



Flexible staffing of physicians with column generation

Melanie Erhard^{1,2}

Published online: 24 April 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In Germany, around 40% of the hospitals do not generate an annual surplus. This leads to an increasing pressure on hospitals' management to reorganize and restructure their processes and resources to decrease the upcoming costs and become profitable. Since personnel, especially physicians, generates a major part of the arising costs, assigning staff efficiently provides an opportunity to decrease associated costs. Up to now, experienced physicians create rosters manually which is cost and time intense due to the problem's complexity and especially the fluctuation in demand. To circumvent this difficulty, it is our main aim to create a new mathematical modeling approach to implement additional flexibility in the rostering process to better match supply and demand. Therefore, we formulate the problem as mixed-integer programming model with the objective to minimize occurring labor costs of physicians over the considered planning horizon subject to coverage of demand to make flexibility monetarily evaluable. In our approach, full flexibility in terms of patterns of working days, shift types, and the placement of the break is provided. To solve the problem under consideration, a column generation heuristic is presented. In our experimental study, the performance of the provided solution approach as well as the effect of additional flexibility in the rostering process are evaluated using real life data. Results indicate the significant impact of implementing flexibility in the scheduling process on the salary costs of the number of required physicians and evidence the superior quality of our solution approach.

Keywords Physicians · Column generation · Flexibility · Hospital · Mixed-integer programming

✉ Melanie Erhard
melanie.erhard@unikat.uni-augsburg.de

¹ Chair of Health Care Operations/Health Information Management, Faculty of Business and Economics, University of Augsburg, Universitätsstraße 16, 86159 Augsburg, Germany

² University Center of Health Sciences at Klinikum Augsburg (UNIKA-T), Neusässer Straße 47, 86156 Augsburg, Germany

1 Introduction

Within the last decades, hospitals face an increasing pressure to become profitable. Almost one third of Germany's hospitals generated a significant loss in 2016. Solely 61% of the hospitals negotiated the year with a surplus. The remaining 10% account a more or less balanced budget (Blum et al. 2017). Since this proportion did only change marginally over the last 10 years and this ongoing negative trend will not stop in near future, the situation is severe (Blum et al. 2007). As a result, hospital's management is forced to counteract the cost pressure it is confronted with, e.g. by reducing upcoming costs. Since more than half of the operating costs are generated by the workforce (especially physicians), this seems to be a leverage point (Statistisches Bundesamt 2016). But, as the number of patients for hospitals is increasing at the same time (Bölt 2014), it is not sufficient to just reduce the number of personnel to save money since this might raise negative consequences on patient care and the quality of the provided service.

Therefore, it is necessary to reduce personnel costs by planning and scheduling staff efficiently to potentially decrease unplanned overtime hours and corresponding expenses (Thungjaroenkul et al. 2007). Up to now, an experienced physician creates schedules commonly manually which is a quite complex task: There are various rules that have to be taken into account, e.g. governmental workforce regulations (Ernst et al. 2004), individual agreements between physicians and hospitals (Brunner et al. 2010), and the length of the schedule as well as personnel requests (Damci-Kurt et al. 2017). Apart from this, it is the main aim to ensure an appropriate number of physicians scheduled to cover demand in each period of every day. In hospitals, demand is liable to heavy fluctuations occurring every day and every hour of the year, which makes it hard to predict in advance. Especially due to peaks in the required level of care, a large number of physicians has to be on duty to handle those (Aggarwal 1982). Since such peaks commonly have a duration of only a few hours, hospitals are overstaffed after its declination. This leads to a lot of idle time during the remaining hours of the day and raises the necessity to construct lines-of-work for employees which better match supply and demand (Brunner and Edenharter 2011). A possible approach to implement more flexibility in the scheduling process is to apply a greater number of shift types having different starting periods and variable lengths instead of solely using predefined 8 h shifts having fixed starting times (Isken 2004). Moreover, additional flexibility can be investigated by employing part time physicians having a reduced number of weekly working hours and can specifically be assigned to cover (expected) peaks in demand (Van den Bergh et al. 2013).

The purpose of this paper is to evaluate flexibility in the physician scheduling process. Our contribution is manifold: We develop a mathematical model that provides full flexibility when scheduling physicians with respect to the patterns of sequences of working days, starting and ending times of shifts and the placement of the break for each physician. Our main objective is to create a (near) optimal schedule that minimizes the total labor costs of physicians assigned to different lines-of-work within the planning horizon. At the same time, the forecasted demand is to be covered and breaks are to be assigned. The problem is formulated as mixed-integer

programming model (MIP) with the objective to minimize total salary costs for personnel and solved applying a column generation (CG) heuristic approach to generate specific lines-of-work for each individual physician. The resulting solution provides a cost minimum staffing policy (when being solved to optimality), or at least a lower bound for estimating the resulting total labor costs. Furthermore, we examine the effect of additional flexibility in the scheduling process on the overall salary expenses and compare this flexible working environment with a more rigid shift system derived from real life.

The paper is structured as follows: In the subsequent section, current literature and research in the field of physician scheduling is reviewed. In Sect. 3, the underlying problem structure is discussed briefly and the stated mathematical model is presented. Afterwards, we present a CG heuristic as appropriate solution approach due to the problem size and complexity (see Sect. 4). Our experimental study is presented in Sect. 5, where the effect of implementing a high degree of flexibility and the assignment of breaks is analyzed. We evaluate the performance of our solution approach and show the positive effect of flexibility in the scheduling process as well as the high quality of the resulting planning decisions. Moreover, we compare our flexible scheduling approach with fixedly working conditions and shift systems which represent real life planning environments to deduce managerial insights which support tactical decision making. Eventually, the main findings and insights are summarized as the paper concludes by identifying some ideas for further research.

2 Overview on current literature

In contrast to the nurse scheduling problem, physician scheduling is still backward in research due to the complexity of the problem reasoned by the multitude of regulations and individual agreements that have to be taken into account. This makes the physician scheduling problem less generalizable and highly constraint. Even though, research has amplified significantly from 1985 up to now, i.e. from one publication in 1985 to nine papers in 2016. In the following, the relevant literature and developments in this field of research are reviewed. For a more detailed bibliography, see Erhard et al. (2018).

In current literature, flexibility in terms of days on and days off duty are mostly considered when scheduling residents, night and weekend shifts, and on-call shifts for residents and/or physicians. Beliën and Demeulemeester (2006) focus on scheduling trainees in a university hospital in Belgium. In their approach, hard constraints, such as labor rules, and soft constraints, e.g. duty preferences in terms of working days, are taken into account. Solving the problem using a Branch-and-Price (B&P) approach results in an optimal schedule of high quality with respect to fulfilled preferences, computational effort, and the length of the planning horizon. A 1-year on-call schedule for residents using five different types of shifts is created in Cohn et al. (2009). Combining professional's expertise with a heuristic solution algorithm leads to an improved schedule which is generated in significantly reduced amount of computation time. A long term planning horizon is also considered in Brunner and Edenharter (2011): To determine the optimal number of personnel required over one

entire year, a mixed-integer programming model is stated and solved applying a column generation heuristic. Since the number of available shifts is generated implicitly by the mathematical formulation, almost a maximum level of flexibility in shift types is provided.

Implicitly generated flexible shifts are defined by their starting and ending period having various different allowed shift lengths. These are also in use in Brunner et al. (2009): To minimize labor costs, a mixed-integer program is formulated and solved using a decomposition heuristic dividing the planning horizon in weekly subproblems. As a result, schedules of high quality reducing staffing costs are created in short amount of time. This research is developed further by Brunner et al. (2010) to analyze the effect of applying different branching strategies within the proposed B&P approach. Focusing on a similar objective function, it is the aim of Stolletz and Brunner (2012) to minimize paid out costs and maximize fairness for physicians. In their approach, flexible shifts are also in use. In contrast to Brunner et al. (2009), shift types were generated by a shift matrix in a preprocessing step and serve as input for the stated reduced set covering formulation. Applying a heuristic which decomposes the considered problem by week leads to a high quality schedule violating significantly less soft restrictions in an appropriate solution time. Investigating uncertainty in terms of patient arrivals in their problem formulation, Ganguly et al. (2014) consider three different emergency departments. Scheduling physicians with different skill levels, exact results of their MIP formulation indicate potential for balanced staffing costs and levels of patients to improve staffing policies. Also considering an emergency department, it is the major aim of El-Rifai et al. (2015) to minimize the waiting time of patients while providing a fair distribution of shifts for physicians at the same time. Their stochastic mixed-integer programming formulation is solved exact and the results are subsequently evaluated using a simulation software. Results indicate potential improvement and differences for the considered staffing strategies.

In contrast, the assignment of breaks is almost neglected in current literature. Even though the positive effect of an appropriate number of break periods within a working day on the performance of personnel is well known (Janaro and Bechtold 1985), recent research hardly deals with it. In general staff scheduling literature, Bechtold and Jacobs (1990) were one of the first to investigate the assignment of breaks to increase utilization of personnel. Therefore, a linear program (LP) is formulated to implicitly assign rest periods flexibly. Comparing their approach to an explicit set-covering formulation results in improved performance with respect to solution time and the number of generated integer solutions, even for large test instances. Thompson (1995) also assigns shifts and breaks implicitly. Extending the formulation of Bechtold and Jacobs (1990) by a larger number of available shift types and the consideration of overtime hours, optimal solutions are created in a shorter amount of time without violating operational constraints on the placement of the break. Adopting the set covering formulations developed by Dantzig (1954), Aykin (1996) stated an integer programming (IP) model to assign multiple rest and lunch breaks to break windows. Considering a predetermined small number of shifts, every feasible combination for a break to shift assignment is defined. Evaluating their modeling approach for five different demand scenarios leads to the

insight that this is a suitable approach to solve larger instances to optimality. Further extension is conducted in Rekik et al. (2010) by considering fractional breaks. When focusing on physician scheduling literature, the situation is even severe: Up to now, solely four papers take breaks into account. Brunner et al. (2009, 2010) propose an implicit approach for the assignment of breaks to physicians whereas Stolletz and Brunner (2012) assign breaks explicitly by a predetermined shift matrix. Comparing implicit and explicit modeling formulation leads to the insight, that an explicit reduced set covering formulation outperforms the implicit modeling approach with respect to solution times and the number of required constraints and variables. A combined approach is in use in Erhard and Brunner (2017): Considering predefined flexible overlapping shifts in a set covering problem formulation, breaks are assigned implicitly by an extended formulation of Bechtold and Jacobs (1990). Evaluating the effect of taking breaks in the scheduling process into account leads to the insight, that even on a strategic level those are especially important. Not considering an appropriate number of breaks in staffing decisions therefore leads to a significant underestimation of the size of the workforce which is required to cover demand. This may lead to a lot of overtime hours and dissatisfaction of physicians, especially under real world circumstances where rather low flexibility in shift types is provided. This effect decreases the more flexibility is implemented in the scheduling process.

To the best of our knowledge, there exists no approach in literature that considers all features of flexibility with respect to the patterns of sequences of working days, starting and ending times of shifts, and the placement of flexible breaks. With our research we close the gap.

3 Problem description and statement of the mathematical model

The problem under consideration is to minimize total salary costs of the required workforce to cover demand for a given planning horizon consisting of $|W|$ weeks, each of them comprising of a set of D days. Each of these days d spans $|P|$ periods, e.g. stated in 1-h increments. To create individual lines-of-work, a set of physicians I is available to be assigned to various shift types $s \in S$. Each physician i can be modeled with his specific characteristics and individual restrictions that are either stated in the labor contract or base on special agreements between the physicians and the hospital they work for. In the following, we consider a homogeneous group of physicians, i.e. assuming similar qualification and experience levels as well as identical labor contracts and working regulations. However, we could also model each physician by individual characteristics, e.g. by implementing part- and fulltime physicians having different working hour regulations: Fulltimers might for example have a higher maximum number of allowed working hours each week which can be integrated by defining \bar{R} individually for each physician, i.e. \bar{R}_i . Other labor aspects can be handled similarly to account for a heterogeneous group of physicians.

Flexibility in terms of working patterns is implemented by the minimum number of consecutive days on duty D^{on} as well as the maximum number of consecutive

days on duty \bar{D}^{on} . Between each consecutive sequence of days on duty for each physician i , a minimum number of days off duty \underline{D}^{off} has to be ensured.

