of using more detailed models are increases in accuracy leading to improved results. Indeed, such formulations regularly outperform aggregated model formulations by construction if the same set of assumptions is used for the evaluation as is used for the construction of the model. Such reasoning can be biased, however, since additional complicating factors not

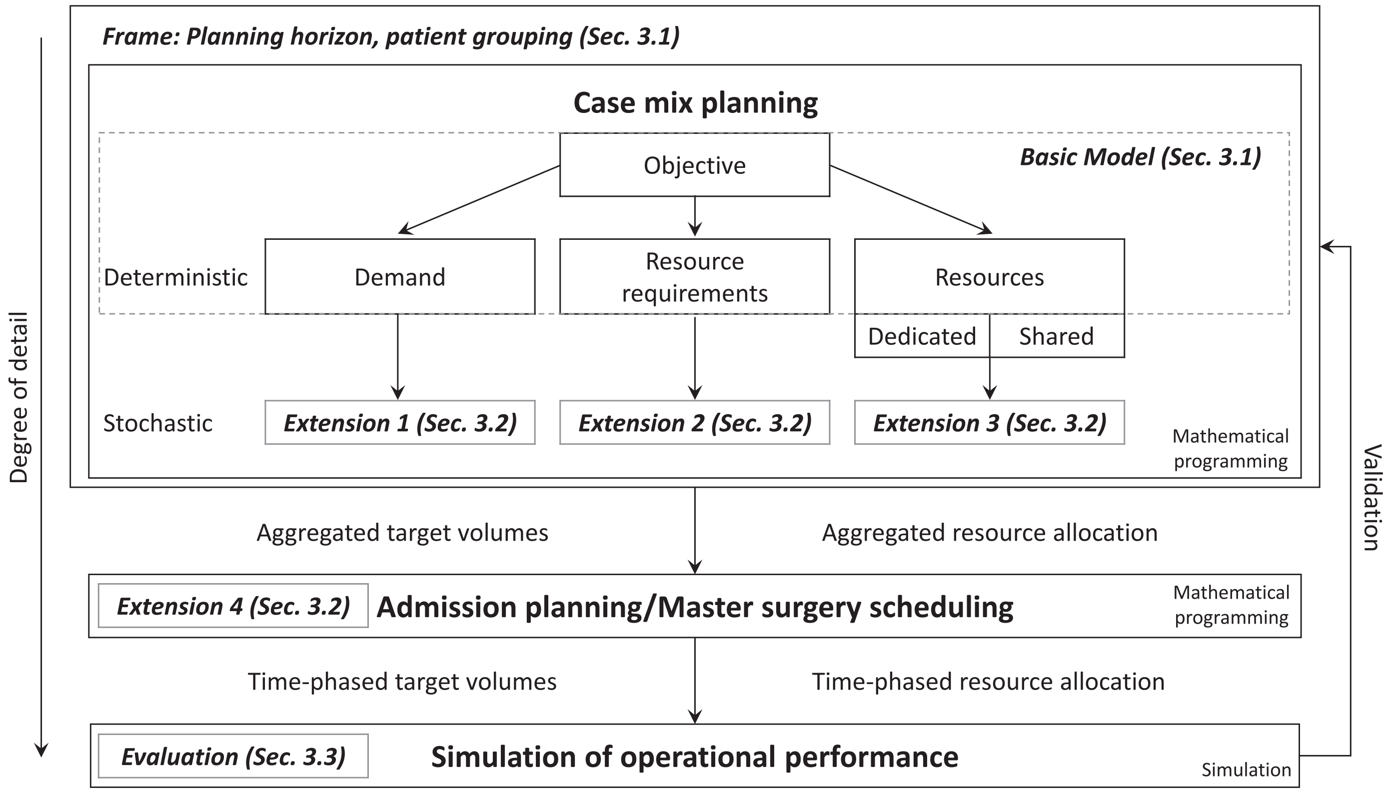

Fig. 1. Framework for case mix planning.

considered in the model can potentially lead to situations where aggregated model formulations outperform more detailed models. Evaluations via simulation realistically reflecting the hospital processes can be used to objectively assess the performance of case mix planning models.

In this section, we present a framework to address these issues by identifying appropriate case mix planning problem formulations that depend on the individual problem setting of a hospital. The framework is depicted in Fig. 1. It is described in three sections and illustrated by case mix planning models presented in the literature. In Section 3.1, considerations related to the frame for case mix planning problems are discussed. In addition, a basic deterministic formulation of the case mix planning problem is presented. Specific complicating aspects when planning the case mix are characterized in terms of model extensions in Section 3.2. Aspects not addressed in the case mix planning phase, but deemed to be potentially relevant for describing hospital processes can be considered in a phase where the performance of hospital operations is simulated. This phase is described in Section 3.3.

3.1. General considerations in modeling case mix planning problems

The modeling aspects discussed in this section concern general issues to be dealt with when planning the case mix. The frame for the case mix problem is set by defining the planning horizon and the definition of patient groups. The case mix planning problem can be interpreted as selecting a mix that maximizes some objective function. Demand and available resources constrain the mix. Resource requirements of patients are used to describe the dependency between patient volumes and resources.

3.1.1. Planning horizon

Case mix planning and related resource allocation decisions typically involve adjustments in staffing levels of qualified personnel and in the organizational structure and are thus typically planned for periods such as one year or longer [17,19,24,29].

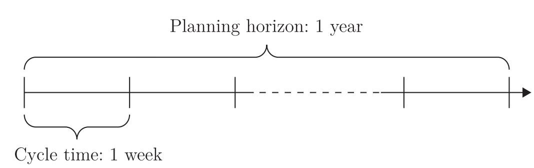

Fig. 2. Example for a cyclic planning approach.

Aggregating patient demand and resource requirements over the whole period and thereby neglecting fluctuations within the planning period can give rise to problems on an operational level. Instead, cyclic planning approaches can be used to model such fluctuations. In cyclic planning, the planning horizon is divided in smaller repeated time periods as illustrated in Fig. 2. The length of such a time period is referred to as the cycle time [30]. Typically, similar periods with a cycle time of one or two weeks are used in stochastic case mix planning approaches [19,20,23]. The cycle time should be chosen deliberately, especially if due times vary between different patient groups as described by Vansteenkiste et al. [31].

Summarizing, the case mix planning model is more accurate when the planning horizon is subdivided into different time periods suitable to model fluctuations.

3.1.2. Patient grouping

Case mix planning decisions can be made at the level of highly aggregated patient groups such as departments or major diagnostic categories (MDCs) or at the level of individual surgeons, DRGs, or procedures.

Highly aggregated patient groups should be considered if there is no potential to influence patient volumes on detailed levels [5]. Roth and Van Dierdonck [26] suggest aggregating patients at the level of MDCs for aggregated planning purposes arguing that this

Table 2
Classification of allocation policies.

Allocation policy		Example for allocation during cycle time of 7 days						
		Mon	Tue	Wed	Thu	Fri	Sat	Sun
Dedication (distinct allocation)	Ward bed	Dept1	Dept1	Dept1	Dept1	Dept1	Dept1	Dept1
Time-sharing (time-phased allocation)	Operating room	Dept1	Dept2	Dept2	Dept1	Dept1	–	–
Sharing (on request)	ICU bed	All departments						

enables the coordination with other strategic activities such as marketing. They emphasize the close relationship between MDCs and hospital specialties. Highly aggregated patient groups can be expected to be heterogeneous with respect to resource use. Grouping patients with more similar resource usage and contributions to the profit potentially leads to better model results due to a lower variability within the patient groups. However, such modeling is only useful if the case mix can be influenced on such a level.

Hybrid model formulations with respect to the granularity of patient groups can be used. Assume that a hospital is only able to influence patient volumes at the level of MDCs and that the shares of detailed resource related groups are largely stable within an MDC. Under these circumstances, a hybrid model formulation making use both of MDCs and resource related groups allows for a tighter scheduling of patients due to more diverse treatments and lower variability within the resource related groups.

3.1.3. Objectives and key performance indicators (KPIs)

Due to the high correlation of the case mix with the revenue of a hospital, goals for case mix planning often address financial aspects [5]. A maximization of revenue in terms of patient volume weighted by case severity not only relates to financial aspects but can also be in line with the goals of patients seeking treatment in hospital systems with scarce resources and long waiting times. Waiting times and referrals of patients to other hospitals due to capacity shortages represent additional performance indicators that address finances but also aspects concerning the convenience of patients and quality of care.

3.1.4. Patient demand

It is important to assess patient demand thoroughly as the optimal case mix is sensitive to patient demand. Long-term decisions concerning the handling of non-elective patient demand and decisions concerning the volume of elective patients are different in nature. Non-elective patient demand is often dealt with by reserving adequate buffer capacity to provide timely access to hospital resources and to avoid extensive overtime. Consequently, non-elective patient demand influences the capacity of resources available for elective patients rather than being an integral part of case mix choices.

The estimation of elective patient demand can be based on historic volumes or forecasts taking demographic, sociological, and further trends into account. Statistics on patients referred to other hospitals or patients opting for treatment in other hospitals due to long waiting list are seldom covered by hospital information systems. Thus, estimating the patient demand based on historic volumes is potentially biased. This is primarily relevant for hospitals with catchment areas overlapping with those of competing hospitals.

3.1.5. Resource requirements and resource availability: modeling constraining resources

Only resource restrictions concerning resources representing bottlenecks need to be considered in case mix planning. If a resource is not scarce, or can be extended at relatively low costs, its consideration as a constraining resource only adds complexity to the problem without adding value. Constraints restricting such resources will not be tight in the optimal solution of the mathematical program and can thus be omitted. Thus, it can be advantageous to deal with sufficiently available resources only in subsequent planning problems. In general, operating room capacity is considered as one of the most scarce and expensive resources in hospitals. Consequently, the allocation of operating room time is frequently considered as an approach to describe the case mix in the literature on case mix planning [5]. In the remainder of this paper, we follow this operating room focused perspective. It has to be noted that the resulting planning problems can easily be extended to include non-surgical specialties.

According to Vissers et al. [27], bottleneck resources can be allocated to organizational units based on different policies, (i) dedication: resources are dedicated to one organizational unit for the duration of the planning horizon, (ii) time-sharing: resources are allocated to different organizational units on different periods, and (iii) sharing: resources are available for all organizational units on request. Table 2 provides examples for these three types of allocation policies.

