Generating insights on including flexibility in the planning processes of scarce resources in hospitals

Stefanie Ebel

Angaben zur Veröffentlichung / Publication details:

Ebel, Stefanie. 2025. "Generating insights on including flexibility in the planning processes of scarce resources in hospitals." Augsburg: Universität Augsburg.

Generating insights on including flexibility in the planning processes of scarce resources in hospitals

Kumulative Dissertation

der Wirtschaftswissenschaftlichen Fakultät
der Universität Augsburg
zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
(Dr. rer. pol.)

Vorgelegt von
Stefanie Susanne Ebel, M.Sc.

29.08.2024

Erstgutachter:	Prof. Dr. Jens O. Brunner
Zweitgutachter:	Prof. Dr. Sebastian Schiffels

Vorsitzender der mündlichen Prüfung:	Prof. Dr. Axel Tuma
Tag der mündlichen Prüfung:	07.01.2025

Adaptability is not imitation.
It means power of resistance and
assimilation.

- MAHATMA GANDHI

List of Contributions

This thesis contains the following contributions published in or intended to be submitted to scientific journals. The categories relate to the journal ranking VHB-Rating of the Verband der Hochschullehrerinnen und Hochschullehrer für Betriebswirtschaft e.V., 2024. The order of the contributions corresponds to the order of print in this thesis.

Contribution 1 Ebel SS, Brunner JO, Schimmelpfeng K (2024): Analyzing the value of flexibility in integrated surgery and staff scheduling on a daily planning level applying a column generation heuristic.

Status: to be submitted soon to INFORMS Manufacturing & Service Operations Management; Category A.

Contribution 2 Ebel SS, Grieger M, Erhard M, Brunner JO (2024): Analyzing staffing decisions subject to flexible scheduling and demand volatility in healthcare.

Status: to be submitted soon to Springer Journal of Scheduling; Category A.

Contribution 3 Bartenschlager CC, Ebel SS, Kling S, Vehreschild J, Zabel LT, Spinner CD, Schuler A, Heller AR, Borgmann S, Hoffmann R, Rieg S, Messmann H, Hower M, Brunner JO, Hanses F, Römmele C (2022): COVIDAL: A machine learning classifier for digital COVID-19 diagnosis in German hospitals.

Status: published in ACM Transactions on Management Information Systems; Category B.

The contribution is available online at https://doi.org/10.1145/3567431.

Table of Contents

1. Introduction

Limited resources in any industry require suitable consideration of their utilization. The scarcer resources are, the more important are efficiency and adaptability in their use. This thesis investigates three different contributions addressing diverse planning processes of scarce resources in hospitals. Intentionally, it appraises how the consideration of flexibility in these processes affects the resources' usage.

Hospitals are currently facing significant organizational and financial challenges in planning and scheduling, threatening their capacities to act and work (cf. Osterloh, 2022, 2024). The medical workload increases from both rising patient demands and a decreasing number of hospitals (cf. Bölt, 2024; Statistisches Bundesamt, 2023a). At the same time, skilled experts, like physicians and nurses, are scarce and difficult to find on the job market (cf. Fuchs and Weyh, 2023; PwC Deutschland GmbH, 2022). Also, operating rooms belong to one of the most vital yet limited resources, with an average surgery-to-patient ratio of 0.9 surgeries per patient (cf. Statistisches Bundesamt, 2023b). Hospitals need to accommodate to varying patient needs and the dynamic nature of medical procedures to ensure timely patient care. The complexity in these settings rises, as hospital resource management involves intricate coordination among various departments, medical staff, and patients. For instance, operating rooms require thorough scheduling taking multiple factors into account, such as diversity of surgical procedures, varying durations of treatments and recovery, availability of specialized staff or equipment, pre- and post-operative care capacities, and the need to integrate emergencies (cf. Harris and Claudio, 2022; Shehadeh and Padman, 2022). Besides that, further challenges can occur in personnel scheduling. A variety of pitfalls may disturb fluent working processes in hospitals, such as internal regulations regarding the quality of care, staff absences, coordination efforts, individual preferences, diverse sense of fairness, and technical issues (cf. Hutfluss et al., 2024). Consequently, planning strategies in hospitals are required to be flexible and dynamic to address effective and efficient utilization of resources.

Considering complex planning processes and wide-ranging managerial tasks, decision support systems have become a sufficient assisting tool. Since the early 20$^{\text{th}}$ century, methods of operations research have been the main component of such systems, modeling real-world-inspired problems and providing adequate solution algorithms. Recently, decision support systems gained another capable component as machine learning methods are constantly improving (cf. Gupta et al., 2022). Either of these fields of studies are incorporated in the contributions of this thesis.