Flexibility in shifts is provided by the various available shift types $s \in \mathcal{S}$. These are determined by specific characteristics, such as the minimum and maximum allowed shift length, i.e. S^{min} and S^{max} . In our case, this results in a shift stretch of 7–12 h as a maximum value since we define the length of a period to correspond to 1 h. The available set of shifts \mathcal{S} is predetermined by a shift matrix which serves as input for the mathematical model by the binary parameter A_{sp} . This parameter indicates if a period p is a working period for a specific shift type s by being equal to 1. Otherwise, the parameter takes 0 as its value for the remaining periods. Figure 1 provides an example how a shift matrix could look like for $S^{min} = 7, S^{max} = 12$, and $|\mathcal{P}| = 20$.

Figure 1 shows an exemplary shift matrix for a minimum shift length of seven and a maximum shift length of twelve periods. As can be seen, a 1 in a cell denotes a working period within a shift whereas a 0 indicates an off period. For example shift two has a length of seven working periods, which starts in period two and has its last working period in period 8.

The binary decision variable x_{isd} ensures the assignment of a specific shift to an individual physician for each day on duty. The resulting shift schedule for each physician i is created in a way that does not exceed a maximum number of weekly working hours \bar{R} which is determined by labor regulations. Moreover, not any two shifts are allowed to be assigned consecutively for each physician i since a minimum number of rest periods P^{rest} between two shift assignments has to be ensured. Note,

		SHIFT s								
A_{sp}		1	2	3	4	5	6	7	...	$ \mathcal{S} $
PERIOD p	1	1	0	0	0	0	0	0	...	0
	2	1	1	0	0	0	0	0	...	0
	3	1	1	1	0	0	0	0	...	0
	4	1	1	1	1	0	0	0	...	0
	5	1	1	1	1	1	0	0	...	0
	6	1	1	1	1	1	1	0	...	0
	7	1	1	1	1	1	1	1	...	0
	8	0	1	1	1	1	1	1	...	0
	9	0	0	1	1	1	1	1	...	1
	10	0	0	0	1	1	1	1	...	1
	11	0	0	0	0	1	1	1	...	1
	12	0	0	0	0	0	1	1	...	1
	13	0	0	0	0	0	0	1	...	1
	14	0	0	0	0	0	0	0	...	1
	15	0	0	0	0	0	0	0	...	1
	16	0	0	0	0	0	0	0	...	1
	17	0	0	0	0	0	0	0	...	1
	18	0	0	0	0	0	0	0	...	1
	19	0	0	0	0	0	0	0	...	1
	20	0	0	0	0	0	0	0	...	1

Fig. 1 Example for possible line-of-work

we are considering a discontinuous planning problem since shifts are not allowed to spill over from one to the subsequent planning day.

Flexibility in the placement of breaks within shifts is provided by the integrated break window for the individual assignment of a break in a period p for each physician i working shift s on day d . The size of the break window is determined by the length of the assigned shift and the minimum number of working periods after a shift has started and before the shift ends, i.e. B^{pre} and B^{post} . The resulting lines-of-work are assigned in a way that ensures demand N_{dp} to be covered in every period p of every day d in the planning horizon.

An example of a line-of-work for one single physician is provided in Fig. 2. We consider a planning horizon of five working weeks Monday to Sunday. Each row denotes a planning period within a (planning) day, i.e. we consider 20 planning periods p . Columns represent the flexible shift assignments for the physician on any day. A 1 denotes a working period. A 0 within a sequence of 1 s indicates a break assignment. For instance, the physician working this line-of-work has a break in period 4 on Monday in the first week. The schedule shows that sequences of working days vary in the first day on duty and the number of consecutive working days, i.e. ranging from four to seven consecutive days on duty. Moreover, shifts start at various planning periods and have different lengths. Each shift has an appropriate break which is flexibly assigned, e.g. in period four on Monday in the first week and in period fourteen on Thursday in the subsequent week. The first positive demand is normalized to period 1.

In the following, we formulate the problem as mixed-integer programming model. In our formulation, we provide a maximum level of flexibility in terms of working days, shift types, and the assignment of breaks. Since we focus on the evaluation and comparison of our mathematical modeling and solution approach, our objective is the minimization of the total salary costs subject to coverage of demand. Appraising the required size of the workforce by occurring labor costs within the planning

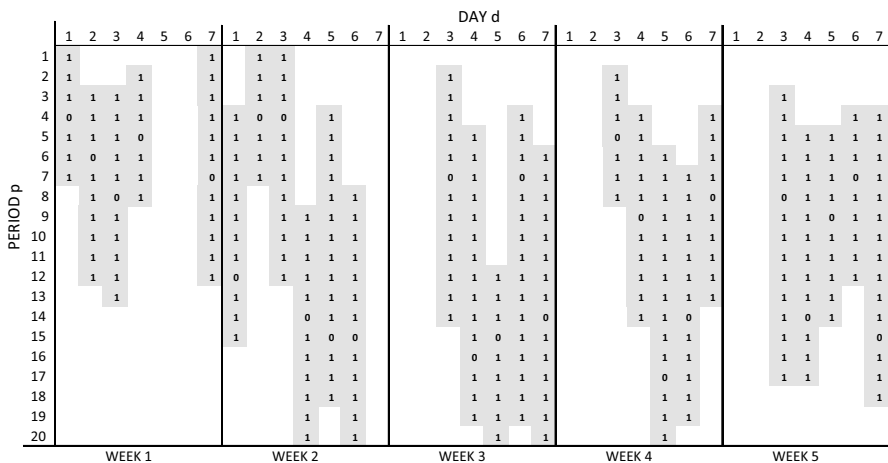


Fig. 2 Example for possible line-of-work

horizon assigns a monetary value to scheduling flexibility. This makes our results more comparable. Therefore, the following binary decision variable is defined, i.e.:

$$Y_i = \begin{cases} 1, & \text{if physician } i \text{ is on duty within the planning horizon} \\ 0, & \text{otherwise.} \end{cases}$$

On a daily basis, we have to decide if a physician works or is off duty. We use a binary variable y_{id} defined as:

$$y_{id} = \begin{cases} 1, & \text{if physician } i \text{ works on day } d \\ 0, & \text{otherwise.} \end{cases}$$

To assure feasible schedules, we define another binary decision variable z_{id} that determines the start of a sequence of working days, i.e.

$$z_{id} = \begin{cases} 1, & \text{if physician } i \text{ starts a sequence of working days on day } d \\ 0, & \text{otherwise.} \end{cases}$$

On each working day, the physician is assigned to a shift. We model this by the following binary variables:

$$x_{isd} = \begin{cases} 1, & \text{if physician } i \text{ is assigned to shift } s \text{ on day } d \\ 0, & \text{otherwise.} \end{cases}$$

Moreover, each physician being assigned to a shift also gets a break assigned which is modeled by:

$$b_{ispd} = \begin{cases} 1, & \text{if physician } i \text{ is assigned to a break in period } p \text{ when working shift } s \text{ on day } d \\ 0, & \text{otherwise.} \end{cases}$$

Now we introduce the essential notation and present the basic model.

Sets with indices

$i \in I$	Set of physicians with index i
$w \in W$	Set of weeks with index w
$d \in D$	Set of days with index d
$d \in D_w$	Subset of days, i.e. Monday to Sunday, within week w with D^{week} as the number of days per week, i.e. 7 days, and $D_w = \{(w-1) \cdot D^{week} + 1, \dots, w \cdot D^{week}\}$
$p \in P$	Set of day-periods with index p
$s \in S$	Set of shifts with index s

Parameters

c^{plan}	Salary costs of a physician (depending on the length of the planning horizon)
S^{min}	Minimum shift length
S^{max}	Maximum shift length
A_{sp}	1 if shift s covers period p , 0 otherwise
F_s	First working period in shift s

- L_s Last working period in shift s
- W_s Number of working periods in shift s
- B^{pre} Amount of working periods before the break is allowed to start
- B^{post} Amount of working periods after the break has ended
- D^{on} Minimum number of consecutive days on
- \bar{D}^{on} Maximum number of consecutive days on duty
- D^{off} Minimum number of days off between consecutive sequences of working days on duty
- P^{rest} Minimum number of rest periods between two consecutive shifts
- \bar{R} Maximum amount of regular working periods for physician i per week
- N_{dp} Demand in period p of day d physician i

Binary decision variables

- Y_i 1 if physician i works in the planning horizon, 0 otherwise
- y_{id} 1 if physician i is on duty on day d , 0 otherwise
- z_{id} 1 if the sequence of working days of physician i begins on day d , 0 otherwise
- x_{isd} 1 if physician i is assigned to shift s on day d , 0 otherwise
- b_{isdp} 1 if physician i working shift s in period p of day d is assigned a break, 0 otherwise

$$\text{Minimize } \sum_{i \in I} c^{plan} \cdot Y_i \tag{3.1}$$

subject to

$$\sum_{s \in S} \sum_{i \in I} (A_{sp} x_{isd} - b_{isdp}) \geq N_{dp} \quad \forall d \in D, p \in P \tag{3.2}$$

$$y_{id} \leq Y_i \quad \forall i \in I, d \in D \tag{3.3}$$

$$z_{id} = y_{id} (1 - y_{i(d-1)}) \quad \forall i \in I, d \in D \tag{3.4}$$

$$\sum_{t=d}^{d+\bar{D}^{on}} y_{it} \leq \bar{D}^{on} \quad \forall i \in I, d \in D \tag{3.5}$$

$$z_{id} \leq y_{it} \quad \forall i \in I, d \in \{1, \dots, |D| - D^{on}\}, t \in \{d, \dots, d + D^{on} - 1\} \tag{3.6}$$

$$z_{id} \leq 1 - y_{it} \quad \forall i \in I, d \in \{1 + D^{off}, \dots, |D| - D^{on} + 1\}, t \in \{d - D^{off}, \dots, d - 1\} \tag{3.7}$$

$$\sum_{s \in S} x_{isd} = y_{id} \quad \forall i \in I, d \in D \tag{3.8}$$

$$x_{is_1d} + x_{is_2(d+1)} \leq 1 \quad \forall i \in I, s_1, s_2 \in S : |P| - L_{s_1} + F_{s_2} - 1 < P^{rest}, d \in D \setminus |D| \tag{3.9}$$

$$\sum_{s \in S} \sum_{d \in D_w} W_s x_{isd} \leq \bar{R} \quad \forall i \in I, w \in W \tag{3.10}$$

$$\sum_{p=F_s+B^{pre}}^{L_s-B^{post}} b_{isdp} = x_{isd} \quad \forall i \in I, s \in S, d \in D \tag{3.11}$$

$$Y_i, y_{id}, z_{id}, x_{isd}, b_{isdp} \in \{0, 1\} \quad \forall i \in I, s \in S, d \in D, p \in P \tag{3.12}$$

The objective function (3.1) minimizes the total salary costs of the number of employed physicians.