Hybrid forms extend these definitions. For example, Dexter et al. [17] and Gupta [18] propose to allocate operating room time for the whole planning horizon but possibly reallocate it to other departments or surgical groups on short notice if there is not enough demand for the department or surgical group the operating room time was previously assigned to. Cross-trained nurses or overflow beds can also be interpreted as shared extensions of dedicated resources. Ma and Demeulemeester [25] and Harper [32] describe the practice of assigning patients to another suitable ward when there is no available capacity.

The choice of whether and, if so, to what extent bottleneck resources should be allocated in advance is an important management decision. This is especially true if there are major fluctuations in resource demand.

3.1.6. Generic formulation of case mix planning models

In general, mixed integer or stochastic mixed integer programs are used to model case mix planning problems. The following conventions for notation are used in the remainder of the paper. Parameters are described with capital letters and decision variables with small letters. An indexed letter in surrounding parentheses represents a matrix with the corresponding dimensions, e.g., the notation (a_{rj}) with surrounding parentheses is used to describe the matrix of allocation decisions of resources indexed with r to organizational units indexed with j. The following indices, sets, parameters, and decision variables are used for a generic description of case mix planning problems. A summary of the complete notation used in the remainder of this paper is provided in the appendix.

Indices and sets	
$r \in \{1, \dots, m\}$	Resources
$j \in \{1, \dots, n\}$	Organizational units
Parameters	
A_r	Available amount of resource r
Decision variables	
a_{rj}	Amount of resource r allocated to organizational unit j

When planning the case mix, the decision maker in charge is faced with the question of how to determine the best allocation (a_{rj}) of resources $r \in \{1, \dots, m\}$ to organizational units $j \in \{1, \dots, n\}$

in the hospital. Note that organizational units are considered as departments in the remainder of the paper. However, the notion of departments can easily be exchanged with subspecialties, wards, physician groups, or individual physicians depending on the hospital setting. The problem can be formulated in a generic form as follows.

$$\max \left\{ V((a_{rj})) \,\middle|\, \sum_{j=1}^{n} a_{rj} \leq A_r \quad \forall\, r \in \{1,\ldots,m\},\, (a_{rj}) \in X \right\} \tag{1}$$

The function $V((a_{rj}))$ describes the value of the allocation that the decision maker expects, A_r is the total amount of available resources of resource r, and X includes further restrictions on the resource allocation. The value $V((a_{rj}))$ depends on the number of patients that are to be admitted in each department in the course of the planning horizon. Target volumes can be derived as a main output of this rough cut allocation and used as input for further planning problems.

3.1.7. Model "N" - Naive approach

The most simple method to allocate resources for the next planning period is to maintain the current assignment of resources to departments, denoted as $A_{rj}^{current}$, i.e., $a_{rj} = A_{rj}^{current}$ for all $r \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$. While this approach sounds naive, it can be optimal under certain circumstances. Dexter et al. [17] describe a setting where allocated operating room time is forbidden to decrease for any subspecialty and there is not enough demand for any subspecialty to efficiently use an expansion of operating room resources. In their setting, the naive approach is argued to be optimal and expanded operating room capacity is simply used as overflow capacity.

3.1.8. Model "B" - Basic deterministic planning

The basic case mix planning model serving as a starting point for further reflections is a simple deterministic approach. The following additional notation is used.

Additional parameters	
M_j	Average contribution margin of department j
Q_{rj}	Required amount of resource r for treating a patient of department j
D_j	Demand for department j
Additional decision variables	
x_j	Number of patients admitted by department j

Allocating resources at the time of planning the case mix or assigning resources when required is identical from a mathematical point of view in basic deterministic models. Thus, a distinction between shared and dedicated resources is, in general, not relevant when discussing such models.

In the basic model, resources are allocated to maximize the sum of patients weighted by the respective contribution margin M_j or similar subject to resource and demand constraints. For each department j and resource r, the patient volume $x_j \in \mathbb{Z}_0^+$ of department j weighted with the average resource requirement Q_{rj} is not allowed to exceed the allocated capacity a_{rj}. This can be modeled using the following constraints.

$$Q_{rj} \cdot x_j \leq a_{rj} \quad \forall\, r \in \{1,\ldots,m\},\, j \in \{1,\ldots,n\} \tag{2}$$

These constraints are equivalent to the condition that the patient volume of any department j cannot exceed $\lfloor \frac{a_{rj}}{Q_{rj}} \rfloor$ for any resource r.

In addition, the patient volumes cannot be larger than the demand. The resulting problem can be formulated in a compact way plugging in the following value function into (1) and setting $X = \mathbb{R}^{m,n}$.

$$V((a_{rj})) = \sum_{j=1}^{n} M_j \cdot \min\left\{ \left\lfloor \frac{a_{1j}}{Q_{1j}} \right\rfloor, \ldots, \left\lfloor \frac{a_{mj}}{Q_{mj}} \right\rfloor, D_j \right\} \tag{3}$$

For example, Dexter et al. [29] describe such a basic deterministic approach. They consider operating rooms, intensive care unit (ICU) hours, ward hours, and implants as resources in their analysis. The only difference to a model described by (1) and (3) is that they use real-valued decision variables and hence the floor functions are omitted.

3.2. Specific modeling considerations

In this section, different model extensions are presented and discussed. The extensions comprise (i) stochastic modeling of demand (Extension 1), (ii) stochastic modeling of resource requirements (Extension 2), (iii) stochastic resource supply (Extension 3), and (iv) incorporation of time-phased allocation (Extension 4). The extensions are discussed one by one. Obviously, different extensions can be combined when formulating a case mix planning model.

Most of the model extensions described in the remainder of this section address decision making under uncertainty. If a parameter or a decision variable depends on the outcome $\omega \in \Omega$ of a random experiment, this is noted explicitly. For example, the notation $D(\omega)$ indicates that this parameter depends on the respective outcome ω of a random experiment, i.e., is a random variable. The notation $x(\omega)$ indicates that this decision variable represents a recourse decision. A recourse decision is a decision rule describing a decision in response to the respective outcome ω of the random experiment. In the remainder, $\omega \in \Omega$ is used to describe the "state of the world", i.e., to simultaneously describe all stochastic influences relevant for the problem under consideration to simplify notation.

3.2.1. Extension 1 - Stochastic demand

Demand for hospital services is naturally highly variable. Consequently, stochastic demand is considered in different papers on case mix planning. The following function describes the value of an allocation scheme (a_{rj}) for the basic model extended by the consideration of fluctuations in demand.

$$V((a_{rj})) = \sum_{j=1}^{n} E\left[M_j \cdot \min\left\{ \left\lfloor \frac{a_{1j}}{Q_{1j}} \right\rfloor, \ldots, \left\lfloor \frac{a_{mj}}{Q_{mj}} \right\rfloor, D_j(\omega) \right\} \right] \tag{4}$$

Case mix planning problems focusing on stochastic demand are described by Dexter et al. [17] and Gupta [18]. They solely consider the allocation of operating room capacity, i.e., $m = 1$, and use a modification of the function (4) assigning a salvage value M to unused capacity. The resulting problem is mathematically equivalent to plugging in the following function into (1) with $X = \left\{ a_{OR,j} \in \mathbb{R}^{m,n} \,|\, a_{OR,j} \geq A_{OR,j}^{current} \quad \forall\, j \in \{1,\ldots,n\} \right\}$.

$$\begin{aligned} V((a_{OR,j})) = \sum_{j=1}^{n} E\Bigg[& M_j \cdot \min\left\{ \frac{a_{OR,j}}{Q_{OR,j}}, D_j(\omega) \right\} \\ & + M \cdot \max\left\{ a_j - Q_{OR,j} \cdot D_j(\omega), 0 \right\} \Bigg] \end{aligned} \tag{5}$$

3.2.2. Extension 2 - Stochastic resource requirements

Uncertainties in the use of resources in aggregated planning can be modeled as the sum of independent and identically distributed random variables $Q_{rj}(\omega)$ describing the random use of resource r of one patient of department j. Conceivable sources for uncertainty are, for example, the surgery duration and the length of stay in the ward or in the ICU. Resource constraints can concern the number x_j of patients planned to be treated before knowing the realized values of $Q_{rj}(\omega)$ or the number $x_j(\omega)$ of patients that can be treated given the realized values of $Q_{rj}(\omega)$.

Let $Q_{rj}^{l}(\omega)$ denote the lth replication of the random variable $Q_{rj}(\omega)$. Resource requirements of resource r by patient volumes

Table 3
Linearization of stochastic resource requirements.

Modeling approach	Variance of resource requirements	Number of decision variables
Exact non-linear	$x_j \cdot Var[Q_{rj}(\omega)]$	n
Exact linear via multiple choice	$x_j \cdot Var[Q_{rj}(\omega)]$	$n \cdot (2 + UB_j)$
Multiplicative approximation heuristic	$x_j^2 \cdot Var\big[Q_{rj}(\omega)\big]$	n
Newsvendor based two stage heuristic	$k_{rj} \cdot Var[Q_{rj}(\omega)]$ (first stage)	n

x_j and $x_j(\omega)$ in department j can be expressed as $\sum_{l=1}^{x_j} Q_{rj}^l(\omega)$ and, respectively, $\sum_{l=1}^{x_j(\omega)} Q_{rj}^l(\omega)$. The former term is used by Choi and Wilhelm [19] arguing that a fixed number of surgeries to be scheduled for each department per day is to be determined in advance. For resource requirements concerning planned patient volumes, it is advisable to relax resource constraints of the form $\sum_{l=1}^{x_j} Q_{rj}^l(\omega) \leq a_{rj}$ since a value of x_j feasible for all possible outcomes of $Q_{rj}(\omega)$ is very likely too restrictive. Possible relaxations include the use of soft constraints where a violation of the constraint is penalized in the objective function or the use of chance constraints. Soft constraints are used by Choi and Wilhelm [19] to describe a setting where overutilization is allowed but comes at certain costs. A situation where capacity of downstream resources is not to be exceeded with a certain probability when planning the case mix is modeled by Fügener [21] using chance constraints. Yahia et al. [20] use a model formulation where the number of patients depends on the realization of ω. Hence, the corresponding decisions are modeled as recourse decision variables $x_j(\omega)$.