The term *flexibility* is defined as “the ability to [...] be changed easily according to the situation” (Cambridge University Press, 2024). For the medical environment, Nickl-Weller et al. state that “the ideal of resource-efficient and patient-centred health care goes hand in hand with flexible adaptability” (Nickl-Weller et al., 2024, p. V). Precisely, flexibility in hospitals’ resource management should, on the one hand, facilitate the handling of unforeseen, short-term circumstances, and on the other hand, provide medium- and long-term coping strategies to deal with resource shortages.

Planning tasks occurring in hospital settings are pointed out in detail in the framework provided by Hans et al., 2012. The framework considers medical planning, resource capacity planning, materials planning, and financial planning. Exemplary tasks for all categories are named with respect to the strategical, tactical, offline operational, and online operational planning levels. A slightly modified subset of the framework is shown in Figure 1 which shows the entries for medical planning and resource capacity planning. The contributions of this thesis evaluate the consideration of flexibility in some of the listed tasks, as outlined subsequently.

	Medical planning	**Resource capacity planning**
Strategic	research, development of medical protocols	case mix planning, capacity dimensioning, workforce planning
Tactical	treatment selection, protocol selection	block planning, staffing, admission planning
Offline operational	diagnosis and planning of an individual treatment	appointment scheduling, workforce scheduling, surgery advance and allocation scheduling
Online operational	triage, diagnosing emergencies and complications, isolation decisions	monitoring, emergency coordination, re-planning

Figure 1.: Reduced and modified version of the framework for healthcare planning and control by Hans et al., 2012

Strategic planning addresses long-term decisions for the next two to ten years. In this step, hospitals plan the dimensions of their resources’ capacities rather than details on the exact realizations. These decisions influence all underlying shorter planning levels. In the next stage, tactical planning covers a mid-term horizon of several months up to one or two years. It provides the frame conditions for short-term working routines. For instance, block capacities for treatments are determined based on forecasted demands. Additionally, staffing for different departments or collaborations between departments are considered (cf. Hans et al., 2012). Contribution 2 refers to tactical and partially to strategic resource capacity planning. It aims to optimize the size of the workforce in several flexible settings. Therefore, it investigates flexible shift scheduling policies and multiple volatile demand profiles that are imaginable for hospital departments. Contrasting the results for several settings with each other provides insights into the benefit of addressing flexibility in the planning process. The outcomes could be useful for tactical or strategic decision making with respect to staff employment or the staffs' working departments.

Operational planning covers short-term planning tasks due within the next weeks, days, or hours. Precisely, the offline operational level addresses the upcoming weeks and days, whereas online operational planning refers to the next days and especially hours. Common offline operational tasks for resource capacity planning are appointment and staff scheduling. For operating room planning, the tasks are separated into advance and allocation scheduling. The former places patients' appointments on a certain day of the upcoming planning period, and the latter sequences and allocates the individual treatments to specific rooms and time slots. Staff rosters are created and schedule employees to certain shifts or tasks. Moreover, online operational tasks consider monitoring and re-planning decisions. Possible unforeseen reasons that require a reorganization of capacities could be emergency occurrences, staff absences, or delays from extended surgery durations (cf. Hans et al., 2012; Rahimi and Gandomi, 2020). Contribution 1 models offline operational planning tasks and concludes insights for both offline and online operational decisions. The objective is to optimize integrated surgery allocation scheduling and surgical team's staff assignments with respect to different flexible scheduling options. In this context, the advantages of flexible room-changing opportunities for both surgical interventions and staff are evaluated. The information gathered may support decisions in operating room management, e.g., in allocation scheduling, staff assignments, and especially re-planning tasks.

In some cases, resource capacity planning is directly linked with medical planning. Considering the online operational level, physicians or nurses make decisions about triage, emergencies, or patient isolation. In this context, they consider both the patients' medical condition and the availability of resource capacities. Exemplary diagnoses requiring timely isolation and rearrangement of resource capacities could be noroviruses, influenza, or most recently COVID-19. Contribution 3 presents a machine learning classifier to predict a patient's COVID-19 infection. The estimation is used to conclude the necessity of isolating a certain patient. To be flexibly adjustable to the hospital's current situation, the contribution tests adaptable performance metrics within the estimation approach and evaluates the outcomes. Physicians and nurses could leverage the estimations of such systems to better utilize isolation capacities.

The key topic of this thesis is flexibility in the planning processes of scarce resources in hospitals. The three contributions of the thesis explore this field and generate insights by answering the following research questions:

1. Which effects do flexible room-changing opportunities for treatments and staff members have on the necessity of highly qualified medical experts in operating room scheduling?