Flexible days The first block of constraints (3.2)–(3.7) models flexible working days. Constraints (3.2) take care about demand being covered in every period p of every day d within the planning horizon. It is therefore necessary to reduce the number of assigned physicians in each period p by the number of employees having their break in this specific period of the day. As a result, the number of assigned physicians is forced to be greater than or equal to the level of demand in each period of every day, even though some physicians are assigned to a break in this period. Moreover, constraints (3.3) indicate if physician i has at least 1 day on duty during the whole planning horizon, i.e. $\forall d \in D, y_{id} = 1 \rightarrow Y_i = 1$. Constraints (3.4) determine the start of a sequence of consecutive working days (i.e. $z_{id} = 1$) by linking variables y_{id} with z_{id} . In other words, if there is a switch from 0 to 1 in the y -variables then the corresponding z -variable is set to 1. Those constraints can be easily linearized (see “Appendix”). The next constraints (3.5) force a maximum number of consecutive working days \bar{D}^{on} . In other words, in any sequence of $(\bar{D}^{on} + 1)$ consecutive y -variables at least one of those must be 0; otherwise the constraints are not fulfilled. Next, we enforce a minimum number of \underline{D}^{on} consecutive working days for each valid schedule. If a sequence of working days starts on day d , i.e. $z_{id} = 1$, then constraints (3.6) assure that the y -variables for the days from d to $d + \underline{D}^{on}$ are bound from below by 1. Mathematically speaking, the constraints ensure the following: $\forall i \in I, d \in D, z_{id} = 1 \rightarrow y_{it} = 1$ for $t \in \{d, \dots, d + \underline{D}^{on} - 1\}$. Ensuring a minimum number of days off after each sequence of days on, we use a similar modeling idea (see constraints 3.6). To be able to utilize the z -variables, we enforce the appropriate number of off days before a sequence of days on starts; otherwise we would need additional variables indicating the end of a working sequence. In particular, constraints (3.7) assure a minimum number of \underline{D}^{off} days off between two consecutive sequences of working days. As before, if a sequence of working days starts on day d , i.e. $z_{id} = 1$, then constraints (3.7) model that the y -variables for the days from $d - \underline{D}^{off}$ to $d - 1$ are forced to 0, otherwise the constraints would not be valid, i.e. $\forall i \in I, d \in D, z_{id} = 1 \rightarrow y_{it} = 0$ for $t \in \{d - \underline{D}^{off}, \dots, d - 1\}$. If one of the y_{it} variables before a working sequence starts is equal to 1, the constraints (3.7) are not

fulfilled since the right hand side would be 0 as well but the left hand side would be 1 due to $z_{id} = 1$. This is a contradiction.

Flexible shifts After the conditions of the various sequences of working days have been regulated, the allowed distribution of shift types to each individual physician i is to be considered. The second set of constraints (3.8)–(3.9) handles the flexible assignment of shifts: Constraints (3.8) ensure that each physician being on duty on a specific day d gets exactly one shift s assigned. If a day is not a working day for a physician, i.e. $y_{id} = 0 \rightarrow x_{ids} = 0, \forall i \in I, d \in D, s \in S$. To ensure a minimum number of rest periods P^{rest} between two consecutive shift assignments, constraints (3.9) do not allow some specific combinations of shift assignments on day d and $d + 1$ where off periods are less than P^{rest} . Not any number of working hours are permitted within a week for each physician due to labor regulations and individual agreements of hospitals and physicians. Therefore, constraints (3.10) ensure a maximum number of working hours within a week for each physician. Depending on the number of assigned shifts and corresponding working hours for each shift W_s , constraints (3.10) sum up all duties within 1 week w for each physician i and force those not to exceed a maximum value \bar{R} .

Flexible breaks Constraints (3.11) ensure exactly one break assignment for each physician on duty. In particular, the break is to be placed within a predetermined break window defined by the minimum number of working hours before a break is allowed to be placed B^{pre} and the minimum number of working periods after a break B^{post} individually for each assigned shift type s . Eventually, variable definitions are given in (3.12).

4 Solution approach

Since standard software (IBM ILOG OPL Studio 8.0 and CPLEX 12.0) is not able to find a feasible solution for the majority of parameter settings, planning horizons of more than 1 week, and demand scenarios except for the mean demand level, a CG heuristic is applied to find good integer solutions in reasonable runtime (Desaulniers et al. 2005). As can be seen easily, our model decomposes by physician. Hence, an individual line-of-work covering the entire planning horizon is constructed, i.e. specifying sequences of working days, days off, the assigned shift type for each day on duty, and the period of the break within the assigned shift.

Due to the specific block structure of our problem, Dantzig-Wolfe decomposition is applied (Dantzig and Wolfe 1960; Desrosiers and Lübbecke 2005). The general procedure of the solution algorithm is shown in Fig. 3.

Therefore, we decompose our initial model and formulate a Master Problem (MP) and a Subproblem (SP) for each physician (Dantzig and Wolfe 1960). In contrast to other decomposition approaches, we do not need a convexity constraint which is reasoned by the following: As we assume a homogenous group of physicians, we have $|I|$ identical Subproblems. This results in a bounded solution space, i.e. a finite number of available working patterns and a set of extreme points which is equal to the set of available lines-of-work. Moreover, the size of staff $|I|$ is not predetermined as input parameter: Instead, we are implicitly deciding about the required number of

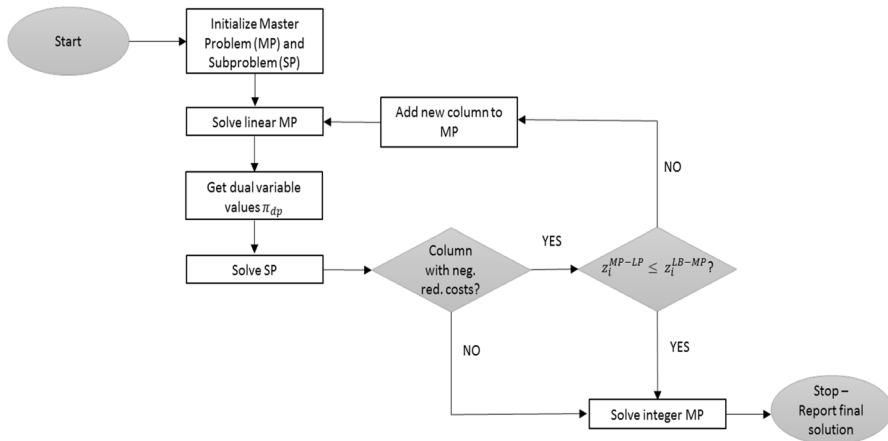


Fig. 3 Flowchart column generation heuristic

physicians by minimizing the total labor costs in our objective function (Barnhart et al. 1998; Desrosiers and Lübbecke 2011; Vanderbeck 2011). Since we solve the extended version using CG, we have to relax the integrality constraints for our linear MP (MP–LP). Considering only a subset of feasible columns, the resulting restricted MP (RMP) does not contain the integrity of all feasible solutions, solely the specific ones which are generated by solving the SP. This leads to a significant reduction in solution times (Lübbecke and Desrosiers 2005). Consequently, two separate optimization problems are formulated and solved iteratively during the process of CG.

The SP in this coherence serves as a generator of new columns (variables) to enter the basis (Barnhart et al. 1998). Here, the physicians are aggregated as the SP creates feasible lines-of-work that are to be assigned to a number of physicians which results in an additional reduction in symmetry. After solving the linear relaxation of the RMP, the resulting dual variable values (π_{dp}) of the demand constraints (4.3) are consigned to the SP which is subsequently solved to optimality as integer program (IP) to determine a new promising column (Desrosiers and Lübbecke 2005). Generated columns in this context are evaluated according to their reduced cost which are to be negative in our minimization problem to generate a positive effect on the objective function value of the RMP. Since our SP is solved to optimality as well, we are able to identify the one column that generated the largest improving (in our case: minimizing) effect on the objective function (Brunner and Edenharter 2011). Afterwards, the new detected column is added the RMP which is solved again. There are two termination criteria for the algorithm: First, the algorithm terminates if no column that prices out can be found. Second, a feasible lower bound for the MP–LP (LB–MP) is implemented to improve the algorithm’s performance (Volland et al. 2017). Since a large computational effort is investigated to prove for optimality when solving MP during the CG heuristic, the so-called tailing-off effect leads to a significant growth in solution times (Gilmore and Gomory 1963). To state

LB–MP, we use information derived in each iteration of the solution process, i.e. the current objective function values of the MP–LP (z_i^{MP-LP}) and the SP (z_i^{SP}). Based on Lübbecke and Desrosiers (2005), the bound is stated in (4.1).

$$z_i^{LB-MP} = \max\{z_{i-1}^{LB-MP}, z_i^{LB-MP}\} \tag{4.1}$$

Calculating the lower bound in each iteration of the solution process, we initialize $z_0^{LB-MP} = -\infty$ and define $z_i^{LB-MP} = \left\lceil \frac{z_i^{MP-LP}}{1 - z_i^{SP}} \right\rceil$. The algorithm terminates due to MP–LB if $z_i^{LB-MP} \geq z_i^{MP-LP}$, i.e. once the lower bound for MP–LP is larger than or equal to the current MP objective function value, before finding the optimal MP–LP solution. This is in around one third of our test instances the case.

As a result, the optimal solution for the linear RMP is simultaneously the optimal solution to the linear MP. But, up to now, the provided optimal solution might be continuous and provides therefore a lower bound. Hence, as last step of our CG heuristic, we solve the RMP as IP to provide a feasible integer solution, i.e. an upper bound. The thereby provided solution must not be optimal any longer. The resulting gap between the solution provided by CG which serves as lower bound (LB) and the objective function value provided by solving the RMP as IP [serving as upper bound (UB)] is small as will be shown soon.

Additional notation for MP and SP:

Sets with indices

$r \in \mathbf{R}$ Set of working rosters with index r

Parameters

X_{dp}^r 1 if roster r covers period p on day d , 0 otherwise

Decision variables

λ_r Number of physicians working roster r in the planning horizon

y_d 1 if day d is a day on duty (working day), 0 otherwise

z_d 1 if the sequence of working days begins on day d , 0 otherwise

x_{sd} 1 if shift s is assigned on day d , 0 otherwise

b_{sdp} 1 if shift s gets a break assigned in period p on day d , 0 otherwise

Mathematical Model—Master Problem

$$\text{Minimize } \sum_{r \in \mathbf{R}} c^{plan} \cdot \lambda_r \tag{4.2}$$

subject to

$$\sum_{r \in \mathbf{R}} X_{dp}^r \lambda_r \geq N_{dp} \quad \forall d \in \mathbf{D}, p \in \mathbf{P} \quad (\pi_{dp}) \tag{4.3}$$

$$\lambda_r \geq 0 \text{ and integer} \quad \forall r \in \mathbf{R} \tag{4.4}$$

The objective function (4.2) minimizes the total labor costs of personnel being assigned to a roster over the planning horizon.

Constraints (4.3) ensure that the number of assigned physicians in every period p is as least as high at the required level of care (or higher) such that the arising demand is covered throughout the entire planning horizon. The last constraints (4.4) define the newly introduced decision variable λ_r which determines the number of physicians working a specific roster $r \in \mathbf{R}$.

The dual variables $\pi_{dp} \geq 0$ derived from constraints (4.3) determine the generic reduced cost of a MP column as given in (4.5).

$$c^{plan} - \sum_{d \in \mathbf{D}} \sum_{p \in \mathbf{P}} \pi_{dp} X_{dp}^r \tag{4.5}$$

In SP notation, the reduced costs are stated in (4.6).

$$c^{plan} - \sum_{s \in \mathbf{S}} \sum_{d \in \mathbf{D}} \sum_{p \in \mathbf{P}} \pi_{dp} \cdot (A_{sp} x_{sd} - b_{sdp}) \tag{4.6}$$

In the following, we state our generic SP. Mathematical Model—Subproblem

$$\text{Minimize } c^{plan} - \sum_{s \in \mathbf{S}} \sum_{d \in \mathbf{D}} \sum_{p \in \mathbf{P}} \pi_{dp} \cdot (A_{sp} x_{sd} - b_{sdp}) \tag{4.7}$$

subject to

$$z_d = y_d(1 - y_{(d-1)}) \quad \forall d \in \mathbf{D} \tag{4.8}$$

$$\sum_{t=d}^{d+\bar{D}^{on}} y_t \leq \bar{D}^{on} \quad \forall d \in \mathbf{D} \tag{4.9}$$

$$z_d \leq y_t \quad \forall d \in \{1, \dots, |\mathbf{D}| - \bar{D}^{on}\}, t \in \{d, \dots, d + \bar{D}^{on} - 1\} \tag{4.10}$$

$$z_d \leq 1 - y_t \quad \forall d \in \{1 + \bar{D}^{off}, \dots, |\mathbf{D}| - \bar{D}^{on} + 1\}, t \in \{d - \bar{D}^{off}, \dots, d - 1\} \tag{4.11}$$

$$\sum_{s \in \mathbf{S}} x_{sd} = y_d \quad \forall d \in \mathbf{D} \tag{4.12}$$

$$x_{s_1 d} + x_{s_2 (d+1)} \leq 1 \quad \forall s_1, s_2 \in \mathbf{S} : |P| - L_{s_1} + F_{s_2} - 1 < P^{rest}, d \in \mathbf{D} \setminus |\mathbf{D}| \tag{4.13}$$

$$\sum_{s \in \mathbf{S}} \sum_{d \in D_w} W_s x_{sd} \leq \bar{R} \quad \forall w \in \mathbf{W} \tag{4.14}$$

$$\sum_{p=F_s+B^{pre}}^{L_s-B^{post}} b_{sdp} = x_{sd} \quad \forall s \in \mathbf{S}, d \in \mathbf{D} \tag{4.15}$$

$$z_d, y_d, x_{sd}, b_{sdp} \in \{0, 1\} \quad \forall s \in \mathbf{S}, d \in \mathbf{D}, p \in \mathbf{P} \quad (4.16)$$

Searching for a new promising column in each iteration, dual values π_{dp} derived from the demand constraints (4.3) are required. Therefore, the objective function of the corresponding Subproblem (4.7) is to minimize the reduced cost of a new column.