Both terms, $\sum_{l=1}^{x_j} Q_{rj}^l(\omega)$ and $\sum_{l=1}^{x_j(\omega)} Q_{rj}^l(\omega)$, are in general difficult to incorporate in mathematical optimization models due to their non-linearity. Table 3 summarizes different modeling approaches for stochastic resource requirements. Theoretically, the addition of independent random variables corresponds to convolutions of probability distributions. Due to the complexity and diversity of the involved probability distributions, it is in general challenging to solve the corresponding non-linear program. For example, Fügener [21] addresses the problem of sums of random variables where the number of terms is a decision variable using convolutions of distribution functions. The resulting distribution is approximated by a normal distribution. Embedding the resulting chance constraints leads to a non-linear program which is linearized using piecewise linear functions.

A computationally expensive alternative for identifying the optimal solution is the use of multiple choice constraints where integer variables concerning patient volumes $x_j(\omega)$ are transformed to binary decision variables $y_{jk}(\omega)$ via $k \cdot y_{jk}(\omega) = x_j(\omega)$ and $\sum_{k=0}^{UB_j} y_{jk}(\omega) = 1$. The linearized term for resource requirements reads as follows.

$$\sum_{k=0}^{UB_j} \left(\sum_{l=1}^{k} Q_{rj}^l(\omega) \right) \cdot y_{jk}(\omega), \tag{6}$$

This approach is only feasible if the upper bound on patient volumes UB_j are not too large.

Yahia et al. [20] suggest a model where the sum of random variables is replaced by a multiplication of the random variable with the decision variable, i.e., with $Q_{rj}(\omega) \cdot x_j$ and $Q_{rj}(\omega) \cdot x_j(\omega)$. A similar approach is pursued by Freeman et al. [23] assuming the same surgery durations for all patients of a patient group when simulating required operating room resources. The expected values for the original term and the multiplicative approximation are the same. However, the variance of the sum is $Var[\sum_{l=1}^{x_j(\omega)} Q_{rj}^l(\omega)] = x_j(\omega) \cdot Var[Q_{rj}(\omega)]$, whereas the variance of the product is $Var[Q_{rj}(\omega) \cdot x_j(\omega)] = x_j(\omega)^2 \cdot Var[Q_{rj}(\omega)]$. The variance is therefore overestimated by a factor of $x_j(\omega)$ and consequently the standard deviation is overestimated by a factor of $\sqrt{x_j(\omega)}$. Gerchak et al. [33] note that the multiplicative approximation is often applied in production models with random yields. Choi and Wilhelm [19] reveal in a case study that using the multiplicative approach can lead to considerably different objective values when compared with other solution approaches.

Another approach to modeling stochastic resource requirements of resource r is presented by Choi and Wilhelm [19] making use of newsvendor based heuristic solutions procedures with underage costs C_{rj}^u and overage costs C_{rj}^o. The idea is to identify an optimal representative number of patient slots k_{rj} to be scheduled by solving the following optimization problems for all departments $j \in \{1, \ldots, n\}$.

$$\min_{k_{rj} \in \mathbb{Z}_0^+} C_{rj}^u E\left[\max\left(A_r - \sum_{l=1}^{k_{rj}} Q_{rj}(\omega), 0\right)\right] + C_{rj}^o E\left[\max\left(\sum_{l=1}^{k_{rj}} Q_{rj}(\omega) - A_r, 0\right)\right] \tag{7}$$

The value of k_{rj} can then be used as a constant in the case mix planning model.

3.2.3. Extension 3 - Stochastic resource capacity

Uncertainties in resource capacity can originate from different sources. Assuming that all medical equipment is fully functional for most of the time, the major factor of uncertainties in resource capacity concerns human labor. Uncertainties in the capacity of resources such as operating rooms, ICU or ward capacity most often originate from variabilities in staffing levels. Potential sources include stochastic influences such as illness of staff and delays of the first surgery in an operating room. Stochastic resource capacity can also be used as a modeling device to account for disruptions of the schedule of elective patients by emergency patients. In this case, the available capacity for elective patients can be modeled as the total capacity minus capacity required by emergency patients depending on the outcome $\omega \in \Omega$.

It is reasonable to assume that the allocation of resources takes place before uncertainties concerning the resource capacity become known since otherwise there is no stochastic influence any more at the time of resource allocation. If a resource with uncertain capacity corresponds to a resource to be allocated, the amount of resources available for a department can be considered as a random variable depending on the value of the decision variable a_{rj} describing the corresponding allocation. In this case, Constraints (2) emerge as $Q_{rj} \cdot x_j \leq f(a_{rj}, \omega)$, where $f(\cdot, \cdot)$ is a function translating the allocated resource into actually available resources. This function can be of a similar form as discussed for the case of stochastic influences on resource requirements. If there is no need to allocate a resource with uncertain capacity, this directly relates to a right hand side value $A(\omega)$ which can be interpreted as a random variable since no decision variables are involved. The resulting constraints read as $\sum_{j=1}^{n} Q_{rj} \cdot x_j \leq A_r(\omega)$. Similar to the case of stochastic resource requirements, this constraint should either be relaxed or the decision variable concerning the patients should be modeled as a recourse decision variable to avoid excessive underutilization since otherwise this constraint has to hold for

$\min_{\omega\in\Omega}\{A_r(\omega)\}$, no matter how small the probability for the corresponding event is. Yahia et al. [20] consider uncertainties in the capacity of nurses in their case mix planning approach. They use a constraint of the form $\sum_{j=1}^{n} a_{Nurse,j} \leq A_{Nurse}(\omega)$ which is from a mathematical perspective equivalent to the deterministic constraint $\sum_{j=1}^{n} a_{Nurse,j} \leq \min_{\omega\in\Omega}\{A_{Nurse}(\omega)\}$.

3.2.4. Extension 4 - Master surgery schedule (MSS)/admission planning

The MSS addresses the assignment of operating rooms to departments on the different days of the planning cycle. It can be considered as the next phase after case mix planning in the process of scheduling surgeries [24]. Similarly, admission planning, such as described in Adan and Vissers [34], relates to the determination of the number of patients from different patient categories to be admitted on different days of the planning cycle.

The main potential benefit of simultaneously determining the optimal case mix and the MSS or an admission plan concerns artificial variations in surgery scheduling due to inactive days, i.e., days without elective surgeries. For example, if most patients stay only one or two days in the ICU and there are no elective surgeries on the weekends, the ICU is likely not to run at full capacity on Mondays. Such an effect can be explicitly modeled taking the weekly structure of inactive days into account when planning the case mix. The main difference to the basic deterministic planning model *B* is the extension of the problem by the dimension of time. The superscript *day* is used to distinguish the time-phased parameters and variables from aggregated parameters and variables used in the other models. The corresponding extended notation reads as follows.

Additional indices and sets	
$t \in \{1,\dots,T\}$	Days of the planning cycle
$i \in \{1,\dots,I\}$	Operating rooms
$S_j \subseteq \{1,\dots,I\}$	Operating rooms equipped for surgeries of department j
Time-phased parameters	
A^{day}_{rt}	Available amount of resource r on day t
$Q^{day}_{rj\tau}$	Required amount of resource r for treating a patient of department j on day τ after admission
Time-phased decision variables	
a^{day}_{rjt}	Amount of resource r allocated to department j on day t
y^{day}_{jti}	1 if operating room i is assigned to department j on day t, 0 otherwise
x^{day}_{jt}	Number of patients admitted by department j on day t

The resulting structure of the model is the following.

$$\max \quad \sum_{j=1}^{n} M_j \sum_{t=1}^{T} x^{day}_{jt} \tag{8}$$

subject to

$$\sum_{\tau=0}^{\infty} Q^{day}_{rj\tau} x^{day}_{j((t-\tau) \bmod T)} \leq a^{day}_{rjt} \quad \forall\, r\in\{1,\dots,m\}, j\in\{1,\dots,n\}, t\in\{1,\dots,T\} \tag{9}$$

$$\sum_{t=1}^{T} x^{day}_{jt} \leq D_j \quad \forall\, j\in\{1,\dots,n\} \tag{10}$$

$$\sum_{j=1}^{n} a^{day}_{rjt} \leq A^{day}_{rt} \quad \forall\, r\in\{1,\dots,m\}, t\in\{1,\dots,T\} \tag{11}$$

$$\sum_{j=1}^{n} y^{day}_{jti} \leq 1 \quad \forall\, t\in\{1,\dots,T\}, i\in\{1,\dots,I\} \tag{12}$$

$$\sum_{i\in S_j} y^{day}_{jti} = a^{day}_{OR,j,t} \quad \forall\, j\in\{1,\dots,n\}, t\in\{1,\dots,T\} \tag{13}$$

$$x^{day}_{jt} \in \mathbb{Z}^+_0 \quad \forall\, j\in\{1,\dots,n\}, t\in\{1,\dots,T\} \tag{14}$$

$$y^{day}_{jti} \in \{0,1\} \quad \forall\, j\in\{1,\dots,n\}, t\in\{1,\dots,T\}, i\in\{1,\dots,I\} \tag{15}$$

$$(a^{day}_{rjt}) \in X \tag{16}$$

Constraints (9) to (11) relate to the constraints of the basic version of the case mix planning problem extended by the dimension of time with index t. Constraints (9) correspond to Constraints (2) of the basic problem and can be read as follows. The parameters $Q^{day}_{rj\tau}$ define the requirements of resource r for a patient of department j on day τ after admission. Consequently, the product of $Q^{day}_{rj\tau} \cdot x^{day}_{j((t-\tau) \bmod T)}$ defines the amount of resource r required by patients admitted on day τ at day $t-\tau$. The latter is to be read modulo T due to the cyclic approach. The sum over all requirements of resource r for all possible days of admission τ days before t results in the required amount of resource r for department j on day t. Using empirical data as input, the sum only contains a finite number of terms. In general, the number of terms can be reduced to T in a preprocessing step by adding the respective values $Q^{day}_{rj\tau}$ with equal values of τ modulo T. Constraints (12) and (13) determine the MSS. In particular, Constraints (12) guarantee that each operating room is assigned to at most one department on each day of the planning cycle. Constraints (13) implicitly ensure that an operating room is only assigned to a department if it is equipped according to the specific requirements of the department. In addition, this set of constraints links the binary decision variables assigning individual operating rooms to departments to the corresponding aggregated decision variables. The data and the decision variables can be linked to the notation of Model "*B*" as follows.

$$A_r = \sum_{t=1}^{T} A^{day}_{rt}, \quad Q_{rj} = \sum_{\tau=0}^{\infty} Q^{day}_{rj\tau}, \quad a_{rj} = \sum_{t=1}^{T} a^{day}_{rjt}, \quad x_{jt} = \sum_{t=1}^{T} x^{day}_{jt} \tag{17}$$

The notion of $Q^{day}_{rj\tau}$ is illustrated using the example of average bed requirements in the ICU ($r = ICU$) for patients of department j on the days following the admission of the patient. Consider the following average bed requirements on the days $\tau = 0, 1, 2, \dots, 9$ after admission.

$$Q^{day}_{ICU,j,\cdot} = (0.61, 0.37, 0.22, 0.13, 0.08, 0.05, 0.03, 0.02, 0.01, 0.01)$$

The coefficients for the optimization model using a cycle length of $T = 7$ are depicted in Table 4. The table is to be read as follows. Consider patients admitted at day $t = 1$ of the planning cycle. The corresponding requirements are listed in the first column. Patients admitted at day $t = 1$ need on average $Q^{day}_{ICU,j,0} + Q^{day}_{ICU,j,7} = 0.61 + 0.02 = 0.63$ beds on the first day of the planning cycle since $0 \equiv 7 \bmod 7$ and so forth. In total, they require a bed in the ICU for 1.53 days on average. The average resource requirements are the same for patients admitted on days $t = 2, 3, 4, 5$ shifted by 1,2,3,4 days, respectively. It is assumed that there are no admissions of elective patients on the weekend. The resource requirements correspond to a total average length of stay of $Q_{ICU,j} = \sum_{\tau=0}^{\infty} Q_{ICU,j,\tau} = 1.53$ in the basic aggregated formulation of Model "*B*". We refer to Adan and Vissers [34] for a more detailed description of these modeling conventions.