2. How does the availability of flexible shift types affect the number of required employees for hospital departments?

3. How does the application of flexible performance metrics affect the reliability of predictions in a multiple classifier machine learning algorithm for disease recognition?

4. Is the consideration of flexibility significantly important in resource planning processes within hospitals?

The subsequent chapters of this thesis are arranged as follows: Section 2 provides an overview of the three contributions that form the core of this research. In Section 3, these contributions are discussed in detail, focusing on how they address the research questions previously introduced. Finally, Section 4 concludes the thesis by summarizing the findings and outlining directions for future research. The full versions of the unpublished contributions, along with a link to the published work, are included in the Appendix.

2. Summary of the contributions

This thesis consists of several contributions to the current literature addressing different variants of flexibility in the planning processes of scarce resources in hospitals. This section summarizes the three individual contributions which are provided in entirety in the appendix.

The first contribution examines integrated surgery allocation scheduling and staff task assignments, which are operational resource capacity planning tasks, as related to Figure 1. The contribution's insights address both offline and online operational planning. In this context, operating rooms and highly qualified staff are the scarce resources to be utilized efficiently. The second contribution considers strategic and tactical resource capacity planning. It targets scarce medical staff and delves into the effects of flexible shift and break scheduling. Lastly, the third contribution relates to scarce isolation capacities during the COVID-19 pandemic. By providing estimations about a patient's infection, it refers to online operational medical and resource capacity planning.

2.1. Analyzing the value of flexibility in integrated surgery and staff scheduling on a daily planning level applying a column generation heuristic

Ebel et al. (2024a) present a mixed-integer linear program (MILP) for integrated surgery allocation scheduling and staff task assignment in an operating room department. The decomposition-based reformulation of the model is solved with a column generation heuristic. The results of the computational study are evaluated with respect to different scopes of flexibility, in order to conclude pertinent insights. The complete research is provided in Appendix A and is going to be submitted to *Manufacturing & Service Operations Management*, ranked in category A, referred to the VHB-Rating score (Verband der Hochschullehrerinnen und Hochschullehrer für Betriebswirtschaft e.V., 2024).

Motivation Hospitals are increasingly facing challenges, like rising demand and staff shortages. At the same time, highly skilled medical staff and operating theaters are among the most critical, expensive, and scarce assets in hospitals. The combination of these aspects accounts for simultaneously optimizing surgery schedules and interrelated staff assignments. Researching this topic on a daily planning level primarily regards the sequencing and room allocation of surgeries, and the assignment of capable surgical teams to specific tasks. As both parts of the problem affect each other, the integrated formulation opens up opportunities to evaluate the benefits of planning different scopes of flexibility.

Flexibility The planning process in this contribution requires operating theaters and surgical teams. Either of these are scarce and costly resources that need to be planned efficiently. Flexibility especially regards the possibilities of surgeries' placements and teams' assignments. Therefore, the effects of varying room-changing opportunities for both surgical procedures and surgical teams are researched.

Methodology First, the authors develop a generic MILP scheduling the surgeries of one medical specialty across multiple operating theaters and the time periods of a day. Concurrently, it assigns surgical teams considering different qualification levels required to fulfill the treatments. To investigate flexibility, there are parameters that limit the maximum daily number of room changes per team and the options to shift surgeries to various rooms and time periods. Secondly, as the problem takes severe efforts to be solved as a MILP, the model is reformulated following Dantzig-Wolfe decomposition. This results in one restricted master problem and two subproblems which are embedded in a column generation heuristic. The task-sequencing subproblem determines surgery sequences, whereas the task-assignment subproblem develops work assignments for surgical teams. The algorithm is initialized using surgery sequences from the underlying data. Thus, the master problem lacks a complete enumeration of all possible schedules. Instead, suitable plans are generated as needed throughout the column generation process. In the end, the master problem is solved as an integer program to receive a valid solution. Thirdly, a problem-specific analytical lower bound is introduced and applied within the model reformulation. The lower bound serves to further accelerate the solution process.

Computational study To specify the relevant scope of the experiments, a real-world data set of a maximum care provider hospital is analyzed in detail. The information gathered is used to filter and label the data set for further steps of the study's setup. The resulting data pool is then used for the creation of artificial study instances. A three-step approach is applied to randomly draw surgical programs from the data pool and modify them systematically. This process addresses various drivers of flexibility, such as the surgeries' expected duration, their complexity, the daily opening hours, treatment's shifting options, and room changes of teams. Overall, 1620 instances are generated and solved by means of the column generation heuristic.