The constraints of our Subproblem correspond to constraints (3.4)–(3.11) which are already discussed in detail in former sections. Therefore, these are not explained any further. Note, we have dropped index i . Constraints (4.16) define the decision variables.

In each iteration, the Subproblem generates specific lines-of-work for generic physicians, including an individual sequence of working days, the assignment of a shift for each day on duty, and the flexible assignment of a break during this shift. In our case, the optimal values of $z_d, y_d, x_{sd}, b_{sdp}$ for all $d \in \mathbf{D}, p \in \mathbf{P}$ are combined to construct a new column r :

$$X_{dp}^r = \left[\begin{array}{c} c^{plan} \\ \sum_{s \in \mathbf{S}} (x_{sd} - b_{sdp}) \end{array} \right]$$

Based on the concept of column generation, only columns having negative reduced costs are added to RMP to improve the objective function value, i.e. revealing an objective function value of the SP smaller than 0.

5 Experimental study

In this section, the performance of our heuristic solution approach in comparison to an exact solution (where disposable) is evaluated by using real life data, i.e. aggregated demand patterns. Therefore, the compact formulation of our model as well as the solution approach are implemented in IBM ILOG OPL Studio 8.0 and CPLEX 12.0. All computations were executed on a 2.60 GHz Intel(R) Xeon(R) CPU E5-2650 v2 Machine with 8 GB RAM running under the Windows 10 Enterprise operating system. The SP is solved to optimality for each test instance while at the end of the algorithm a time limit is set to 600 s when solving the RMP as IP.

5.1 Underlying demand pattern

In our experimental study, we use data derived from a large teaching hospital in Germany. In more detail, the occurring demand of an operating theater over one entire year is analyzed. We aggregate the demand (from Monday to Sunday) over 1 year and build three different demand scenarios, i.e. mean demand pattern (50% quantile), 75% quantile, and maximum demand (100% quantile). Depending on the part of the experimental study, these different scenarios are in use: In the first part of the study, we analyze the effect of different lengths of the planning horizon by using mean demand level. Similarly, the 50% quantile demand pattern is in use to test for

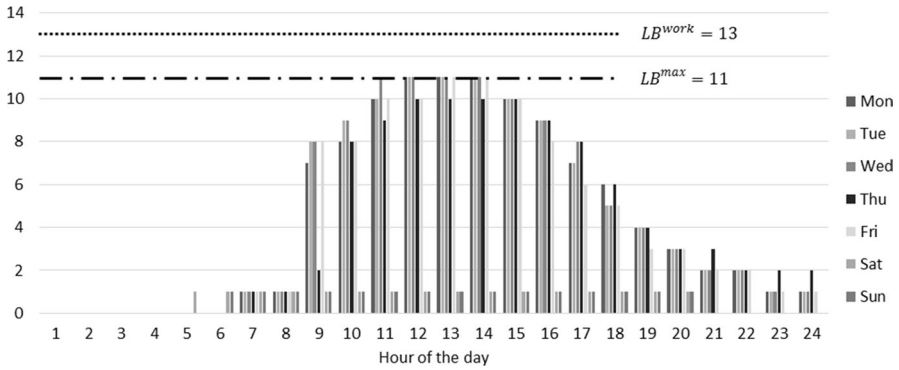


Fig. 4 Real life demand—mean (50% quantile) demand scenario

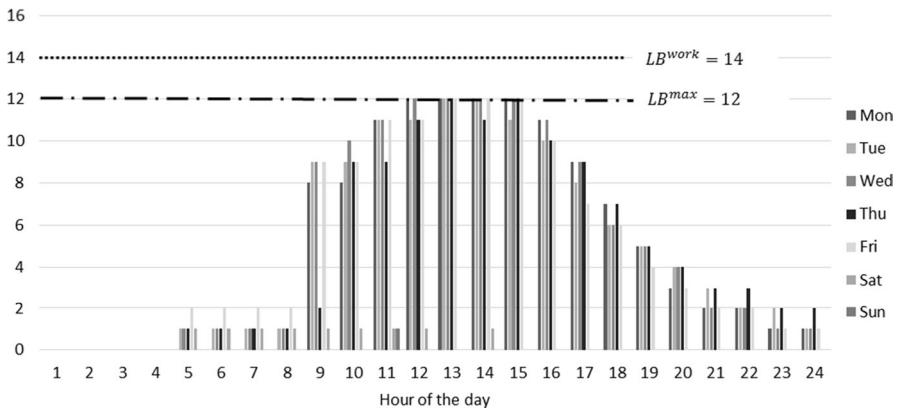


Fig. 5 Real life demand—75% (quantile) demand scenario

the effect of various types and levels of flexibility in our factorial analysis (third part of the experimental study) and in the fourth step of our study when comparing our flexible scheduling environment with a more rigid shift system derived from real life surroundings. In contrast, the second step of the conducted study investigates robustness by implementing the previously mentioned scenarios (see Figs. 4, 5, 6). Moreover, Figs. 4, 5 and 6 exhibit two different bounds for approximating the required workforce size in addition to the structure of the considered demand scenarios. The workload bound (LB^{work}) is determined by the maximum amount of allowed working periods per physician per week and the maximum demand bound (LB^{max}) which is defined by the overall peak in demand within the planning horizon (for more detail we refer to Brunner and Edenharter 2011). Note, for each test instance, the required workforce size provided by the workload bound is larger than provided by the maximum demand bound which might lead to a significant increase in solution times. According to Brunner and Edenharter (2011), a personnel scheduling problem becomes NP-hard, if $LB^{work} \geq LB^{max}$. This means, LB^{work} is the binding bound

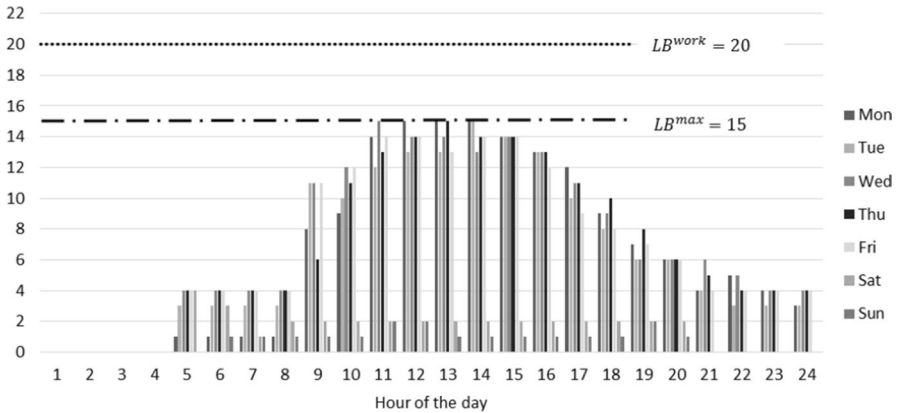


Fig. 6 Real life demand—maximum (100% quantile) demand scenario

for the overall workforce size due to the rather low level of the peak in demand (for a detailed proof see Brunner and Edenharter 2011). We include requirement for care on weekend days in our data to be able to evaluate our theoretical approach especially with respect to flexibility in sequences of working days, even though, this specific demand is covered by on-call shifts in practice.

Decomposing annual demand in 1 week planning periods (Monday to Sunday), demand of each hour of every day is aggregated. As can be seen easily, each day consists of 24 1-h periods. In general, demand starts occurring at 5 am and lasts until 12 am, as we consider a stretch of 20 periods for the upcoming demand. At the beginning of the day, the requirement for care is rather low, but increases significantly at 9 am to its maximum value around noon. During this time of the day, demand is fairly constant, at least for 4 h. Subsequently, the required level of care decreases slowly to its minimum value at 12 am. For the specific case of our mean demand scenario, there is barely no requirement for care in the first 2 h, i.e. 5 am and 6 am. This is not the case for the 75% (quantile) and the maximum demand level (100% quantile). Additionally, the maximum demand varies between the considered scenarios, i.e. eleven for the mean demand scenario, twelve for 75% quantile demand scenario, and 15 for the maximum demand scenario (100% quantile) respectively.

Remaining hours were not taken into account when creating our mean demand level since demand during these hours is commonly handled by the assigned on-call physicians. As we are focusing on a discontinuous planning problem, shifts do not spill over to the next planning day. Our planning day is defined from 5 am to midnight.

To make our results comparable, we additionally define a standard parameter setting which is in use throughout the entire experimental study, except in Sect. 5.3 where we conduct a factorial analysis to determine the effect of an increasing respectively decreasing level of flexibility in the scheduling process. For more detail, see Table 1.

As a result, created lines-of-work have to assign at least three but at most six consecutive working days to a physician. Moreover, between two subsequent sequences

Table 1 Standard parameter setting

Day parameter	$\bar{D}^{on} = 3$	$\bar{D}^{off} = 6$	$D^{off} = 2$	
Shift parameter	$S^{min} = 7$	$S^{max} = 12$	$F_s = 1 \dots 14$	$L_s = 7 \dots 20$
Break parameter	$B^{pre} = 3$		$B^{post} = 3$	
Ergonomic/Legal regulated parameter	$\bar{R} = 45$	$W_s = 6 \dots 11$	$P^{rest} = 11$	$c^{plan} = W \cdot 1'625$

of working days, a minimum of 2 days has to be off. Since the assignment of a break is not mandatory until a shift length of 6 h, we define the minimum length for our shifts to be seven periods, i.e. 7 h. The maximum allowed shift length is 12 h. The remaining shift dependent parameter values are determined implicitly by the defined values for S^{min} and S^{max} , such as the binary values for the shift matrix A_{sp} , the first and last working period for each shift (F_s and L_s respectively), and the number of on duty periods for each shift type W_s .

In our approach, only one type of break with a length of 1 h is considered. But, this formulation can easily be modified and extended, e.g. by varying the length of the assigned break. The placement is determined according to the break parameters. In our standard parameter setting, a physician being assigned to a shift has to work at least three periods ($B^{pre} = 3$) until a break is assigned but no later than 3 h before the shift ends ($B^{post} = 3$).

The subsequent set of parameters ensures a minimum number of rest periods between two consecutive shifts by defining $P^{rest} = 11$, as regulated by law as well as a maximum amount of weekly working hours for each physician, i.e. $\bar{R} = 45$. Moreover, the weekly average salary is assumed to be 1625€ (Marburger Bund 2016) which are to be multiplied by the number of weeks in the planning horizon to determine the labor cost of one entire roster. Note, in our experimental study, we suppose a homogeneous group of physicians.

5.2 Evaluation of solution approach

In this subsection, the performance of our solution approach in comparison to the exact solution of our compact formulation is evaluated. To do so, the standard parameter setting as well as the three predefined demand scenarios are in use. First, various lengths for the planning horizon are analyzed with respect to solvability, solution times, and the quality of the provided solution. Second, a labor cost minimal schedule is determined for the different demand patterns by solving the compact formulation to optimality (if possible) and by applying the suggested CG heuristic.

The performance of the proposed solution approach is tested by implementing various lengths for the planning horizon, i.e. 1, 2, 4, and 6 weeks. Here, the standard parameter setting as well as the mean demand pattern serve as input. For more detail, see Table 2.