In the case of identically equipped operating rooms, i.e., $S_1 = \dots = S_n = \{1,\dots,I\}$, the determination of an MSS timetable can be skipped in the master surgery schedule extension (Extension 4) and postponed to a post-processing step. This relates to discarding Constraints (12) and (13) in the master surgery schedule extension and solving the following feasibility problem where the parameter $A^{day}_{OR,j,t}$ is equal to the corresponding value $a^{day}_{OR,j,t}$ of the case mix

Table 4
Coefficients of resource requirements.

Avg. resource requirement on day t	Admission on day t of the planning cycle						
	1	2	3	4	5	6	7
1	0.63	0.03	0.05	0.08	0.13	–	–
2	0.38	0.63	0.03	0.05	0.08	–	–
3	0.23	0.38	0.63	0.03	0.05	–	–
4	0.13	0.23	0.38	0.63	0.03	–	–
5	0.08	0.13	0.23	0.38	0.63	–	–
6	0.05	0.08	0.13	0.23	0.38	–	–
7	0.03	0.05	0.08	0.13	0.23	–	–
Sum	1.53	1.53	1.53	1.53	1.53	–	–

planning problem for all $j \in \{1, \ldots, n\}$ and $t \in \{1, \ldots, T\}$.

$$\Big\{ (y_{jti}^{day}) \in \{0,1\}^{n \times T \times I} \; \Big| \; \sum_{j=1}^{n} y_{jti}^{day} \leq 1 \quad \forall\, t \in \{1, \ldots, T\}, i \in \{1, \ldots, I\}; \\ \sum_{i \in S_j} y_{jti}^{day} = A_{OR,j,t}^{day} \quad \forall\, j \in \{1, \ldots, n\}, t \in \{1, \ldots, T\} \Big\} \tag{18}$$

Incorporating the structure of a week is potentially relevant for case mix planning if there are inactive days without surgeries and at the same time the set of scarce resources includes resources required on consecutive days such as beds in the ward or in the ICU.

3.3. Simulation of operational performance

A high level of abstraction using simplifying assumptions is regularly used when planning the case mix due to the complexity of the involved processes. Thus, results of case mix planning models should be validated before applying them in practice. It is tempting to use the same set of assumptions for the evaluation of the models as was used for modeling the case mix planning problem. However, this approach can lead to biased results. When evaluating the case mix it is often not necessary to make assumptions that are as strict as those imposed in case mix planning models. Evaluations via simulation, for instance, can often be modeled on a much more detailed level than optimization problems. Thus, constraints neglected at the level of case mix planning can be incorporated in simulations. For example, Freeman et al. [23] use randomly generated surgery durations and lengths of stay for individual patients when simulating the relevant processes for evaluating the case mix plan. In the case mix planning model, a simplifying assumption of deterministic length of stay is made. Surgery durations are also drawn from a probability distribution but are assumed to be similar for all patients of a patient group when planning the case mix.

A simulation model for evaluating the operational performance of a case mix planning model should cover all relevant bottleneck resources of the hospital. It should be tailored to the processes of the individual hospital. Thus, only few generic points relevant for any hospital setting are addressed in this section. A more detailed description of a simulation applied to a real world problem setting is provided in Section 4.3.

Summarizing, assumptions made in the case mix planning phase to simplify the case mix planning model can often be relaxed when evaluating case mix strategies via simulation. Further, the processes in the hospital can be modeled in much more detail than suitable for aggregated planning purposes. Consequently, it is possible that resource allocation schemes assumed to be optimal can be outperformed by allocation schemes considered to represent suboptimal solutions.

4. Experimentation

The framework is applied to the setting and data of a large German hospital. The value of incorporating the different extensions discussed in Section 3.2 is discussed using a factorial experimental study. In particular, the decisions derived by different case mix planning models are evaluated using a simulation of the relevant hospital processes to identify which extension adds value.

4.1. Problem setting

In this section, we describe the problem setting of a large German hospital planning the case mix. This description is based on detailed expert interviews with the physicians in charge of managing the operating rooms and the ICU at this hospital.

In hospital systems with case payment mechanisms, hospitals have an incentive to influence the case mix. Operating rooms, including the surgical staff, are often considered as the most expensive resource in the hospital. Thus, planning the case mix and allocating operating room resources to organizational units such as departments within the hospital are strongly interrelated problems. The treatment of different types of patients requires different qualifications for the staff. It is a common approach to allocate operating room time and adjust the staff accordingly for medium-term to long-term planning horizons such as one year or six months. We assume that an operating room is allocated to at most one department per day. All operating rooms can be equipped with instruments according to the specific requirements of the departments in the setting of the case study hospital. Thus, all rooms are applicable to serve any department. Further, we consider the same allocation for every week of the planning horizon on the assumption that seasonal influences on the demand and the supply of hospital services can be neglected. Please note that the model can easily be adjusted to take seasonal fluctuations into account if required. The goal is to determine the optimal number of operating room days to be allocated to a department per week.

For each department, a fixed number of elective surgeries to be scheduled per day is determined. The goal is to find the optimal balance between high capacity utilization and limited overtime. Distributions of surgery durations as well as disruptions through emergencies are taken into account. If emergencies are served in a dedicated operating room, they only have a minor influence on the schedule of the elective patients [35]. Otherwise, the elective schedule is interrupted by emergency arrivals, since emergencies typically have to undergo surgery within a short time span. It is assumed that costs of an increase in the length of stay by one day exceed the costs of overtime. These costs not only relate to monetary aspects but also to the convenience of patients and to consecutive disruptions of the surgery schedule the next day. Thus, elective patients are usually treated the day they are scheduled. Scheduled surgeries that end after or are not started before the end of the regular opening hours induce overtime.

Situations can occur where there is a shortage of beds in the ward a patient scheduled for surgery is allocated to. It is assumed that patients can be assigned to appropriate alternative wards if necessary and then be relocated to the ward the patient was previously assigned to once capacity is available. Such alternative assignments can be reduced using refined master surgery schedules as proposed by Ma and Demeulemeester [25]. All in all, bed and nursing capacity in the regular wards are not considered as major bottlenecks resources in our study. It should be noted that this assumption applies to the case site but does not hold in all circumstances. Thus, it is important to highlight that the results of the case study are conditioned by this assumption.

The only situation where scheduled surgeries are canceled regularly is when the patient is likely to require monitoring in the ICU after surgery and there is no capacity left in the ICU. If necessary, surgeries can be preponed or postponed on short notice on the same day they are scheduled. Note that the capacity of staff in the operating rooms and in the ICUs is considered implicitly. The capacity of surgical teams influences the number of surgeries to be regularly scheduled and the availability of ICU staff determines the number of staffed beds available for monitoring.

4.2. Implementation of the case mix planning model

In this section, an integer program based on the formulation of the case mix problem (Model B) making use of all extensions (Extension 1) to (Extension 4) is presented. This integer program is considered as the "full model" since it incorporates all types of extensions discussed in Section 3.2. The major decision to be derived by the model is the allocation of operating room time to departments. The decision on the number of operating room blocks to be assigned to each specialty j on each day t of the planning horizon is denoted as $a_{OR,j,t}$. The recourse variable $x_{jt}^{day}(\omega)$ represents an auxiliary decision variable determining the number of elective patients of department j that can be admitted to surgery on day t given the realization of the "state of the world" $\omega \in \Omega$.

$$\max \quad \sum_{j=1}^{n} M_j E\left[\sum_{t=1}^{T} x_{jt}^{day}(\omega)\right] \tag{19}$$

subject to

$$x_{jt}^{day}(\omega) \le K_j a_{OR,j,t}^{day} \quad \forall\ j \in \{1,\dots,n\}, t \in \{1,\dots,T\} \tag{20}$$

$$\sum_{j=1}^{n}\sum_{\tau=0}^{\infty} Q_{ICU,j,\tau}^{day}(\omega) x_{j((t-\tau)\bmod T)}^{day}(\omega) \le A_{ICU,t}^{day}(\omega) \quad \forall\ t \in \{1,\dots,T\} \tag{21}$$

$$\sum_{t=1}^{T} x_{jt}^{day}(\omega) \le D_j(\omega) \quad \forall\ j \in \{1,\dots,n\} \tag{22}$$

$$\sum_{j=1}^{n} a_{OR,j,t}^{day} \le A_{OR,t}^{day} \quad \forall\ t \in \{1,\dots,T\} \tag{23}$$

$$a_{OR,j,t}^{day} \le A_{OR,j}^{\max} \quad \forall\ j \in \{1,\dots,n\}, t \in \{1,\dots,T\} \tag{24}$$

$$a_{OR,j,t}^{day} \in \mathbb{Z}_0^+ \quad \forall\ j \in \{1,\dots,n\}, t \in \{1,\dots,T\} \tag{25}$$

$$x_{jt}^{day}(\omega) \in \mathbb{Z}_0^+ \quad \forall\ j \in \{1,\dots,n\}, t \in \{1,\dots,T\} \tag{26}$$

The Objective (19) is the maximization of the expected revenue. It is assumed to depend linearly on the number of patients treated in each department. Constraints (20) to (22) in combination with the Objective (19) define the value $V((a_{rjt}^{day}))$ of the resource allocation scheme (a_{rjt}^{day}).