Results and findings Evaluating the benefit of the analytical lower bound, it significantly speeds up processing times from an average of 78.182 seconds to an average of 2.115 seconds. For 84.9% of all instances, the solution's quality stays the same despite the reduced runtime. Embedding the lower bound, optimal solutions are identified in 76.1% of all cases and the optimal aggregated number of surgical teams is reached in 94.9% of all instances. The main insights based on the results' comparison can be summarized as follows:

- The more flexible an instance's setting is, the more highly qualified teams can be replaced with lower qualified staff.
- Flexible room-changing options for the surgical teams are especially beneficial in rigid surgery scheduling settings.
- Flexible room-shifting options for surgical procedures are beneficial in any setting. The highest impact is reached with expanded opening hours and rigid staff scheduling settings.
- Flexible expansions of the opening hours are especially valuable with flexible surgery scheduling settings.

Overall, it turns out that including flexibility in surgery planning has a higher impact than flexibility in staff scheduling.

2.2. Analyzing staffing decisions subject to flexible scheduling and demand volatility in healthcare

Ebel et al. (2024b) provide an integer program (IP) for flexible shift scheduling of healthcare employees including flexible break assignments. Column generation is used as a solution approach for the model's decomposition-based reformulation. Analyzing the results of the computational study in the context of several settings regarding flexibility, staffing insights are derived. The complete research work is provided in Appendix B and is close to submission to the *Journal of*

Scheduling, ranked in category A, referred to the VHB-Rating score (Verband der Hochschullehrerinnen und Hochschullehrer für Betriebswirtschaft e.V., 2024).

Motivation A severe topic in hospitals is the accomplishment of workload despite staff shortages and absences. In this context, appropriate staffing decisions should not only consider efficiency but also viability of workforce schedules regarding shift and break assignments, as well as legal and contract regulations. The overall planning process relates to different settings such as flexible scheduling options and the underlying demand structure. Accordingly, by investigating both of these aspects in a flexible planning model, information about their impact on the resulting workforce sizes can be evaluated.

Flexibility The planning task examined in this contribution regards suitable staffing of hospital departments. Adequate medical personnel is scarce due to the job market's circumstances and accounts for an essential part of hospital costs. Flexibility can be induced into the planning process by enabling flexible shift regulations and break placement. To receive valuable long-term evidence regarding the workforce size, flexibility is evaluated with respect to various demand profiles.

Methodology At first, an IP is developed to assign a homogeneous set of medical staff to shifts for one week. To address flexibility, all possible shifts respecting a minimum and maximum duration are stored in an extensive set of shifts. Moreover, constraints for an implicit scheduling of breaks are included and assign a number of breaks to the time periods of a day. Next, the model is decomposed by means of Dantzig-Wolfe decomposition to facilitate the computational effort to solve the problem. Consequently, a restricted master problem and a pricing subproblem are formulated and incorporated into a column generation heuristic. The subproblem is used to generate weekly shift schedules that each employee could work. Besides that, the master problem determines how many employees should work a certain weekly schedule and decides how many breaks are scheduled for each day and time period. Suitable shift schedules are unknown at the beginning of the algorithm and are created iteratively throughout the algorithm's process. Hence, the master problem is unaware of all possible shift schedules, making the column generation procedure a heuristic approach. By the end of the column generation procedure, a valid solution is provided solving the master problem as an IP.

Computational study The study focuses on resulting workforce sizes with various settings of flexibility. Therefore, two types of flexibility profiles are included in the setup. First, demand profiles influence the instances' demand structure and levels. Second, shift profiles influence the instances' set of shifts, constraints for shift limitations, and the consideration of break assignment. In every demand profile, the demand increases in the morning until it reaches a constant level, the so-called block, and then decreases in the evening. Nevertheless, the demand profiles vary in the flexibility of demand blocks, the occurance of peaks in the demand pattern, and the quantity of staff required. Shift profiles either consider break scheduling or neglect it. Additionally, they either apply an inflexible, semi-flexible, or flexible shift system. Inflexible systems contain two or three shift types in the set of shifts. Semi-flexible and flexible systems contain all possible shifts, however, in semi-flexible settings only one shift per employee and week, i.e. per schedule, is allowed. In total, applying eight demand and eight shift profiles, 1600 instances are generated, solved, and analyzed.

Results and findings The column generation heuristic solves the problem instances in seconds. The average solution times for different settings vary between 0.1 second to 44.9 seconds. Moreover, the results are near-optimal with average gaps per setting ranging between 0.0% to 2.3%. Relating the gaps with the mean number of employees required in each settings, the results deviate by less than one worker from optimality on average. The key findings identified comprise the following:

- The number of peaks in the demand profiles has a minor impact on the staff size, whereas the flexibility of demand blocks slightly influences the workforce size.

- The more flexible shift profiles are, the lower the resulting workforce size. This effect is stronger when break planning is considered than when breaks are disregarded.