Table 2 presents the solution values and algorithm performance measures for the considered CG approach. The first column represents the length of the planning period of the underlying demand pattern. Column two displays the optimal solution

Table 2 Objective value three demand scenarios

	CG					
	Objective MP-LP	Objective MP-IP	Optimality gap	# columns	Solution time: MP-LP	Solution time: MP-IP (s)
1 week roster	24,375	26,000	1625	509	3.8 min.	600
2 week roster	48,750	55,250	6500	1790	43.5 min.	600
4 week roster	97,500	110,500	13,000	2950	5.3 h	600
6 week roster	146,250	165,750	19,500	7654	8.9 h	600

value for MP-LP whereas column three contains the integer solution for the CG heuristic (MP-IP). Column four evaluates the quality of the provided solution by calculating the absolute optimality gap, i.e. $opt.gap(absolute) = (MP - IP) - (MP - LP)$. The last three columns provide information about the number of required columns for the CG heuristic, and solution times for MP-LP and MP-IP.

An exact solution derived by the compact formulation is only provided for a 1 week roster (also implemented in IBM ILOG OPL Studio 8.0 and CPLEX 12.0). The optimal solution of 24,375€ is provided in 2 h solution time. Planning periods of more than 1 week cannot be solved to optimality, i.e. even a feasible solution cannot be found within a time limit of 3 h. Moreover, an optimal solution for the LP relaxation requires more than 3 h for some test instances. With respect to solution times, run times providing an exact solution are more than ten times higher for the MIP compared to our CG heuristic (for the 1 week roster).

In contrast, the CG heuristic is able to find a good solution in acceptable amount of time, i.e. 13 min, 55 min, 5 h, and 9 h for a 6 week planning period respectively. Note, solution times might be significantly lower for test instances with a high peak due to the dominance of LB^{peak} over LB^{work} for these specific test instances. Considering the objective function values, a salary cost minimal roster provided by the CG heuristic ranges from 26,000€ to 165,750€, i.e. corresponding to one respectively two additionally required physicians. But, the smaller value of 26,000€ is only appropriate when focusing on a planning horizon of 1 week which is rather short. Since the resulting minimum salary costs depend on the length of the planning period, the resulting objective value increases the longer the length of the roster.

Figure 7 displays the development of the implemented LB [stated in (4.1)] exemplarily for the CG heuristic when generating a 2-week roster:

The figure shows the lower bound z_i^{LB-MP} for each iteration when running the CG heuristic as well as the maximum value which is in use as early termination criterion during processing the algorithm. This value of the LB is rather small in the first iteration but increases significantly when generating the next column. Afterwards, the LB increases successively.

When solving the different demand scenarios (i.e. 50%, 75%, and 100% quantile) in the second step of the algorithmic evaluation, a planning horizon of 4 weeks is considered for each instance (see Table 3). Table 3 gives the solution and computational values for the CG heuristic. The first column represents the underlying

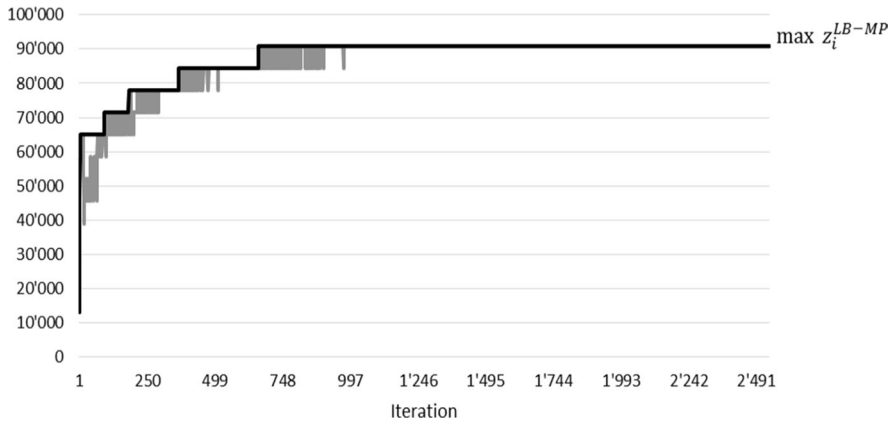


Fig. 7 Real life demand—maximum (100% quantile) demand scenario

Table 3 Objective value three demand scenarios for 2 weeks

	CG					
	Objective MP-LP	Objective MP-IP	Optimality gap	# columns	Solution time: MP-LP	Solution time: MP-IP (s)
50% quantile	97,500	110,500	13,000	2950	5.3 h	600
75% quantile	104,000	117,000	13,000	3133	5.8 h	600
100% quantile	117,000	123,500	6500	3493	6.5 h	600

scenario for the demand pattern. Again, column two to seven provide the objective function values, absolute optimality gap, number of iterations and solution times.

Solving the formulated problem using our CG heuristic for mean demand, an objective function value of 97,500€ is proposed as MP-LP solution, whereas total labor costs of 110,500€ is proposed when forcing integer values (MP-IP).

For the remaining two demand patterns (75% and 100% quantile), a total salary for physicians of 117,000€ respectively 123,500€ are required. Both values are provided in quite an acceptable amount of time for a roster covering an entire month, even though computation time is higher for the maximum demand level, i.e. 5 h compared to 6 h respectively.

Due to the complexity derived by the three different types of implemented flexibility, the compact formulation of our mathematical program is not solvable for the majority of parameter settings, a planning horizon of more than 1 week, and other demand pattern than mean level. Since our column generation approach performs quite satisfying with respect to solution quality and computational effort, it is further in use for the following steps of our experimental study.

5.3 Factorial analysis of flexibility parameters

In this subsection, we conduct a factorial analysis to test for the effect of the different types and degrees of flexibility in the staffing process on the total salary costs of the workforce. In our analysis, we subdivide our flexibility parameters into two discrete subsets according to the effect we focus on: day parameters and shift parameters.

In the following, each of these values is varied successively while keeping the remaining parameters constant. For the minimum number of consecutive working days, the minimum number of consecutive days off as well as the minimum shift length, the value of the specific parameter is increased iteratively. For the maximum number of consecutive days on duty and the maximum shift length, we vary the appropriate value in the opposite direction. This results in a total of 168 test instances whose coherence is stated in Table 4. Note, some of the tested parameter-combinations might not be applicable in real life to build a working roster, but for completeness of our factorial design, these are also conducted.

We solve each variation of the stated parameters applying our CG heuristic whereas the former stated mean demand scenario serves as input for the underlying requirement for care having a 2-week planning horizon.

5.3.1 Variation of day parameters

For each parameter setting of D^{on} and \bar{D}^{on} , the minimum number of days off is additionally increased iteratively from one to five to investigate the effect additional flexibility in terms of daily working patterns on total workforce size. In each step, the maximum number of consecutive days on duty is decreased from the initial value of six to the current value of \underline{D}^{on} , until $\bar{D}^{on} = \underline{D}^{on}$. This results in a total of 105 different settings. A detailed overview on the objective function values, absolute optimality gap, number of columns, and computational effort is provided in Table 5 in the ‘‘Appendix’’. Again, \underline{D}^{on} is increased iteratively whereas the parameter value of \bar{D}^{on} is decreased.

In general, solution times vary depending on the parameter setting in use, but the majority is rather low, i.e. ranging from few seconds to more than 10 h. Around 70% of the considered test sets cannot be solved within the defined time limit when solving MP-IP. In these cases, the solution process is aborted after a computation time of 600 s. The remaining 28 instances are solvable in < 10 min. It is therefore possible, to find a good feasible solution with an appropriate computational effort for solving a 4 week planning problem. This again serves as evidence for the satisfying

Table 4 Factorial design

Parameter	D^{on}	\bar{D}^{on}	\underline{D}^{off}	S^{min}	S^{max}	B^{pre}	B^{post}	168 test instances
Standard parameter setting	3	6	2	7	12	3	3	
Variation in day parameters	1–6	6–1	1–5	7	12	3	3	
Variation in shift parameters	3	6	2	7–12	12–7	3–1	3–1	

performance of our provided solution approach in contrast to the compact MIP formulation.

Concerning the objective function value, the minimum labor costs for physicians range from 110,500€ to 396,500€ at its maximum. The resulting salary costs vary depending on the provided level of flexibility in patterns of working days: Full flexibility results in an objective value of 110,500€ whereas a parameter setting with a lower level of flexibility provides 396,500€ as solution, i.e. for $\bar{D}^{on} = \underline{D}^{on} = 1$ and $\underline{D}^{off} = 5$. A more realistic parameter setting, defining $\underline{D}^{off} = 2$, results in personnel costs ranging from 110,500€ to 221,000€ when varying the remaining day parameter \bar{D}^{on} and \underline{D}^{on} . Due to the reduction in flexibility as the minimum and the maximum number of allowed consecutive working days converge, the available combination of assignable days on and off duty decrease. As a result, labor costs increase significantly. Even though, the spread of the minimum and the maximum total wage costs is not that significant for each set of test instances, similar results can be ascertained for the remaining parameter combinations, e.g. $\bar{D}^{on} = 6, \underline{D}^{on} = 2$ and $\underline{D}^{off} = 1$ in comparison to $\bar{D}^{on} = 2, \underline{D}^{on} = 2$ and $\underline{D}^{off} = 1$. As a result, if the appropriate workforce size derived by the personnel costs provided by the CG heuristic for a specific parameter setting is employed, the average weekly utilization of personnel ranges from 20.07% to 72.03% at its maximum (see column ten in Table 5 in the ‘‘Appendix’’). Utilization levels are derived by the ratio of available hours of supply and hours of occurring demand (in each week), given in (5.1):

$$\text{Weekly utilization} = \frac{\sum_{d \in D} \sum_{p \in P} N_{dp}}{\bar{R} \cdot \sum_{r \in R} \lambda_r} \quad (5.1)$$

Since we use our 50% quantile demand scenario as input data, the requirement for care is 551 h each week. These are confronted with the working hours of the employed personnel. As an example, weekly utilization for $\bar{D}^{on} = 6, \underline{D}^{on} = 1$ and $\underline{D}^{off} = 1$ is calculated: The labor costs for a 4 week period provided by our CG heuristic is 110,500€ which leads to a required number of 17 physicians. This results in a total of 765 h of available working hours each week. Proportioning these two measures devotes an average utilization of staff each week of $\frac{551}{765} = 0.7203$, i.e. 72.03%.

Contrary to the increasing labor costs when decreasing the level of flexibility, the average utilization of staff each week drops the lower the degree of flexibility. This is reasoned by the growing number of required personnel and with this, an increasing number of available working hours weekly (in case of less flexibility) to cover demand. But, opposed to the increasing available working hours, demand for care is constant in each test instance. Therefore, average utilization decreases. Although utilization fluctuates, the achieved levels in settings providing an appropriate level of flexibility seem to be quite reasonable especially in a hospital care providing context. Since surgeries commonly have a duration of several hours and require a great amount of concentration, utilization should not be too high to create hours of idle time (around the scheduling of surgeries) which can be used to handle administrative tasks as well as emergency patients. Note, weekly utilization of staff is strongly

influenced by the predefined underlying scheduling rules: The higher the degree of flexibility, the higher the utilization of physicians due to the increasing capability to better match supply and demand. The more rigid the rostering regulation, the less options for the assignment of days on duty and different shifts occur.

With respect to the quality of the provided solutions, our column generation approach seems to work quite well: Optimality gap is 6190€ on average, i.e. approximately one physician. The absolute optimality gap is at most 13,000€ for all 105 test instances. Moreover, 40% of the considered test instances provides an absolute optimality gap of around 6500€, almost one third of our test sets have an absolute optimality gap of 13,000€, and more than 30% are solved to optimality by the CG heuristic, i.e. optimality gap equals 0€.

An exemplary depiction of the development of the resulting objective function values for MP–LP as well as the solution of MP–IP for a minimum number of 1 day on duty, i.e. $\bar{D}^{on} = 1$, is shown in Fig. 8.

The vertical axis shows the minimum number of days off (duty) ranging from one to 5 for each value of $\bar{D}^{on} = 1, \dots, 6$. As the figure displays, the objective function value for MP–LP as well as for MP–IP increases for each value of \bar{D}^{on} when reducing flexibility by increasing \bar{D}^{off} from one to five. Moreover, there is a trend discernible: Despite the increasing objective function value for each manifestation of \bar{D}^{on} itself, the total labor costs of physicians grow linearly in the reduction of flexibility in the maximum number of days on duty. As a result, the lower the degree in flexibility in \bar{D}^{on} , the higher the total salary expenses.