Uncertain surgery durations $Q_{OR,j}(\omega)$ are considered using a two stage heuristic as described in Section 3.2. In the first stage, the following newsvendor based optimization model is solved for every department $j \in \{1,\dots,n\}$ weighing underage costs $C_{OR,j}^{u}$ and overage costs $C_{OR,j}^{o}$ for idle time and overtime, respectively, to identify the optimal representative number of surgeries $k_{OR,j}$ related to the opening hours A_{OR} of the operating rooms.

$$\begin{aligned}\min_{k_{OR,j}\in\mathbb{Z}_0^+} \ & C_{OR,j}^{u} E\left[\max\left(A_{OR} - \sum_{l=1}^{k_{OR,j}} Q_{OR,j}(\omega), 0\right)\right] \\ & + C_{OR,j}^{o} E\left[\max\left(\sum_{l=1}^{k_{OR,j}} Q_{OR,j}(\omega) - A_{OR}, 0\right)\right]\end{aligned} \tag{27}$$

In the second stage, the optimal values are then used as constants $K_{OR,j} = k_{OR,j}$ in Constraints (20) in the case mix planning model. The number of surgeries to be scheduled per department and day is restricted by the product of $K_{OR,j}$ and $a_{OR,j,t}^{day}$. This approach to addressing uncertain surgery durations is a heuristic approach as combining the two stages in one model can lead to a better solution.

Constraints (21) limit the number of patients by the available capacity of the ICU. It is assumed that staff and equipment in the ICU is suitable to monitor patients from all departments. Thus, there is no need to allocate ICU beds to different departments in advance. The number of patients that require monitoring in the ICU on day t of the planning cycle is determined by the number of patients that are admitted to surgery on the same day or on previous days but still require monitoring on day t. Let $Q_{ICU,j,\tau}^{day}(\omega)$ be the relative share of patients of department j that still requires monitoring τ days after surgery. Multiplying the share with the number of surgeries τ days before t, i.e., at time $t-\tau$, results in the number of patients that underwent surgery on day $t-\tau$ and still require treatment on day t. Note that these computations are to be read modulo T. Summing up these products over all days $\tau = 0, 1, \dots$ before surgery results in the required number of beds for each day t of the planning cycle. Similar to Constraints (20), the multiplicative approach used in Constraints (21) is a heuristic approach to include stochastic resource requirements as illustrated in Section 3.2.

Constraints (22) restrict the total number of patients admitted by each department during the cycle time. The distribution $D_j(\omega)$ represents the aggregated demand adjusted for appropriate waiting times. Constraints (23) assure that the maximum number of operating rooms available on each day is not exceeded. In addition, Constraints (24) set bounds on the maximum number of operating room slots per day for each department. This is relevant due to a restricted number of operating teams in each department that can work at the same time. These constraints define the set X of feasible allocations such as denoted in (1). There are no restrictions limiting the assignment of operating rooms to departments. Consequently, the detailed MSS timetable assigning individual operating rooms to departments on different days is derived in a post-processing step such as described in Section 3.2 without disadvantage. The domains of the variables are given in (25) and (26). All decision variables are assumed to be integer. The stochastic programs are transformed into deterministic approximations using SAA. All integer programs are modeled in CPLEX 12.6.3.

4.3. Implementation of the simulation

The relevant processes in the hospital providing the basis for the simulation are illustrated in Fig. 3 using the process modeling language Business Process Model and Notation (BPMN) 2.0. The starting point of the process is an incoming request for surgery

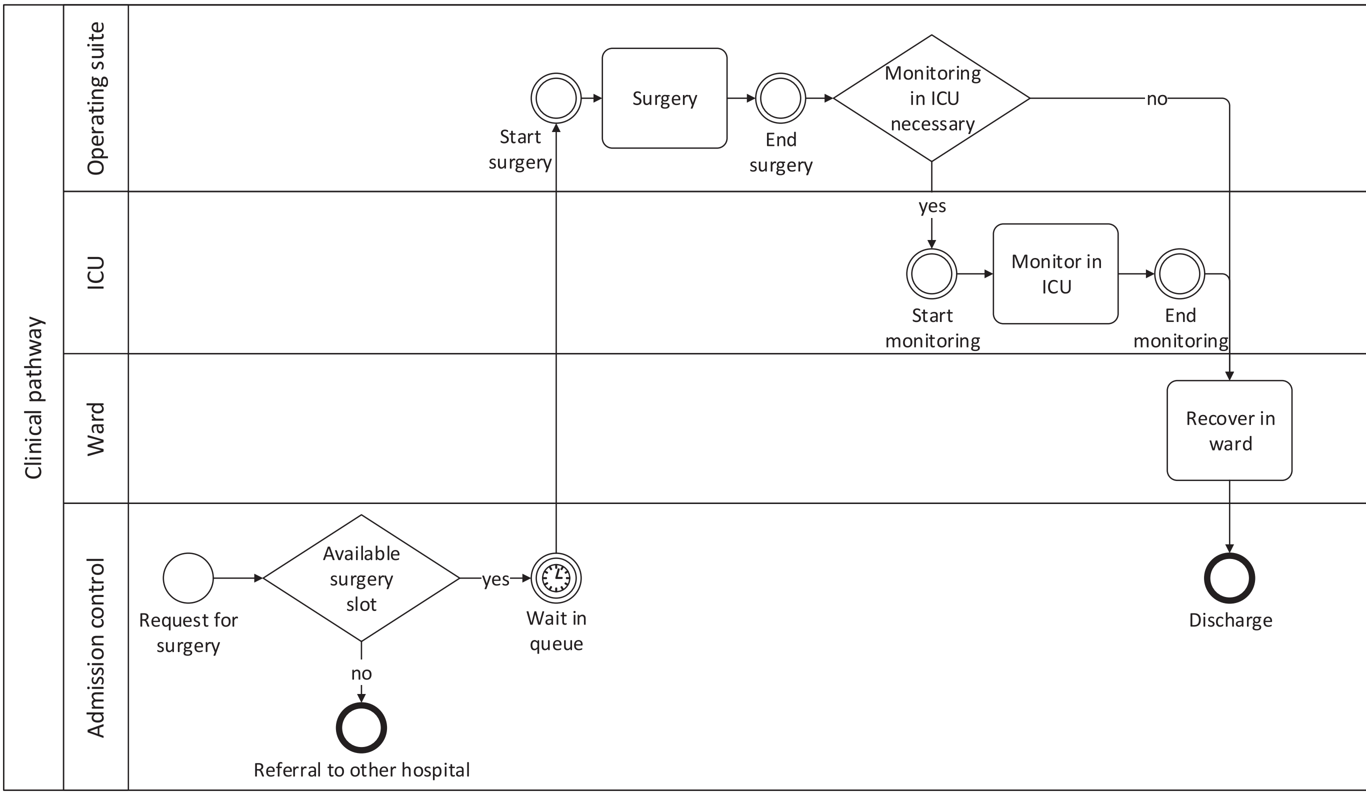

Fig. 3. Model of the process used for the simulation of operational performance of the CMP models.

registered by the hospital admission. Requests for elective patients are added to the queue of patients to be scheduled for the corresponding department. As soon as it becomes clear that there is no available capacity to book the patient for surgery within the maximum waiting time of the patient, the patient is referred to another hospital. For each day, a fixed number of elective patients is scheduled for each department. These patients are served the day they are scheduled, irrespective of the actual duration of the surgeries and independent on the number of incoming emergencies on that day. Overtime is used to cover any potential capacity bottleneck. The only exception where an elective patient does not undergo surgery the day he or she is scheduled is when the patient is likely to require monitoring on the ICU but there is not enough capacity to accommodate the patient. Once surgery is finished, a check is made to determine if monitoring on the ICU is required. This decision can differ from the assessment before the surgery. If the patient is admitted to the ICU, he or she is monitored until his condition is assessed as stable. Afterwards, he or she is sent to the regular ward for further recovery. If no monitoring on the ICU is required, the patient is directly sent to the ward. After recovery in the ward, the patient is discharged.

The processes for emergency and urgent patients differ slightly. When an emergency patient arrives at the hospital, he or she is added to the top of the queue so that he or she can be served as soon as an operating room gets available. He or she is assigned to the next available operating room regardless of the department this operating room is assigned to. In addition, he or she does not wait until there is capacity in the ICU for monitoring after surgery, since it is assumed that there is always capacity on hold in the ICU to accommodate emergency patients. This is justified by reserving a certain share of beds as buffer capacity for emergencies. In the worst case in which buffer capacity is insufficient to accommodate an emergency patient, it is assumed that extra overflow capacity can be made available on short notice. In practice, this can relate to temporarily increasing patient to nurse ratios or to assigning the emergency patient to a different ward until capacity is available in the ICU. Alternatively, other ICU patients with stable conditions can be discharged early to release capacity [36]. Urgent patients have to be treated within one day but do not have to be treated immediately. Depending on their time of arrival, they are added to the daily elective program of a department the day they arrive or the following day and are otherwise treated similar to emergency patients in the simulation.

In practice, patient flows can differ. For example, it can be necessary to treat patients again if their condition worsens during their stay in the ICU or during their stay in the regular ward. Such patients are implicitly treated in the model as new requests for elective surgery or as emergency patients if immediate action is required. The ICU is a surgical intensive care unit, implying that only surgical patients are sent to this unit. We observed only very rare exceptions of this in the real-world data. In addition to the surgical ICU, there is a medical ICU available for non-surgical patients. The medical ICU is not part of the simulation since there is no major interference with the operating theater or the surgical ICU. The subprocess of recovery in the regular wards is not modeled explicitly either, since it is assumed that there is enough bed capacity to accommodate patients as described in Section 4.1. If there is a shortage of beds in the ward a patient would normally be assigned to, he or she is placed in an alternative ward and redirected to the appropriate ward as soon as capacity gets available.

The simulation is modeled in Java 8. The resource allocation scheme determined by the case mix planning model is used as input for the simulation. Elective and non-elective patient demand, surgery durations, and lengths of stay in the ICU are simulated based on parametric and empirical distributions as described below.