- The incorporation of flexible shift types enables workforce savings of more than 15% compared to inflexible shift systems.

- Incorporating flexible breaks into the model, more staff is required. As breaks are mandatory in real-world scenarios this gives a better estimation for planners than models without breaks.

- Most efficient staffing and high utilization are reached in flexible shift profiles with a working hours to workload ratio close to one. In inflexible shift profiles ratios up to 2.9 are observed indicating potential inefficiencies.

2.3. COVIDAL: A machine learning classifier for digital COVID-19 diagnosis in German hospitals

Bartenschlager et al. (2023) introduce a decision support algorithm estimating patients' COVID-19 infections in order to adequately utilize hospital isolation capacities. They developed a machine learning classifier based on patients' laboratory blood parameters. The algorithm's performance is tested by investigating data preparation processes, test set selection and flexible objective functions. This contribution has been published in *ACM Transactions on Management Information Systems*, ranked in category B, referred to the VHB-Rating score (Verband der Hochschullehrerinnen und Hochschullehrer für Betriebswirtschaft e.V., 2024).

Motivation By the beginning of the COVID-19 pandemic hospitals faced the challenge of how to cope with symptomatic patients. Patients infected with SARS-CoV-2 should be isolated to avoid further spread of the virus. In contrast, rooms for isolation are a scarce resource that should not be blocked by patients with other respiratory diseases. Using common test methods, POC antigen tests are fast but lack reliable results, and PCR tests take a severe amount of time until the result returns. Therefore, methods based on laboratory parameters are considered. Since blood samples can be evaluated faster than PCR tests, isolation time can be reduced and capacities can be released earlier for patients testing negative.

Flexibility The planning decision focused by this contribution concerns scarce isolation capacities. The handling of isolation beds and rooms is influenced by changes and insecurities in hospital environments. In this context, flexibility regards the adaptability of the algorithm's estimation to various circumstances. Therefore, the method is evaluated in multiple settings applying flexible objectives. Different goals and metrics are tested like sensitivity, specificity, F1-score, accuracy, and balanced accuracy.

Methodology The algorithm uses multi-center data of three different real-world sources. All of them contain multiple blood parameters and the patients' PCR results. To harmonize the structure, four different strategies are applied for data preparation. In the training phase, the algorithm creates any pairs of laboratory parameters available in the data set. Next, several supervised machine learning models are trained with each combination. Accordingly, a pool of 780 classifiers is created ready to be used in a multiple classifier system

(MCS). In the prediction phase, input weights and thresholds specify the decision makers' preferences regarding the selection of classifiers to be included in the calculation of the final prediction. In its basic version, COVIDAL considers two metrics for this, sensitivity and accuracy. Therefore, only highly sensitive or highly accurate classifiers are considered in the selection. Calculating the final prediction, the relative frequency over all classifiers' results in the selected set is calculated. The outcome ranging between 0.0 and 1.0 is then mapped to a COVID-negative estimation if the result is lower than 0.5 or a COVID-positive estimation, otherwise.

Computational study COVIDAL's performance is evaluated with respect to the different data preparation strategies, test sets, and objectives. The data preparation strategies comprise ordinally scaled features and three variants of cardinally scaled features. All of them use the same input data but represent it differently. There are two types of test sets applied, one common 65/35 split for training and test data, and one 95/5 split where COVIDAL's results are compared against the estimations of seven specialized physicians. Additionally, COVIDAL's results are benchmarked against basic machine learning classifiers. Diverse objectives are referred to as optimization strategies and affect the selection of classifiers for the final prediction. The focus on sensitivity can be replaced by a focus on specificity or the positive or negative F1-score. Moreover, accuracy can be replaced by balanced accuracy, i.e. the mean of sensitivity and specificity, which addresses imbalanced datasets.

Results and findings Validating various data preparations, COVIDAL reaches the best results for cardinally scaled features using real-valued data as received from the laboratory or normalized with mean and standard deviation. In these cases, sensitivity, specificity, and accuracy range between 84% to 90%, independent of the underlying training/test split. Similar performance is reached for the other data preparations if the 65/35 training/test split is applied, resulting in sensitivity, specificity, and accuracy between 80% to 90%. In the remaining cases, COVIDAL performs highly sensitive but shows deficits in specificity. Benchmarking COVIDAL's performance against common machine learning approaches, like neural networks, linear regression, or support vector machines, COVIDAL provides more consistency, contrasting with the fluctuating performance of standard AI algorithms concerning various data preparations. Comparing machine learning methods against the classification performance of medical experts, standard AI methods reach significantly worse specificity and accuracy than physicians. However, COVIDAL outperforms most of the physicians. Finally, comparing different objectives, sensitivity and specificity vary by about 20% and accuracy by about 10%, making adaptable objectives a suitable flexible decision making option in COVIDAL.