Once employees have to have at least some consecutive days off after being on duty, the size of the workforce and with this the total salary costs increase dramatically. Generally speaking, the higher the convergence between the minimum and the maximum number of allowed consecutive working days, the lower the computational effort due to the decreasing number of feasible assignments. Additionally, the total number of required physicians and salary costs increase. Less flexibility in patterns of working days results in a significant increase in the value of the objective. But, even though a higher number of physicians is needed to cover demand for less flexible working patterns, especially the number of consecutive days off might

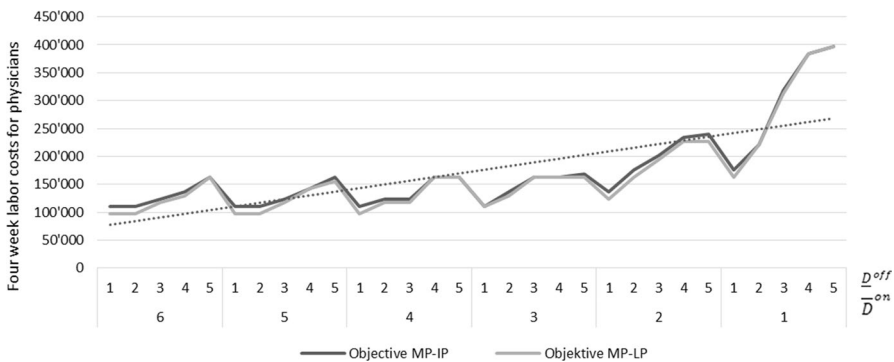


Fig. 8 Trend in in day parameter flexibility (for $\bar{D}^{on} = 1$)

be of major importance for the satisfaction of the workforce. Having only one single day off between two consecutive sequences of working days leads to discontent and a higher probability for employee's turnover. Therefore, it is sufficient to find the balance between the desired level of flexibility and an acceptable burden for staff to ensure their wellbeing.

5.3.2 Variation of shift parameters

In this section, the effect of flexibility in the allowed length of shifts is decreased and analyzed while the flexibility in starting times of shifts remains. Similar to the former subsection, the values of the different shift parameters (concerning its length) are varied iteratively. In the first step, S^{min} is increased from 7 to 12 h whereas the value of S^{max} is decreased from an initial maximum length of twelve to the current value of S^{min} . Due to the effect of the shift length on the placement of the break and with this, coverage of demand, the size of the break window is additionally considered. Therefore, B^{pre} and B^{post} are varied simultaneously, i.e. $B^{pre} = B^{post} = 3$ in the initial step. Subsequently, the values are decreased iteratively to enlarge the size of the break window. Increasing the number of possibilities for the placement of the break provides additional flexibility in the scheduling process. This results in a total of 63 different combinations for the setting of the shift parameters (see Table 6 in the "Appendix" for a detailed overview on the solution and evaluation values, computational effort, and the average weekly utilization as performance indicator).

Similar to the variation in day parameters, the computational effort to solve the different shift parameter settings varies significantly based on the considered parameter setting but is still appropriate, i.e. ranging from < 1 min to 17 h. Moreover, solving MP-IP in < 10 min is only achievable for < 15% of the test instances. For the remaining ones, the predetermined time limit of 600 s is effective. Generally, the higher the convergence between minimum and maximum allowed shift length, the smaller the solution time due to the reduced number of available shift types. For a maximum level of flexibility, the computational effort is still rather low which is reasonable. Increasing the flexibility for the assignment of the break by enlarging the size of the break window does not lead to a significant growth in solution times.

Moreover, there is a significant effect of the degree of shift flexibility on workforce costs. A larger number of shift types results in lower expenses for physicians. For the maximum level of flexibility with $S^{min} = 7$ and $S^{max} = 12$, solely 110,500€ labor costs are required to employ enough personnel to cover demand. Since the number of allowed shift types is reduced due to the convergence of S^{min} and S^{max} , total salary costs increase to a maximum of 169,000€. In this extreme, the fluctuation in demand can only be encountered by assigning one single type of shift having various starting periods, e.g. with a duration of twelve periods. Considering the case that a hospital does solely consider flexible starting periods for shifts rather than flexible shift lengths to schedule their personnel for whatever reason, the results provided by the table indicate that shorter shifts might be more suitable to cover standard demand compared to shifts with a long duration. In general, implementing a higher amount of flexibility in the length of available shifts results in a smaller objective function value and with this a decreasing number of physicians required.

Concerning the average weekly utilization of staff, a minimum value of 48.98% and a maximum of 72.03% is provided which in turn seems to be quite appropriate for hospital settings.

Inserting additional flexibility in the scheduling process by enlarging the size of the break window does either have a positive or no effect on the objective function value. For almost 20% of the different parameter settings, the maximum size of the break window decreases the total labor costs. Especially for long shift durations providing a minimum amount of flexibility, e.g. $S^{min} = 11$ and $S^{max} = 12$, enlarging the feasible combinations for the placement of the break does not affect the resulting objective. As a result, flexibility in break assignment particularly affects settings that do solely provide shorter shifts (< 11 h) with a low level of flexibility, i.e. especially for real life assumptions. For instances, allowing a high degree of flexibility in the number of available shift types, additional flexibility by enlarging the size of the break window does not have a significant effect. In contrast to the factorial analysis of day parameters, the solution quality of the shift parameter analysis is not of similar performance, but still appropriate. On average, optimality gap is 12,484€, i.e. corresponding to an additional employment of at most two physicians. This disparity is rather small and serves additionally as proof for the high performance of our solution approach. Around 15% of the considered test sets provide a total optimality gap of 19,500€, 13,000€ additional salary costs are required in more than 65% and 6500€ higher wage costs are required in 13% of the test instances. Actually, three parameter settings are solved to optimality, i.e. with a MP–LP solution equal to the MP–IP solution. An example for the development of the objective function values of the CG heuristic and the MP–LP in case of $S^{min} = 7$ is provided in Fig. 9.

Again, the total labor costs and with this, the number of required personnel either stays constant or decreases due to the implementation of additional flexibility by enlarging the break window. But this effect is rather small. But, similar to the variation in day parameters, the higher the convergence between S^{min} and S^{max} , the higher the required number of physicians to cover demand.

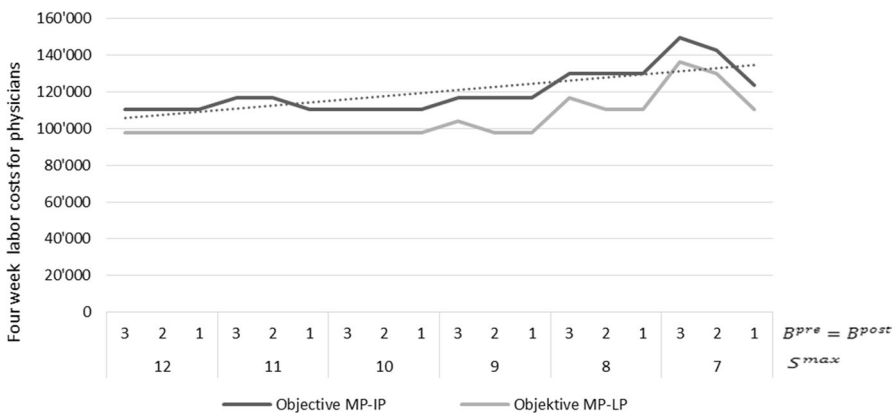


Fig. 9 Trend in in shift parameter flexibility (for $S^{min} = 7$)

Generally, in the scheduling process, the considered amount of flexibility in terms of shift types and the size of the break window provide a significant improvement concerning the total labor costs and therefore the overall number of physicians that is required to cover mean demand. Even though the computational effort rises due to the increasing level of assignment opportunities, the generated benefit is valuable for hospitals to reduce their expenses for personnel while ensuring the requested level of service.

5.4 Real life comparison and deduced managerial insights

As further studies already prove, implementing flexibility in terms of flexible sequences of working days, flexible shifts, and a flexible assignment of the obligatory break leads to a significant reduction in total labor costs. Especially due to the tactical nature of our planning problem, the increase in computational effort is outweighed by the positive effect of flexibility on personnel costs.

Therefore, this section comprises of two parts: In the first part, flexibility in sequences of working days as well as available shift types and the placement of the break is decreased to account for real life surroundings. In the second part, managerial insights for the considered tactical planning problem derived from the results of our experimental study are discussed.

In contrast to our factorial analysis, when accounting for real life scheduling circumstances, we first decrease the level of flexibility: Hospitals commonly do not account for flexible sequences of working days. Instead, a standard 5 days working week is considered. In case of the large teaching hospital in Germany which we were focusing on, physicians are assigned to a regular 5-day working week from Monday to Friday. Demand in off periods (e.g. throughout the night) is covered by on-call service whereas requirement for care on weekend days is handled by additionally cyclic assigned weekend duties for the workforce. Moreover, this hospital does not provide any flexibility in shift types as there is only one single type of shift available, starting at 7:30 am with a duration of 9 or 10 h, to be assigned to personnel. Breaks are only planned implicitly rather than being assigned individually for each physician on duty. To account for this rigid scheduling environment, we first of all decrease the level of flexibility in terms of sequences of working days to a 5-day working week, i.e. physicians are assigned to five consecutive working days. Afterwards, there are 2 days off assigned.

Since other departments in hospitals as well as relevant literature apply a three 8-h shift approach as common practice, we assume three different 8-h shift systems to compare this rigid planning environment with our flexible scheduling approach: First, we assume twelve different 8-h shifts with flexible starting periods, i.e. starting in each period of the (planning) day. Second, three equally overlapping 8-h shifts which are majorly in use in current literature and practice are implemented. Third, a combined shift system with one 8-h shift starting in period four (corresponding to regular operating room opening hours) and two overlapping 10-h shifts which start in period one and period nine respectively is considered. As input data, we assume a 4 week planning period, standard parameter setting, and mean demand pattern.

The first shift system having twelve 8-h shifts available results in total labor costs of 123,500€. This means, reducing flexibility in shift length to a minimum level leads to an increase in personnel costs by 13,000€ compared to the labor budget required by our standard parameter setting, i.e. 110,500€. These rather low total salary costs derive from two main drivers: As already pointed out in our factorial analysis, if there is no flexibility in the length of shifts, short shifts are advantageous compared to longer shifts with respect to the resulting personnel costs. Moreover, since we still allow for flexibility in terms of starting periods, the considered 8-h shifts can start in each period of the day, i.e. from period one to period 12, which results in twelve different 8-h shifts. This means, this shift system allows for a degree of flexibility which is still appropriate to improve the ability to match supply and demand.

In contrast, assuming solely three 8-h shifts leads to total salary costs of at least 195,000€ for the considered planning horizon. This means, compared to our standard parameter setting, the minimum total labor costs decrease by 84,500€ per month if flexibility is implemented in the scheduling process, i.e. a reduction of 43% in total costs if flexibility is integrated in the planning procedure. In contrast to a three 8-h shift system, the combination of an 8-h and two 10-h shifts decreases personnel costs to 165,000€. But, compared to our standard parameter setting, this still represents an increase in total labor costs by 33%, i.e. from 110,500€ to 165,000€.

Based on the previously discussed results of the conducted experimental study, several key insights supporting managerial decision making concerning staffing and scheduling policies and labor budget can be extrapolated: The major insight derived by our study is the effect of flexibility in the scheduling process. When facing labor staffing and scheduling decisions, hospital's management is confronted with a trade-off arbitration between a cost- and an employee-perspective. This means, from a cost-perspective, the more flexibility is implemented in the scheduling process, the lower the resulting total labor costs. In contrast, from a personnel-perspective, the higher the degree of flexibility, the lower the overall satisfaction level of labor due to the unpredictability of working times. As a result, job motivation and the quality of care might suffer. It is therefore important for hospital's management to investigate an appropriate level of flexibility to ensure personnel satisfaction on the one hand and reduce total labor cost on the other hand.

An additional benefit provided by flexible scheduling assumptions is an increase in utilization. The larger number of available shift types with various starting periods and different lengths leads to a better match of supply and demand and therefore, to a decrease in unnecessary overtime hours for personnel. A standard three 8-h shift system, for example, leads to a utilization of around 40% which is rather low. As a result, the more flexible the considered scheduling environment, the higher labor utilization which leads to an efficient assignment of personnel.