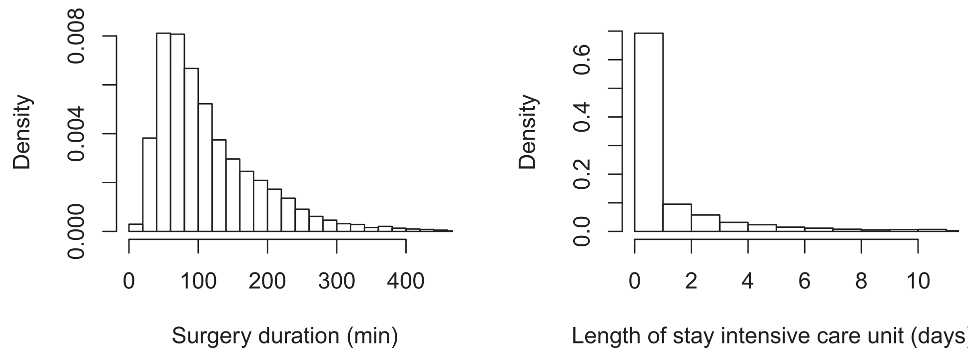

Fig. 4. Distributions of surgery duration and length of stay in the ICU aggregated over all departments.

4.4. Data

The hospital under consideration is a large German hospital with $n = 8$ departments with considerable patient volumes in the central operating theater. For empirical distributions of resource requirements of patients, two years of data were used, resulting in a total of 35,153 surgical patient records. This corresponds to records of about 17,500 surgical patients per year. The operating theater and the surgical ICU are the main sources of revenue for the hospital. At the same time, these resources are considered as the major bottleneck resources. Thus, these resources constitute the set of constraining resources when planning the case mix. Consequently, non-surgical patients are not part of the analysis. Target volumes for such patient groups can be derived in a subsequent planning phase.

There are 16 operating rooms available for surgeries for the eight departments. Each operating room is open for 8 h from Monday until Friday. Opening hours can be extended using overtime to cover fluctuations. No elective surgeries are scheduled on Saturdays and Sundays. During regular hours, emergency patients are treated in the next available operating room. Outside of the regular hours, such as during the night and on the weekend, a staffed operating room is on hold for emergency patients.

The surgical ICU has 40 regularly staffed beds. The care ratio recommended by the German Interdisciplinary Association for Intensive Care and Emergency Medicine (DIVI) is two patients per intensive care nurse [37]. Consequently, if one shift cannot be covered, two ICU beds have to be blocked and the currently available capacity is reduced by two beds. Table 5 gives an overview over the number of blocked beds. On average, 1.5 beds with a standard deviation of 1.8 were blocked. It is common practice to reserve 4 beds for newly arriving emergencies until the end of the regular opening hours. These slots are available for emergency patients arriving during the night and requiring monitoring. This data is used as input for the stochastic availability of ICU beds in the case mix planning problem.

The distributions of resource requirements used as input for the models are department-specific. In this paragraph, resource requirements aggregated over all departments are reported to provide a brief summary of the data. Elective patients account for 71%, urgent patients for 21%, and emergency patients for 7% of the patient records. In the case mix planning phase, urgent and emergency patients are treated similarly since they both cannot be postponed. The empirical distributions of the observed operating room times and the length of stay in the ICU are illustrated in Fig. 4. The distributions are split according to the urgency category and the department of the patients, when used as input for the case mix planning models and the simulation. About 24% of all patients have to be monitored in the ICU after surgery. Out of these patients, 67% are elective patients. Emergency patients are much more likely to require intensive care monitoring than urgent and elective patients. Almost 47% of all emergency patients are sent to the ICU after surgery as opposed to 23% and 20% ICU patients among elective and urgent patients. This results in an average of 9.9 elective patients sent to the ICU per day from Monday until Friday with a standard deviation of 3.2. There are on average 4.7 ICU patients per day from Monday until Sunday stemming from urgent or emergency surgeries with a standard deviation of 2.4. Most patients stay in the ICU for only a short period of time as can be seen in Fig. 4. Short-stay patients with a length of stay of up to 1 day account for 63% and patients with a length of stay of 2 days or less account for 77% of all transfers to the ICU. The average length of stay of a patient is 2.5 days with a standard deviation of 5.0 days. Consequently, utilization rates are potentially lower on Sundays and Mondays compared to other days when simulating the utilization of the ICU.

The estimation of patient demand is based on population census and disease rates. We assume that the probabilities of people living in the catchment area of the hospital for the demand of services for a department are independent and identically distributed for each day of the year. The resulting binomial distribution for the weekly demand of services of a department is approximated using a Poisson distribution. While the assumption of independent and identical distributions can be questioned, the Poisson distribution is often used in the literature to model patient demand [19,38,39]. The assumption was made since only limited information on the true demand for hospital services is available. The empirical observation of the observed patient volumes only accounts for the portion of the demand which is satisfied by the hospital. In addition, the observed volumes represent already smoothed values of the actual demand.

Table 5
Blocked ICU beds.

Blocked beds per day	Available beds	Relative frequency
0	40	49.0%
2	38	29.3%
4	36	17.2%
6	34	4.6%

4.5. Results

A factorial design was used to evaluate the impact of the model extensions (Extension 1 to 4) on the hospital performance for the given problem setting and data. The factorial design consists of 4 factors. Each factor describes whether this model extension is included in the case mix problem. This results in $2^4 = 16$ different models. In addition, these 16 models are compared with the status quo allocation (Model "*N*"). A separate MSS problem is solved afterwards in case Extension 4 is not part of the model. This MSS is constituted by solving a modification of Extension 4 where the output of the case mix planning problem regulates the allocation. It is used to assign the operating room blocks granted to each department via the case mix planning model to different days. The derivation of a master surgery time table assigning individual operating rooms to departments is made in a post-processing step as described in Section 3.2.

The case mix planning model described in Section 4.2 is considered as the "full model" since it contains all extensions. This model is simplified in the different experiments as follows. Extension 1, 2, and 3 are dropped by replacing $D_j(\omega)$, $Q^{day}_{ICU,j,\tau}(\omega)$, and $A^{day}_{ICU,t}(\omega)$ with the corresponding average values D_j, $Q^{day}_{ICU,j,\tau}$, and $A^{day}_{ICU,t}$, respectively. Extension 4 is dropped by aggregating Constraints (20), (21), (23), and (24), over the cycle length and replacing the decision variables with their aggregated counterparts, i.e., $\sum_{t=1}^{T} a^{day}_{OR,j,t} = a_{OR,j}$ and $\sum_{t=1}^{T} x^{day}_{jt} = x_j$. For example, Constraints (21) are replaced by the single following constraint.

$$\sum_{j=1}^{n}\left(\sum_{\tau=0}^{\infty} Q^{day}_{ICU,j,\tau}(\omega)\right)x_j(\omega) \le \sum_{t=1}^{T} A^{day}_{ICU,t}(\omega) \tag{28}$$

The computations were performed on Windows 10 with a 2.70 GHz CPU (Intel® Core™i7-3740QM) and 16 GB RAM. The stochastic integer programs are solved using SAA. The number of scenarios is chosen as $|\Omega| = 1000$ to obtain robust results. The resulting MIPs are solved using CPLEX 12.6.3 with a relative MIP gap of 1%. Each case mix model is evaluated by 1000 replications representing the simulation of 1000 weeks. The allocations of all models are evaluated using the same simulation model.

4.5.1. Computational performance

Fig. 5 illustrates the performance of the SAA for the full model with an increasing number of scenarios. With less than 125 scenarios, the approximation is not reliable. This result is in the same order of magnitude as the results described by Yahia et al. [20]. The SAA converges with an increasing number of scenarios. The solution times increase with an increasing number of scenarios. Due to the manageable number of $T \cdot n \cdot (|\Omega| + 1) = 7 \cdot 9 \cdot (|\Omega| + 1)$ decision variables, computational times are low. However, it has been shown that computational times are extensive if more detailed patient groups or organizational structures are considered [25].

4.5.2. Discussion of the results

The main results of the case mix planning problems and the simulations are displayed in Table 6. The reported figures solely relate to elective patients since it is assumed that the case mix strategy and the resource allocation scheme have no major influence on non-elective patients. Results are reported per week. On average, 95.9 non-elective patients arrive per week in addition to the elective patients reported in the table.

Each row of the table corresponds to one case mix planning model. The first column serves as an identifier for the 17 models that are to be evaluated. Columns 2 to 5 describe which complicating aspects, i.e., which extensions, are part of the respective case mix planning model. The next three columns report on KPIs and the computational performance of the case mix planning problems. The average revenue per week and the average number of patients per week serve as a benchmark to compare the different models. The highest revenue and the largest patient volumes are projected by the basic aggregated model "*B*". In general, the runtime is considerably higher when time-phased allocation decisions are part of the model (Extension 4). The benchmark values for the revenue and the patient volumes as simulated in the evaluation phase are reported in the fourth and the third last column. The simulated revenue is to be interpreted as the revenue obtained by the simulation model when using the output of the respective case mix planning model as input for the simulation. The simulated patient volumes are to be interpreted correspondingly. In addition to average values, the standard deviation of the revenue of the simulation runs is reported for each model.

There is a general tendency of increasing performance in terms of simulated revenue with increasing number of extensions that are considered. The last two columns address the overall performance of the case mix planning models. The prediction error describes the relative difference between the revenue predicted by the case mix planning problem and the simulated revenue. There are considerable differences for higher aggregated model formulations. Thus, when planning the case mix using basic deterministic formulations, one should be aware that highly aggregated planning approaches often constitute overestimations of achievable results. The potential for improvement is measured as the relative difference between the average simulated revenue of the respective case mix model and the average simulated revenue of the best performing model.

The output of the model is illustrated in the following example. Consider model 6. This model refers to a case mix planning model where the basic model is extended by a stochastic modeling of demand (Extension 1) and a stochastic modeling of resource availability (Extension 3). Recall that the cycle length is one week. The objective value of this problem is equal to 898,171. This figure corresponds to one optimization problem where the decisions are optimized with regard to 1000 randomly generated scenarios. The average number of patients of 240.5 is to be interpreted as the number of patients averaged over these scenarios. The model requires 1.9 s to be solved with a relative MIP gap of 1%. The output is translated to an MSS as described in the beginning of this section. This MSS serves as input for a simulation. The MSS is simulated for 1000 weeks. The simulated revenue is on average EUR 852,450 per week with a standard deviation of EUR 89,765, generated by on average 229.4 patients. The prediction error is calculated as $|898{,}171 - 852{,}450|/852{,}450 \approx 5\%$. The potential for improvement when compared to the best performing model in the simulation, i.e., model 15, is equal to $|852{,}450 - 865{,}338|/865{,}338 \approx 1.5\%$.