3. Discussion of the contributions

This thesis aims to provide insights regarding flexibility in medical decision making and hospital planning processes. In this context, the focus is set on scarce resources, like room capacities, bed capacities, and medical staff, that should be deployed as efficiently as possible. The following subsections discuss how the contributions introduced account for valuable insights in resource planning by answering the following research questions.

1. Which effects do flexible room-changing opportunities for treatments and staff members have on the necessity of highly qualified medical experts in operating room scheduling?

2. How does the availability of flexible shift types affect the number of required employees for hospital departments?

3. How does the application of flexible performance metrics affect the reliability of predictions in a multiple classifier machine learning algorithm for disease recognition?

4. Is the consideration of flexibility significantly important in resource planning processes within hospitals?

3.1. Flexible room-changing opportunities

Which effects do flexible room-changing opportunities for treatments and staff members have on the necessity of highly qualified medical experts in operating room scheduling?

Ebel et al. (2024a) provide an integrated surgery and surgical team scheduling approach which addresses interdependencies between both planning problems. The study focuses on different flexible options, like room-changing possibilities for either surgical teams, or surgeries, or both, as well as an enlarged planning horizon. The results are evaluated regarding the number of highly qualified and lower qualified teams required. In the following, all percentage values represent average outcomes over 108 cases, referred to as entities in the contribution.

In rigid staff settings, surgical teams stay in the same room throughout the whole day. Room-changing opportunities **for teams** mean that a team finishes their treatment in room A and then switches to room B and starts a new surgery. This turns out to be especially beneficial in rigid surgery scheduling settings where surgeries have a fixed initial room and time slot that must not be changed. In this context, team flexibility leads to staff scheduling improvements in about 30% of all cases. With increasing flexible scheduling options for surgeries the benefit of staff flexibility shrinks. If surgeries are allowed to be shifted in time but not to change the room, about 10% of all cases benefit from team flexibility. For all the beneficial cases, on average one highly qualified team can be replaced by a lower qualified team. Thus, the highly qualified specialists may be utilized elsewhere in the hospital. If surgeries are permitted full flexibility the effects of staff flexibility become dispensable.

Surgeries have an intended room and time slot up front. Room-changing opportunities **for surgeries** mean that a surgery may be shifted to another room or time slot to better fit into the overall schedule. The major effect of this option crystallizes with expanded opening hours and rigid staff scheduling settings, i.e., improvements in about 78% of all cases. However, also for usual opening hours and flexible staff scheduling settings, there is a reasonable improvement in about 60% of all cases. For other circumstances, values in between are reached. For the beneficial cases, on average 1.0 to 1.2 highly qualified teams can be replaced by a lower qualified team.

For further specification, the effects of extending **opening hours** flexibly are considered. If surgeries must not change rooms but teams are allowed to do so, enlarging the opening hours is beneficial for about 8% of all cases and can replace on average one highly qualified team with a lower qualified team. If surgeries may be shifted flexibly the extension is more beneficial as about 21% of all cases profit from an average replacement of 0.5 highly qualified teams, regardless of rigid or flexible team settings.

To conclude, options for surgery flexibility are more valuable with fixed team settings and vice versa. In summary, the results state that considering the explicit planning of surgery flexibility is more influential in reducing the required number of highly qualified teams than accounting for extra staff flexibility. This finding could provide promising information for re-planning decisions in operating room departments.

3.2. Flexible shift types

How does the availability of flexible shift types affect the number of required employees for hospital departments?

Ebel et al. (2024b) present an IP and a column generation heuristic to investigate the effects of flexible scheduling in healthcare. The settings within the experimental study regard several demand patterns and shift scheduling options. The latter comprise inflexible, semi-flexible, and flexible shift systems, as well as the consideration of flexible break scheduling. In inflexible settings, only two or three shift types are in the overall shift set. In semi-flexible and flexible settings, all possible shifts within a minimum and a maximum duration are eligible, however, in semi-flexible settings, an employee's schedule may only contain one shift type per week. The evaluations consider the required size of the workforce in various settings. In this context, the mean size of the workforce in each setting is an average result of 25 instances each.

Contrasting the results for **inflexible and semi-flexible settings** with respect to the total number of staff required, the semi-flexible settings always employ fewer staff. However, the potential reduction of the workforce differs depending on the underlying demand structure and the consideration of break planning. Settings having two demand peaks throughout a day and inflexible demand blocks reach only a slight improvement of 1.1% if break planning is ignored and 8.7% if break planning is incorporated. Besides that, settings with only one demand peak and flexible block demand reach the highest profit from the semi-flexible option. Their staff can be reduced by 19% without break planning and by 24% with break planning considered. Using the overall demand as a basis for staffing estimations, working capacities of 138% to 290% of the demand are required in inflexible settings. In semi-flexible settings the estimation can be reduced to capacities of 128% to 140%.