It is therefore necessary for hospital's management to define an appropriate level of flexibility to reduce total labor costs while maintaining job satisfaction and personnel motivation. Moreover, as the total salary costs determined by a specific level of flexibility determines the number of required physicians implicitly, this number represents the adequate number of required fulltimers. This information is of central relevance since hospital's management can also decide which kind of contract types to employ, e.g. one fulltime physician, two part time employees having a 50%

workload, or appropriate contract types., which might additionally improve matching supply and demand.

6 Summary and conclusion

In this research, we formulated a MIP based on a reduced set covering approach to consider the implementation of flexible sequences of working days, flexible shifts, and flexible break assignments when scheduling physicians in hospitals. Due to the complexity of the stated model, a column generation heuristic is proposed to provide solutions of high quality in reasonable amount of solution time. To evaluate the performance of the solution approach as well as the effect of implementing a high degree of flexibility in the scheduling process (i.e. in terms of total salary costs as monetary value for the required workforce size), real life data from a large teaching hospital in Germany of the year 2010 is used. In our experimental study, demand of each period of every day over one entire year is aggregated and three different demand scenarios are defined: 50% quantile, 75% quantile, and maximum (100% quantile) demand. Moreover, the performance of the proposed CG heuristic is evaluated by varying the length of the planning horizon from 1 to 6 weeks. Finally, to test for the effect of additional flexibility in the scheduling process, a factorial analysis varying day, shift, and break parameters is conducted.

In general, our CG approach performs quite well: Due to the complexity of our mathematical model, it is not possible to solve the MIP to optimality for at most all test instances. Albeit only a heuristic solution is provided by the CG approach, each of our test sets is solved in an appropriate amount of time. Additionally, heuristic solutions of high quality are provided which is proven by the optimality gap between MP–LP and MP–IP solution. Furthermore, as an outcome of our sensitivity analysis, the lower the degree of flexibility in terms of sequences of working days and shift types, the higher the upcoming labor costs due to the increasing number of required physicians to cover demand. Focusing on the variation of shift parameters, an increasing level of flexibility leads to a reduction in salary costs. Besides, additional reduction is achievable by enlarging the size of the break window. But, this effect is only of reduced significance compared to the level of flexibility in day and shift parameters. As a result, the optimal scheduling strategy depends on the policy of hospitals management and the hospital's ambition in terms of quality of care, desired service level, and employee satisfaction. Especially in case of employing a workforce size below the maximum level, it is barely impossible to stick to a predefined inflexible 8-h shift system to prevent for huge wage expenses. Due to the reduced number of physicians, flexible shifts have to be implemented to create an appropriate flexible shift system adopting the fluctuation in demand.

For future research, a sophisticated relax-and-fix approach may be considered as a decomposition method to tackle the formulation's time-horizon problem. By doing so, the complexity of the problem might be reduced significantly, which leads to a high quality integer solution in short computation time. Moreover, a B&P approach can be implemented to generate optimal integer solutions for the considered problem. Concerning the underlying demand pattern, further research might

be investigated to build a more realistic estimation of demand profiles. Up to now, research commonly uses randomly generated demand patterns based on predefined assumptions or real life data derived from a hospital. But, even if real life demand patterns are in use, these may be biased due to specific regulations and other schedules in hospital: In general, especially when scheduling elective patients, available operating room allocations based on specialty defined by the Master-Surgery-Schedule (MSS) and the number of assignable available shift types determine the schedule of surgeries, i.e. demand. Therefore, demand is partially driven by the shift design of a hospital and might change, if number and type of available shift type changes.

Despite the large potential to decrease hospitals' operating and labor costs, on an operational level, additional aspects concerning the wellbeing of employees should be taken into account to prevent from turnover of physicians. Moreover, a time window constraint regulating the starting times of consecutive shifts can be implemented to ensure more balanced starting periods for each physician within one sequence of working days. In addition to this, on an operational level, a fair distribution of working hours has to be ensured. Since the number of consecutive days on duty and with this, the number of days on duty varies within the planning horizon from physician to physician, the workload (measured in working hours) of some physicians might be higher than for others. As a result, a minimum number of working periods for each physician as mechanism to ensure an equal distribution of workload is to be implemented. It is also important to integrate the patients' wellbeing and therefore the quality of the provided service. Implementing a lot of flexibility especially in terms of sequences of working day might result in its extreme to a roster where physicians are 1 or 2 days on duty and are off for the subsequent day(s). Therefore, an appropriate level of flexibility is to be determined for the trade of between monetary value of flexibility and quality of care. As a result, it might make sense to provide a behavioral template/scheme for hospital's management how to operate depending on the institution's goal.

Appendix I

Linearization of constraint 7.1.

$$z_{id} = y_{id}(1 - y_{i(d-1)}) \quad \forall i \in I, d \in D \quad (7.1)$$

$$z_{id} - y_{id} \leq 0 \quad \forall i \in I, d \in D \quad (7.2)$$

$$z_{id} + y_{i(d-1)} \leq 1 \quad \forall i \in I, d \in D \quad (7.3)$$

$$y_{id} - y_{i(d-1)} - z_{id} \leq 0 \quad \forall i \in I, d \in D \quad (7.4)$$

Appendix II

See Table 5.

Table 5 Factorial analysis—flexibility in day parameter

D^{on}	D^{off}	\bar{D}^{on}	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
1	1	6	97,500	110,500	13,000	2715	313.05	600	72.03
	2		97,500	110,500	13,000	2831	355.55	600	72.03
	3		117,000	123,500	6500	680	36.02	600	64.44
	4		130,000	136,500	6500	600	30.77	600	58.31
	5		162,500	162,500	0	71	2.60	357	48.98
	1	5	97,500	110,500	13,000	2889	824.78	600	72.03
	2		97,500	110,500	13,000	2817	817.40	600	72.03
	3		117,000	123,500	6500	1035	151.45	600	64.44
	4		143,000	143,000	0	585	70.05	600	55.66
	5		156,000	162,500	6500	78	6.68	413	48.98
	1	4	97,500	110,500	13,000	1629	166.83	600	72.03
	2		117,000	123,500	6500	716	36.47	600	64.44
	3		117,000	123,500	6500	1285	86.52	600	64.44
	4		162,500	162,500	0	271	11.37	600	48.98
	5		162,500	162,500	0	66	2.37	232	48.98

Table 5 (continued)

\bar{D}^m	\bar{D}^n	\bar{D}^{diff}	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
3	1	1	110,500	110,500	0	921	52.33	600	72.03
		2	130,000	136,500	6500	828	45.92	600	58.31
		3	162,500	162,500	0	258	11.17	600	48.98
		4	162,500	162,500	0	292	13.75	600	48.98
		5	162,500	169,000	6500	64	2.62	226	47.09
2	1	1	123,500	136,500	13,000	709	39.08	600	58.31
		2	162,500	175,500	13,000	404	19.30	600	45.35
		3	195,000	201,500	6500	206	0.04	45	39.50
		4	227,500	234,000	6500	102	0.02	0	34.01
		5	227,500	240,500	13,000	253	0.03	31	33.09
1	1	1	162,500	175,500	13,000	461	22.23	600	45.35
		2	221,000	221,000	0	203	0.43	38	36.01
		3	312,000	318,500	6500	174	0.28	37	24.99
		4	383,500	383,500	0	69	0.30	26	20.75
		5	396,500	396,500	0	214	0.30	19	20.07
2	6	1	97,500	110,500	13,000	2787	761.37	600	72.03
		2	97,500	110,500	13,000	2817	828.22	600	72.03
		3	117,000	123,500	6500	656	81.22	600	64.44
		4	130,000	136,500	6500	578	69.17	600	58.31
		5	162,500	162,500	0	60	5.25	96	48.98

Table 5 (continued)

\bar{D}^m	\bar{D}^n	\bar{D}^{diff}	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
5	1	1	97,500	110,500	13,000	2945	873.30	600	72.03
			97,500	110,500	13,000	2838	841.23	600	72.03
			117,000	123,500	6500	1093	167.15	600	64.44
			143,000	143,000	0	551	66.93	600	55.66
			162,500	162,500	0	193	18.73	81	48.98
4	1	1	97,500	110,500	13,000	1618	122.00	600	72.03
			117,000	123,500	6500	676	32.00	600	64.44
			117,000	123,500	6500	1266	82.40	600	64.44
			162,500	162,500	0	261	9.87	600	48.98
			162,500	162,500	0	70	2.37	35	48.98
3	1	1	110,500	117,000	6500	985	123.23	600	68.02
			130,000	136,500	6500	837	75.20	600	58.31
			162,500	162,500	0	289	20.82	600	48.98
			162,500	162,500	0	316	27.20	600	48.98
			162,500	169,000	6500	67	6.07	228	47.09
2	1	2	123,500	130,000	6500	741	34.45	600	61.22
			162,500	169,000	6500	407	15.93	600	47.09
			227,500	227,500	0	56	1.73	187	34.98
			227,500	227,500	0	102	2.15	31	34.98
			227,500	234,000	6500	121	2.75	32	34.01

Table 5 (continued)

\bar{D}^m	\bar{D}^n	\bar{D}^{diff}	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
3	6	1	97,500	110,500	13,000	3029	417.18	600	72.03
		2	97,500	110,500	13,000	2950	318.00	600	72.03
		3	117,000	123,500	6500	646	36.55	600	64.44
		4	130,000	130,000	0	583	32.57	600	61.22
		5	162,500	162,500	0	56	2.20	26	48.98
	5	1	97,500	110,500	13,000	2664	791.28	600	72.03
		2	97,500	110,500	13,000	2957	414.52	600	72.03
		3	117,000	123,500	6500	1080	71.25	600	64.44
		4	143,000	143,000	0	640	31.52	600	55.66
		5	162,500	162,500	0	221	9.32	162	48.98
	4	1	104,000	110,500	6500	1429	104.27	600	72.03
		2	117,000	123,500	6500	734	38.35	600	64.44
		3	117,000	123,500	6500	1294	89.02	600	64.44
		4	162,500	162,500	0	267	11.30	600	48.98
		5	162,500	162,500	0	161	6.35	401	48.98
	3	1	110,500	123,500	13,000	1151	75.65	600	64.44
		2	130,000	136,500	6500	1027	63.68	600	58.31
		3	162,500	162,500	0	296	12.28	600	48.98
		4	162,500	162,500	0	298	12.47	600	48.98
		5	162,500	169,000	6500	66	2.37	215	47.09

Table 5 (continued)

\bar{D}^m	\bar{D}^n	\bar{D}^{diff}	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
4	6	1	97,500	110,500	13,000	2775	354.82	600	72.03
		2	97,500	110,500	13,000	2660	333.97	600	72.03
		3	117,000	123,500	6500	639	31.68	600	64.44
		4	130,000	136,500	6500	576	28.13	600	58.31
		5	162,500	162,500	0	61	2.12	375	48.98
	5	1	97,500	110,500	13,000	2773	361.68	600	72.03
		2	97,500	110,500	13,000	2847	372.30	600	72.03
		3	117,000	123,500	6,500	1010	60.88	600	64.44
		4	143,000	143,000	0	569	27.82	600	55.66
		5	162,500	162,500	0	73	2.50	265	48.98
	4	1	104,000	117,000	13,000	1172	79.40	600	68.02
		2	117,000	123,500	6500	794	42.68	600	64.44
		3	117,000	123,500	6500	1306	95.25	600	64.44
		4	162,500	162,500	0	333	15.17	600	48.98
		5	162,500	162,500	0	75	2.82	65	48.98

Table 5 (continued)

\bar{D}^m	\bar{D}^n	\bar{D}^{diff}	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
5	6	1	97,500	110,500	13,000	2728	368.10	600	72.03
		2	97,500	110,500	13,000	2902	391.37	600	72.03
		3	117,000	123,500	6500	762	41.27	600	64.44
		4	130,000	136,500	6500	683	35.77	600	58.31
		5	162,500	162,500	0	83	2.92	69	48.98
5	6	1	97,500	110,500	13,000	2583	325.78	600	72.03
		2	97,500	110,500	13,000	2787	354.57	600	72.03
		3	123,500	130,000	6500	757	1'360.00	600	61.22
		4	143,000	149,500	6500	572	1'633.00	600	53.24
		5	162,500	162,500	0	429	18.85	172	48.98
6	6	1	104,000	117,000	13,000	1723	165.17	600	68.02
		2	110,500	117,000	6500	1111	71.37	600	68.02
		3	123,500	130,000	6500	920	54.12	600	61.22
		4	136,500	143,000	6500	808	43.93	600	55.66
		5	136,500	143,000	6500	209	7.67	87	55.66

Appendix III

See Table 6.