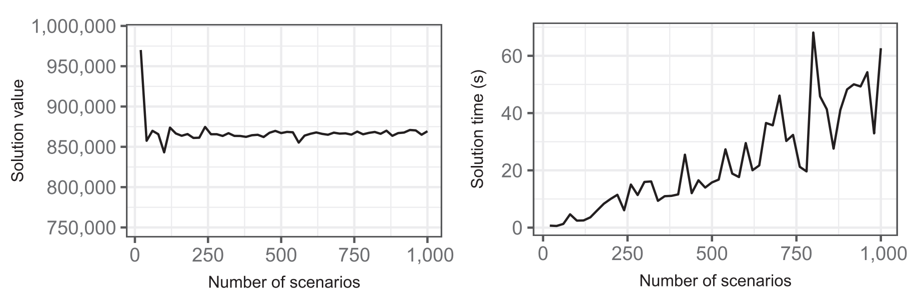

Fig. 5. Computational performance.

Table 6
Operational evaluation of case mix planning results (weekly average).

Id.	Aspects considered at the level of CMP				Prediction (CMP)			Simulation			Evaluation of revenue	
	E1	E2	E3	E4	Revenue (mean)	Patient volume	Runtime (s)	Revenue (mean)	Revenue (sd)	Patient volume	Pred. error	Improvement potential
N	–	–	–	–	901,238	251.0	13.5	801,069	92,010	223.9	13%	7.4%
B	–	–	–	–	934,614	253.0	1.6	850,394	83,990	230.3	10%	1.7%
1	✓	–	–	–	904,862	241.2	2.5	857,123	85,956	230.5	6%	0.9%
2	–	✓	–	–	912,387	238.2	0.8	848,986	103,452	221.0	7%	1.9%
3	–	–	✓	–	919,729	250.2	1.8	841,021	87,580	228.2	9%	2.8%
4	–	–	–	✓	892,905	249.0	48.5	825,848	70,374	230.2	8%	4.6%
5	✓	✓	–	–	877,991	232.3	1.0	853,045	98,950	222.9	3%	1.4%
6	✓	–	✓	–	898,171	240.5	1.9	852,450	89,765	229.4	5%	1.5%
7	✓	–	–	✓	863,508	236.9	318.3	832,702	71,737	229.0	4%	3.8%
8	–	✓	✓	–	904,866	235.4	1.1	862,589	88,791	223.8	5%	0.3%
9	–	✓	–	✓	898,590	233.6	22.5	856,786	101,074	222.5	5%	1.0%
10	–	–	✓	✓	883,582	245.5	59.5	838,222	76,080	230.1	5%	3.1%
11	✓	✓	✓	–	881,658	233.0	1.2	850,594	98,146	225.2	4%	1.7%
12	✓	✓	–	✓	857,852	224.3	13.6	857,093	85,650	224.2	0%	1.0%
13	✓	–	✓	✓	858,327	235.2	236.7	837,527	77,008	228.8	2%	3.2%
14	–	✓	✓	✓	899,648	233.7	20.0	863,062	89,421	223.9	4%	0.3%
15	✓	✓	✓	✓	867,814	227.4	16.5	865,338	91,978	225.9	0%	–

Legend: *N* - Naive approach, *B* - Basic deterministic model, *E*1 - Stochastic demand, *E*2 - Stochastic resource consumption, *E*3 - Stochastic resource availability, *E*4 - master surgery scheduling

The optimal case mix of model 15 is illustrated in Table 7. From an organizational perspective, the case mix is affected by the allocation of operating room slots to departments. Stating targets for volumes, case mix points, or revenues can support the implementation of a coherent case mix strategy. A corresponding master surgery schedule is derived solving the post-processing procedure described in Section 3.2 and illustrated in Table 8.

The following analysis addresses the performance of the models as measured by the simulated revenue and the number of patients. The most potential for improvement is present for the naive approach. This suggests that a change in the case mix policy and allocation of resources might be promising. Even the use of the most basic deterministic case mix planning problem (Model "*B*") suggests compelling benefits. There is a relative improvement of 6% between the status quo and the simulated revenue when applying this case mix planning problem. The sole consideration of time-phased allocation in addition to the basic model produces only moderate results when compared to the other models. The model performing best is the full model which includes all extensions, i.e., model 15. However, comparable results can be achieved with the simpler model 8 which considers stochastic resource consumption and stochastic resource availability. Another simple model in terms of model structure, model 1, is also within a range of 1% of the best performing model. The simulated number of patients varies between the different models without any clear pattern. This discrepancy between revenue and patient volume indicates that the total volume of patients can be an insufficient indicator for the economic success of a hospital. In addition, the mix of patients should be considered.

Besides the analysis of the performance of the case mix planning models when evaluating the model, the accuracy of case mix planning models can be of interest for planning purposes. For example, case mix planning model "*B*" suggests an unrealistically high revenue of EUR 934,614 due to the highly aggregated formulation of the problem. There is a gap of 10% between this predicted value and the value that can actually be achieved when simulating the bottleneck resources. The prediction error is in the range between 6% and 9% for models involving only one type of extension (models 1–4). This value is improved by more than one third on average to values between 3% and 5% when adding another extension. The best results in terms of prediction accuracy are achieved for models incorporating three or more extensions. It is not surprising that there is a general tendency to higher accuracy for models with a higher degree of detail in terms of the number of extensions considered at the level of case mix planning. However, it is remarkable that all models with two extensions are more accurate than any model with only one extension regardless of the type of extension considered. With the exception of model 5, this observation holds similarly when comparing models with three extensions to models with two extensions.

Resource planners should be aware of these gaps so that the future financial performance and patient volumes are not overestimated. For example, Kuo et al. [40] and Mulholland et al. [41] describe aggregated deterministic case mix planning problems. They identify increases in financial performance of 15% and 16%, respectively, when comparing the optimal solution of their case mix planning problem to the observed status quo. These figures are in the same magnitude as our observations when comparing the objective value of the optimal solution of the basic problem "*B*" (EUR 934,614) to the revenue observed in the simulation of the status quo (EUR 801,069) resulting in an estimated improvement of $(934{,}614 - 801{,}069)/801{,}069 \approx 16.7\%$. The actual improvement is only $(850{,}394 - 801{,}069)/801{,}069 \approx 6.2\%$ in our simulation. Similarly, the improvements of financial performance when applying the above mentioned case mix planning approaches in practice can be expected to be lower than the reported figures of 15% and 16%.

Table 7
Optimal case mix in terms of operating room allocation and average weekly targets.

Department	Allocated operating room slots	$K_{OR,j}$	Target volume	Target case mix points	Target revenue (EUR)
Cardiothoracic surgery (CAR)	15	3	36.9	76.2	210,150
General, Visceral, and Transplant Surgery (GVT)	17	3	48.2	54.7	150,860
Gynecology (GYN)	10	4	39.1	33.9	93,406
Neurosurgery (NS)	9	3	22.3	40.7	112,232
Oral and Maxillofacial Surgery (OMS)	3	3	7.5	10.9	30,034
Pediatric Surgery (PED)	3	4	11.9	7.7	21,375
Trauma Surgery (TS)	14	3	40.9	46.5	128,274
Vascular Surgery (VS)	9	3	20.6	44.0	121,484
SUM	80	–	227.4	314.6	867,814

Table 8
Example of a master surgery schedule.

	OR1	OR2	OR3	OR4	OR5	OR6	OR7	OR8	OR9	OR10	OR11	OR12	OR13	OR14	OR15	OR16
Mo	GVT	GVT	GVT	GYN	VS	VS	GYN	GYN	OMS	CAR	CAR	CAR	NS	TS	TS	TS
Tu	GVT	GVT	GVT	GVT	VS	VS	GYN	GYN	PED	CAR	CAR	CAR	NS	NS	TS	TS
We	GVT	GVT	GVT	TS	VS	TS	GYN	GYN	OMS	CAR	CAR	CAR	NS	NS	TS	TS
Th	GVT	GVT	GVT	GYN	VS	VS	GYN	GYN	OMS	CAR	CAR	CAR	NS	NS	TS	TS
Fr	GVT	GVT	GVT	GVT	VS	VS	PED	PED	TS	CAR	CAR	CAR	NS	NS	TS	TS

Considering the gap between simplified aggregated models and results that can be realistically expected to be achieved is also relevant for planning subsequent resources related to patient volumes such as diagnostic services.

Theoretically, higher accuracy and higher performance can also be achieved by adjusting the parameters of simple aggregated deterministic case mix planning models. For example, the upper bounds D_j on patient volumes or the amount of available resource A_r can be modified in the case mix planning phase to identify more realistic case mix plans. The drawback of this approach is that it is a priori not clear which and what magnitude of adjustments leads to improved solutions.

Note that the evaluation of the results is mainly based on revenue which is closely related to the weighted sum of patients. Dexter et al. [42] argue that improving the financial performance in terms of higher-margin services can enable cross-financing of underpaid services or other hospital assets such as digitization or research projects. Additional performance measures as described in Section 3.1 can be applied without major modification of the problem structure. For example, Testi et al. [24] use waiting time of patients as a major criterion to allocate operating room time when planning the case mix.

Concluding, case mix planning is a valuable tool for hospital demand and resource planning in the given problem environment. Applying even the most basic case mix planning approach adds persuasive value. Increasing the model complexity promises only limited additional benefits. The accuracy of case mix planning is considerable higher for more detailed planning problems. Consequently, more detailed models should be considered when using case mix planning as a tool for financial planning. When interpreting the results, a word of caution is necessary. The results depend to a high degree on the processes of the individual hospital and the involved uncertainties. However, the framework can support decisions on selecting convenient and promising models.

5. Conclusion

In this paper, a framework for evaluating the influence of considering stochastic aspects and different levels of aggregation on the performance of case mix planning for hospitals is presented. Stochastic influences are categorized according to whether they relate to demand, resource consumption, and resource availability. Besides stochastic influences, the impact of different aggregation levels with regard to time is considered. We formulate a generic mixed integer problem to case mix planning and compare different approaches proposed in the literature in terms of model structure. The different modeling approaches are compared and applied using data of a large German hospital. The performance of the different models is evaluated using a simulation.