Evaluating further results by facing **inflexible and flexible settings**, even more workforce savings can be observed. Again, the minimum enhancement is reached for settings with two demand peaks and inflexible demand blocks. These settings profit from an average workforce reduction of 17.8% without consideration of breaks and 22% if breaks are included in the planning process. Moreover, settings with one demand peak and flexible block demand reach average workforce reductions of 28.9% and 35.4% excluding and including the consideration of break planning, respectively. While the staffing level of inflexible settings is about 138% to 290% of the demand on average, flexible settings get along with estimation factors of 110% to 120% on average.

The comparison of **semi-flexible and flexible settings** shows more constant improvements for different demand profiles. In all settings, the workforce reductions range between 12% to 17% regardless of considering or neglecting break planning. Hence, opening a semi-flexible setting up to full flexibility increases the benefit.

Consequently, the availability of flexible shift types in general lowers the number of employees required. This effect is stronger with complete flexibility than with semi-flexibility. However, the precise scope of the workforce savings depends on the underlying demand structure. Thus, hospital decision makers need to keep the trade-off between the shift settings discussed in mind when staffing personnel for hospital departments. The key estimators provided for each setting could be used as an indication to facilitate their decision process.

3.3. Flexible performance metrics

How does the application of flexible performance metrics affect the reliability of predictions in a multiple classifier machine learning algorithm for disease recognition?

Bartenschlager et al. (2023) developed an MCS to estimate COVID-19 infections in symptomatic hospital patients. In the computational study, flexible optimization strategies address the selection of suitable classifiers based on their performance metrics. The COVIDAL classifier is tested with several optimization strategies. The outcomes are evaluated with respect to the overall sensitivity, specificity, and further key metrics. In the following, ranges of results depict the algorithm's outcomes across multiple data preparation strategies and test sets.

In the basic version of COVIDAL, the classifiers for the final prediction are selected considering their **sensitivity and accuracy**. E.g., related to the input parameters, only classifiers reaching either a sensitivity of at least 80% or an accuracy of at least 90% are included to calculate the final prediction. The aim of this is to receive highly sensitive results so that patients are not mistakenly relieved from isolation and spread the virus.

Assuming an imbalanced dataset, as is the case for COVIDAL, the accuracy of a single classifier in the selection could still be high even if it has poor specificity. Therefore, one approach is the replacement of accuracy with **balanced accuracy**, i.e., the mean of sensitivity and specificity. Considering the performance metrics sensitivity and accuracy for the classifier selection leads to similar performance as applying sensitivity and balanced accuracy. The final predictions lead to algorithm sensitivity between 80% to 100%, and specificity between 45% to 90%. Consequently, with a focus on sensitivity both accuracy and balanced accuracy are suitable selection criteria.

Since infected patients should not only be detected but the positive prediction should also be reliable, another approach is to replace the sensitivity score with the **F1-score**. Utilizing the performance metrics F1-score and accuracy for classifier selection, the final results' sensitivity takes values between 82% to 100%, and specificity between 49% to 93%. With selection criteria F1-score and balanced accuracy, the prediction's sensitivity ranges between 85% to 100% and specificity between 49% to 94%. Hence, the algorithm tends to identify more uninfected patients if the classifier selection addresses not only a correct detection but also reliable positive predictions. From a practical perspective, this could help in releasing blocked isolation capacities.

Moreover, in some circumstances, decision makers could be especially interested in uninfected patients, so that sensitivity as a selection criterion is replaced with **specificity**. Incorporating specificity and accuracy as selection criteria leads to final results with sensitivity between 71% to 95% and specificity between 50% to 89%. While this outcome is similar to the former evaluation, the results turn out to be more specific and less sensitive, if the performance metrics specificity and balanced accuracy are used. In this case, the results lead to final sensitivity between 80% to 95% and specificity between 63% to 90%. With an interest in uninfected patients, this improves the algorithm's estimation.

In summary, the option to apply flexible performance metrics in the classifier selection of an MCS influences the algorithm's outcome without deterioration. Especially algorithms with an imbalanced underlying dataset, like COVIDAL, benefit from flexible adaptability. Changing the optimization strategy enables the decision maker to align the MCS' output with the practical situation and interest. Furthermore, flexible performance metrics provide a more extensive perspective on the algorithm's estimation which can be desirable in planning situations.