Table 6 Factorial analysis—flexibility in shift parameter

S^{min}	S^{max}	$B^{pre} = B^{post}$	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
7	12	3	97,500	110,500	13,000	2950	318.00	600	72.03
		2	97,500	110,500	13,000	2510	312.75	600	72.03
		1	97,500	110,500	13,000	1583	149.03	600	72.03
	11	3	97,500	117,000	19,500	2960	784.98	600	68.02
		2	97,500	117,000	19,500	3158	935.88	600	68.02
		1	97,500	110,500	13,000	2122	219.00	600	72.03
	10	3	97,500	110,500	13,000	2897	705.35	600	72.03
		2	97,500	110,500	13,000	3662	1073.93	600	72.03
		1	97,500	110,500	13,000	2895	331.75	600	72.03
	9	3	104,000	117,000	13,000	1274	134.32	600	68.02
		2	97,500	117,000	19,500	2161	371.57	600	68.02
		1	97,500	117,000	19,500	2890	673.17	600	68.02
	8	3	117,000	130,000	13,000	1227	127.62	600	61.22
		2	110,500	130,000	19,500	1799	265.98	600	61.22
		1	110,500	130,000	19,500	2297	414.42	600	61.22
	7	3	136,500	149,500	13,000	135	0.76	215	53.24
		2	130,000	143,000	13,000	295	0.86	600	55.66
		1	110,500	123,500	13,000	683	1.62	432	64.44

Table 6 (continued)

S^{min}	S^{max}	$B^{pre} = B^{best}$	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
8	12	3	97,500	110,500	13,000	2883	364.06	600	72.03
		2	97,500	110,500	13,000	2090	214.87	600	72.03
		1	97,500	110,500	13,000	1107	145.32	600	72.03
	11	3	97,500	110,500	13,000	2859	327.27	231	72.03
		2	97,500	110,500	13,000	3098	394.42	54	72.03
		1	97,500	110,500	13,000	1559	241.32	312	72.03
	10	3	97,500	110,500	13,000	2327	196.32	600	72.03
		2	97,500	110,500	13,000	3367	403.05	600	72.03
		1	97,500	110,500	13,000	2321	481.80	600	72.03
	9	3	117,000	130,000	13,000	1012	68.83	600	61.22
		2	97,500	117,000	19,500	1827	112.15	600	68.02
		1	97,500	110,500	13,000	2310	183.88	289	72.03
	8	3	117,000	130,000	13,000	1051	34.65	600	61.22
		2	110,500	123,500	13,000	1647	89.02	600	64.44
		1	110,500	123,500	13,000	1635	91.13	600	64.44

Table 6 (continued)

S^{min}	S^{max}	B^{pre}	B^{post}	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utilization (average) (%)
9	12	3	3	97,500	110,500	13,000	2271	293.68	600	72.03
		2	2	97,500	110,500	13,000	2649	367.23	600	72.03
		1	1	91,000	110,500	19,500	1838	177.75	600	72.03
	11	3	3	97,500	117,000	19,500	1505	104.52	600	68.02
		2	2	97,500	110,500	13,000	2707	308.35	600	72.03
		1	1	97,500	110,500	13,000	3611	98.18	600	72.03
	10	3	3	104,000	117,000	13,000	1210	53.63	600	68.02
		2	2	97,500	110,500	13,000	1631	93.48	600	72.03
		1	1	97,500	110,500	13,000	2256	187.67	600	72.03
	9	3	3	104,000	117,000	13,000	940	28.40	600	68.02
		2	2	97,500	110,500	13,000	1267	51.37	600	72.03
		1	1	97,500	110,500	13,000	1471	150.35	600	72.03
10	12	3	3	117,000	130,000	13,000	731	22.95	600	61.22
		2	2	117,000	123,500	6500	690	21.35	600	64.44
		1	1	117,000	123,500	6500	636	19.07	600	64.44
	11	3	3	123,500	130,000	6500	442	8.93	600	61.22
		2	2	123,500	130,000	6500	384	7.25	600	61.22
		1	1	123,500	123,500	0	396	7.77	600	64.44
	10	3	3	130,000	136,500	6500	478	8.35	600	58.31
		2	2	130,000	136,500	6500	497	8.93	600	58.31
		1	1	130,000	130,000	0	486	8.80	600	61.22

Table 6 (continued)

S^{min}	S^{max}	$B^{pre} = B^{best}$	Objective MP-LP	Objective MP-IP	Opt. gap (absolut)	# columns	Solution time: MP-LP (min)	Solution time: MP-IP (s)	Weekly utiliza- tion (average) (%)
11	12	3	117,000	130,000	13,000	589	14.52	600	61.22
		2	117,000	130,000	13,000	485	10.58	600	61.22
		1	117,000	130,000	13,000	517	11.63	600	61.22
	11	3	123,500	136,500	13,000	306	3.78	600	58.31
		2	123,500	136,500	13,000	240	2.53	600	58.31
		1	123,500	136,500	13,000	259	2.93	600	58.31
12	12	3	162,500	169,000	6500	145	1.20	97	47.09
		2	162,500	169,000	6500	132	1.02	124	47.09
		1	162,500	162,500	0	128	0.93	19	48.98

References

- Aggarwal SC (1982) A focused review of scheduling in services. *Eur J Oper Res* 9(2):114–121
- Aykin T (1996) Optimal shift scheduling with multiple break windows. *Manag Sci* 42(4):591–602
- Barnhart C, Johnson EL, Nemhauser GL, Martin WP, Savelsbergh VPH (1998) Branch-and-price: column generation for solving huge integer programs. *Oper Res* 46(3):316–329
- Bechtold SE, Jacobs LW (1990) Implicit modeling of flexible break assignments in optimal shift scheduling. *Manag Sci* 36(11):1339–1351
- Beliën J, Demeulemeester E (2006) Scheduling trainees at a hospital department using a branch-and-price approach. *Eur J Oper Res* 175(1):258–278
- Blum K, Löffert S, Offermanns M, Perner P (2007) Krankenhaus Barometer—Umfrage 2017. Technical report, Deutsches Krankenhaus Institut e.V., Düsseldorf
- Blum K, Löffert S, Offermanns M, Perner P (2017) Krankenhaus Barometer—Umfrage 2017. Technical report, Deutsches Krankenhaus Institut e.V., Düsseldorf
- Böhl U (2014) Statistische Krankenhausdaten: Grund- und Kostendaten der Krankenhäuser 2011. In: Klauber J, Geraedts M, Friedrich J, Wasem J (eds) *Krankenhaus-Report 2014: Schwerpunkt: Patientensicherheit*. Schattauer, Stuttgart, pp 293–328
- Brunner JO, Edenharter GM (2011) Long term staff scheduling of physicians with different experience levels in hospitals using column generation. *Health Care Manag Sci* 14(2):189–202
- Brunner JO, Bard JF, Kolisch R (2009) Flexible shift scheduling of physicians. *Health Care Manag Sci* 12(3):285–305
- Brunner JO, Bard JF, Kolisch R (2010) Midterm scheduling of physicians with flexible shifts using branch and price. *IIE Trans* 43(2):84–109
- Cohn A, Root S, Kymissis C, Esses J, Westmoreland N (2009) Scheduling medical residents at boston university school of medicine. *Interfaces* 39(3):186–195
- Damci-Kurt P, Zhang M, Marentay B (2017) Improving physician schedules by leveraging equalization: cases from hospitals in US. Working Paper, Kennesaw State University, Kennesaw, Georgia
- Dantzig GB (1954) Letter to the editor—a comment on Edie’s “Traffic delays at toll booths”. *J Oper Res Soc Am* 2(3):339–341
- Dantzig GB, Wolfe P (1960) Decomposition principle for linear programs. *Oper Res* 8(1):101–111
- Desaulniers G, Desrosiers J, Solomon MM (eds) (2005) *Column generation*. Springer, New York
- Desrosiers J, Lübbecke ME (2005) A primer in column generation. In: Desaulniers G (ed) *Column generation, GERAD 25th anniversary series, vol 5*. Springer, New York, pp 1–32
- Desrosiers J, Lübbecke ME (2011) Branch-price-and-cut algorithms. *Encyclopedia of operations research and management science*. Wiley, Chichester, pp 109–131
- El-Rifai O, Garaix T, Augusto V, Xie X (2015) A stochastic optimization model for shift scheduling in emergency departments. *Health Care Manag Sci* 18(3):289–302
- Erhard M, Brunner JO (2017) Physician staffing under special consideration of breaks—a case study on anesthetists. Working Paper, University of Augsburg, Augsburg, Germany
- Erhard M, Schoenfelder J, Fügner A, Brunner JO (2018) State of the art in physician scheduling. *Eur J Oper Res* 265(1):1–18
- Ernst AT, Jiang H, Krishnamoorthy M, Sier D (2004) Staff scheduling and rostering: a review of applications, methods and models. *Eur J Oper Res* 153(1):3–27
- Ganguly S, Lawrence S, Prather M (2014) Emergency department staff planning to improve patient care and reduce costs. *Decis Sci* 45(1):115–145
- Gilmore PC, Gomory RE (1963) A linear programming approach to the cutting stock problem—part II. *Oper Res* 11(6):863–888
- Isken MW (2004) An implicit tour scheduling model with applications in Health-care. *Ann Oper Res* 128(1):91–109
- Janaro RE, Bechtold SE (1985) A study of the reduction of fatigue impact on productivity through optimal rest break scheduling. *Hum Factors* 27(4):459–466
- Lübbecke ME, Desrosiers J (2005) Selected topics in column generation. *Oper Res* 53(6):1007–1023
- Marburger Bund (2016) Entgelttabelle zum TV-Ärzte/VKA. Marburger Bund, Berlin. <https://www.marburger-bund.de/sites/default/files/tarifvertraege/2017/tv-aerzte-vka/entgelttabelle-zum-tv-aerztevka.pdf>. Accessed 18 July 2018
- Rekik M, Cordeau JF, Soumis F (2010) Implicit shift scheduling with multiple breaks and work stretch duration restrictions. *J Sched* 13(1):49–75

- Statistisches Bundesamt (2016) Gesundheit—Kostennachweis der Krankenhäuser 2015. Fachserie 12 Reihe 6.3, Statistisches Bundesamt, Wiesbaden, 1–65
- Stolletz R, Brunner JO (2012) Fair optimization of fortnightly physician schedules with flexible shifts. *Eur J Oper Res* 219(3):622–629
- Thompson GM (1995) Improved implicit optimal modeling of the labor shift scheduling problem. *Manag Sci* 41(4):595–607
- Thungjaroenkul P, Cummings GG, Embleton A (2007) The impact of nurse staffing on hospital costs and patient length of stay: a systematic review. *Nurs Econ* 25(5):255–265
- Van den Bergh J, Beliën J, De Bruecker P, Demeulemeester E, De Boeck L (2013) Personnel scheduling: a literature review. *Eur J Oper Res* 226(3):367–385
- Vanderbeck F (2011) Branching in branch-and-price: a generic scheme. *Math Program* 130(2):249–294
- Volland J, Fügener A, Brunner JO (2017) A column generation approach for the integrated shift and task scheduling problem of logistics assistants in hospitals. *Eur J Oper Res* 260(1):316–334

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

Melanie Erhard was a researcher and PhD student at the chair for Health Care Operations/Health Information Management at the University of Augsburg, Germany. She is experienced in research in Operations Research/Management Science. She holds a PhD degree in economics as well as a diploma in information-oriented Business Administration studies. Her research focus is the planning of resources in hospitals, especially the physician scheduling problem with a specific focus on the investigation of breaks and a flexible planning and scheduling environment.