The framework developed in this paper can be used to formulate reasonable case mix strategies. We find that even the consideration of a simple deterministic planning approach can add significant value as opposed to maintaining current allocation schemes. Stochastic influences on the demand proved worthy when planning the case mix. The incorporation of stochastic resource consumption and stochastic resource availability can add additional benefits. Besides the challenge to identify the best performing model, it can also be interesting to identify models that provide a realistic estimation of the revenue that can be generated with certain case mix strategies. The consideration of the weekly structure, including inactive days without surgery in case mix planning did not add persuasive benefit in terms of revenue but increased the accuracy of the revenue suggested by the case mix planning model when compared to the simulation. When transferring the results of the case study to case mix planning in other hospitals, a word of caution is in order. In different hospitals, different types of resources can be scarce and in some hospitals demand can be the major restraining factor while other hospitals are faced with long waiting lines. Thus, the importance of different stochastic influences can vary considerably among hospitals.

Hospitals in Western countries face increasing economic pressure due to changes in demography and technology. In addition, shortages of qualified medical personnel such as nurses or physicians further aggravate the situation. The efficient utilization of available resources takes a key role in providing high quality care at affordable costs. Efficient utilization is not restricted to the efficient organization of workflows but also relates to the efficient allocation of resources to organizational units within the hospital and the choice of a case mix that best fits to the hospital resources and strategy. The implications of controlled alterations of the case mix are often unclear due to the complexity of the involved processes. A simulation of the relevant processes can be a valuable alternative to costly field studies or pilot projects. In addition, hospitals can respond more quickly to current developments when using decision support based on information technology. Current approaches to case mix planning make use of inaccurate approximations of stochastic influences on the consumption of resources due to the complex structure of the resulting non-linear problems. Therefore, the development of advanced solution methods is a promising direction for future research. Regarding model formulations, the development of models making use of different due times for different patient groups can help to derive more realistic case mix targets when planning the case mix.

Appendix A

The following table contains a summary of the notation used in the paper.

Indices and sets	
$r \in \{1, \ldots, m\}$	Resources
$j \in \{1, \ldots, n\}$	Departments
$t \in \{1, \ldots, T\}$	Days of the planning cycle
$i \in \{1, \ldots, I\}$	Operating rooms
$S_j \subseteq \{1, \ldots, I\}$	Operating rooms equipped for surgeries of department j
Parameters	
M_j	Average contribution margin of department j
C^o_{rj}	Overage costs of resource r for department j
C^u_{rj}	Underage costs of resource r for department j
A_r	Available amount of resource r
A^{day}_{rt}	Available amount of resource r on day t
$A^{current}_{rj}$	Amount of resource r currently allocated to department j
$A^{\max}_{rj}$	Maximum amount of resource r that can be allocated to department j
Q_{rj}	Required amount of resource r for treating a patient of department j
$Q^{day}_{rj\tau}$	Required amount of resource r for treating a patient of department j on day t after admission
K_{rj}	Number of slots of resource r to be scheduled by department j
D_j	Demand for department j
Decision variables	
a_{rj}	Amount of resource r allocated to department j
a^{day}_{rjt}	Amount of resource r allocated to department j on day t
y^{day}_{jti}	1 if operating room i is assigned to department j on day t, 0 otherwise
x_j	Number of patients admitted by department j
x^{day}_{jt}	Number of patients admitted by department j on day t
k_{rj}	Number of slots of resource r to be scheduled by department j

References

[1] Busse R, Schreyögg J, Smith PC. Editorial: hospital case payment systems in europe. Health Care Manage Sci 2006;9(3):211–13.
[2] May JH, Spangler WE, Strum DP, Vargas LG. The surgical scheduling problem: current research and future opportunities. Prod Oper Manage 2011;20(3):392–405.
[3] Guerriero F, Guido R. Operational research in the management of the operating theatre: a survey. Health Care Manage Sci 2011;14(1):89–114.
[4] Hulshof PJ, Kortbeek N, Boucherie RJ, Hans EW, Bakker PJ. Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. Health Syst 2012;1(2):129–75.
[5] Hof S, Fügener A, Schoenfelder J, Brunner JO. Case mix planning in hospitals: a review and future agenda. Health Care Manage Sci 2017;20(2):207–20.
[6] McRae S, Brunner JO, Bard JF. Analyzing economies of scale and scope in hospitals by use of case mix planning. Health Care Manage Sci 2019. doi:10.1007/s10729-019-09476-2.
[7] Abdel-Malek LL, Montanari R. On the multi-product newsboy problem with two constraints. Comput Oper Res 2005;32(8):2095–116.
[8] Lau H-S, Lau AH-L. The multi-product multi-constraint newsboy problem: applications, formulation and solution. J Oper Manage 1995;13(2):153–62.
[9] Lau H-S, Lau AH-L. The newsstand problem: a capacitated multiple-product single-period inventory problem. Eur J Oper Res 1996;94(1):29–42.
[10] Murray CC, Gosavi A, Talukdar D. The multi-product price-setting newsvendor with resource capacity constraints. Int J Prod Econ 2012;138(1):148–58.
[11] Nahmias S, Schmidt CP. An efficient heuristic for the multi-item newsboy problem with a single constraint. Nav Res Logist Q 1984;31(3):463–74.
[12] Erlebacher SJ. Optimal and heuristic solutions for the multi-item newsvendor problem with a single capacity constraint. Prod Oper Manage 2000;9(3):303–18.
[13] Turken N, Tan Y, Vakharia AJ, Wang L, Wang R, Yenipazarli A. The multi--product newsvendor problem: Review, extensions, and directions for future research. In: Handbook of newsvendor problems. Springer; 2012. p. 3–39.
[14] Khouja M. The single-period (news-vendor) problem: literature review and suggestions for future research. Omega 1999;27(5):537–53.
[15] Qin Y, Wang R, Vakharia AJ, Chen Y, Seref MM. The newsvendor problem: review and directions for future research. Eur J Oper Res 2011;213(2):361–374.
[16] Olivares M, Terwiesch C, Cassorla L. Structural estimation of the newsvendor model: an application to reserving operating room time. Manage Sci 2008;54(1):41–55.
[17] Dexter F, Ledolter J, Wachtel RE. Tactical decision making for selective expansion of operating room resources incorporating financial criteria and uncertainty in subspecialties' future workloads. Anesth Analg 2005;100(5):1425–1432.
[18] Gupta D. Surgical suites' operations management. Prod Oper Manage 2007;16(6):689–700.
[19] Choi S, Wilhelm WE. On capacity allocation for operating rooms. Comput Oper Res 2014;44:174–84.
[20] Yahia Z, Eltawil AB, Harraz NA. The operating room case-mix problem under uncertainty and nurses capacity constraints. Health Care Manage Sci 2016;19(4):383–94.
[21] Fügener A. An integrated strategic and tactical master surgery scheduling approach with stochastic resource demand. J Bus Logist 2015;36(4):374–387.
[22] Leeftink G, Hans EW. Case mix classification and a benchmark set for surgery scheduling. J Scheduling 2018;21(1):17–33.
[23] Freeman N, Zhao M, Melouk S. An iterative approach for case mix planning under uncertainty. Omega 2018;76:160–73.
[24] Testi A, Tanfani E, Torre G. A three-phase approach for operating theatre schedules. Health Care Manage Sci 2007;10(2):163–72.
[25] Ma G, Demeulemeester E. A multilevel integrative approach to hospital case mix and capacity planning. Comput Oper Res 2013;40(9):2198–207.
[26] Roth AV, Van Dierdonck R. Hospital resource planning: concepts, feasibility, and framework. Prod Oper Manage 1995;4(1):2–29.
[27] Vissers JM, Bertrand J, De Vries G. A framework for production control in health care organizations. Prod Plann Control 2001;12(6):591–604.
[28] Wortman J, Euwe M, TAAL M, Wiers V. A review of capacity planning techniques within standard software packages. Prod Plann Control 1996;7(2):117–28.
[29] Dexter F, Blake JT, Penning DH, Sloan B, Chung P, Lubarsky DA. Use of linear programming to estimate impact of changes in a hospital's operating room time allocation on perioperative variable costs. Anesthesiology 2002;96(3):718–24.
[30] Beliën J, Demeulemeester E. Building cyclic master surgery schedules with leveled resulting bed occupancy. Eur J Oper Res 2007;176(2):1185–204.
[31] Vansteenkiste N, Lamote C, Vandersmissen J, Luysmans P, Monnens P, De Voldere G, et al. Reallocation of operating room capacity using the due-time model. Med Care 2012;50(9):779–84.
[32] Harper PR. A framework for operational modelling of hospital resources. Health Care Manage Sci 2002;5(3):165–73.
[33] Gerchak Y, Gupta D, Henig M. Reservation planning for elective surgery under uncertain demand for emergency surgery. Manage Sci 1996;42(3):321–34.
[34] Adan I, Vissers J. Patient mix optimisation in hospital admission planning: a case study. Int J Oper ProdManage 2002;22(4):445–61.
[35] Sandbaek BE, Helgheim BI, Larsen OI, Fasting S. Impact of changed management policies on operating room efficiency. BMC Health Serv Res 2014;14(1):224.
[36] Bai J, Fügener A, Schoenfelder J, Brunner JO. Operations research in intensive care unit management: a literature review. Health Care Manage Sci 2018;21(1):1–24.
[37] German Interdisciplinary Association for Intensive Care and Emergency Medicine (DIVI). Empfehlungen zur Struktur und Ausstattung von Intensivstationen. https://www.divi.de/empfehlungen/publikationen/intensivmedizin/399-empfehlungen-zur-struktur-von-intensivstationen-langversion/file; 2010. Last accessed May 27, 2018.
[38] Lovejoy WS, Li Y. Hospital operating room capacity expansion. Manage Sci 2002;48(11):1369–87.
[39] Chow VS, Puterman ML, Salehirad N, Huang W, Atkins D. Reducing surgical ward congestion through improved surgical scheduling and uncapacitated simulation. Prod Oper Manage 2011;20(3):418–30.
[40] Kuo PC, Schroeder RA, Mahaffey S, Bollinger RR. Optimization of operating room allocation using linear programming techniques. J Am Coll Surg 2003;197(6):889–95.
[41] Mulholland MW, Abrahamse P, Bahl V. Linear programming to optimize performance in a department of surgery. J Am Coll Surg 2005;200(6):861–8.
[42] Dexter F, Blake JT, Penning DH, Lubarsky DA. Calculating a potential increase in hospital margin for elective surgery by changing operating room time allocations or increasing nursing staffing to permit completion of more cases: a case study. Anesth Analg 2002;94(1):138–42.