3.4. Flexibility in resource planning

Is the consideration of flexibility significantly important in resource planning processes within hospitals?

The research contributions introduced point out different types of flexibility in hospitals' planning processes. Namely, the prospects of flexible room-changing opportunities, flexible shift types, and flexible performance metrics in diagnosis are discussed.

Referring to the strategic and tactical planning level, **staffing decisions** benefit from flexible shift types and systems. Flexibility in shift scheduling yields workforce savings so that less staff is required to fulfill the workload. Since hospitals face severe staff shortages and suitable employees are scarce in the job market, flexibility in shift scheduling could enable more efficient staffing decisions.

Once employed, the decisions on the operational planning level regard **treatment scheduling and the staff's task assignments**. Specifically educated employees are working on tasks and treating patients, e.g. in operating rooms. Taking their qualifications into account, the treatments' sequencing and room allocation affect the number of highly skilled staff required in the operating room department. In this context, flexibility in terms of room-changing opportunities provides the chance to replace highly qualified surgical teams with lower qualified teams as surgeries are scheduled more efficiently. Moreover, the options for room-changing and information about their utility are beneficial in operational re-planning decisions where managers need to decide whether to switch treatments or staff.

Treating patients on the operational level, hospital physicians can be supported in **diagnosing infections and detecting the necessities of isolation** by medical decision making tools. Exploiting flexible performance metrics in such systems aligns the system's estimation with the hospital's circumstances. Hence, physicians can relate their decision to block or relieve resources, like isolation beds, to an estimation provided by a flexibly adjustable system.

Overall, the consideration of flexibility in hospitals' planning processes discloses options for improvements regarding both staff efficiency and material resource utilization. Thus, addressing flexibility and adaptability is of major importance for medical decision makers as it reveals fundamental information for further planning tasks.

4. Conclusion

Hospital management comprises a wide scope of strategic, tactical, and operational planning tasks. This thesis outlines the chances and benefits of including flexibility in the planning processes of scarce resources, like operating rooms, medical staff, or isolation beds. In the introduction, the complexity of hospital planning tasks is sketched and the necessity for flexibility is motivated. Alongside, the intention and the research questions of the thesis are presented. Next, in the main part of the thesis, the three contributions are summarized presenting their consideration of flexibility, the methodologies, and the main findings. The third section answers the research questions introduced and critically discusses them with respect to the contributions.

The contributions of the thesis address the following planning tasks within hospital management: First, an integrated surgery and staff scheduling optimization model along with an efficient solution algorithm is presented, which sequences surgeries in an operating room department and assigns qualified staff. This contribution addresses offline and online operational resource capacity planning. Second, optimal workforce sizes are determined using integer programming and a column generation algorithm. This concerns strategic and tactical resource capacity planning. Third, a machine learning algorithm to estimate patients' COVID-19 infections in order to handle isolation capacities is presented and deals with online operational medical and resource capacity planning.

All of the three contributions successfully consider flexible options within the respective planning process and enable the generation of valuable insights. The first contribution analyzes the effects of room-changing opportunities in operating rooms. It shows that room-changing possibilities for surgical teams are especially useful if surgeries are fixed. In contrast, room changes for surgeries have the highest impact if teams stay in a fixed room. Overall, it turns out that room-changing flexibility has a higher impact on surgery allocation than on team assignments. The second contribution evaluates the impact of flexible shift types on the required workforce size in different demand profiles. In this context, inflexible, semi-flexible, and flexible shift profiles are tested. For all demand structures, promising reductions of the workforce are reached if shift systems are enriched with flexibility. The highest reductions are reached in flexible settings as the workforce can be reduced by more than 30% in some settings. However, also the semi-flexible shift types provide adequate savings of more than 20% in some settings. Finally, the third contribution makes use of diverse performance metrics that could be applied in an MCS for COVID-19 diagnosis estimation. This estimation helps medical staff to assess the necessity to isolate a patient in a separate room. In its basic version, it evaluates classifiers' sensitivity and accuracy to include their results in the final calculations. Taking different circumstances in hospital departments and the training data set into account, the algorithm could be flexibly adapted to use F1-scores, specificity, or balanced accuracy. All of these options turn out to lead to justified results and provide a wider view on the patients' health condition.

The contributions discussed represent a sufficient foundation for the consideration of flexibility in planning processes of scarce resources in hospitals. Future research could intensify the topic by adding further studies towards this field. With respect to staffing levels, the impact of shift flexibility could be amplified by analyzing shift design and selection for concrete departments in more detail. Besides, additional types of flexibility, like part-time workers, could be addressed. These employees could further improve the ratio between staffs' working capacity and work demand while still meeting all regulations. Room-changing opportunities within operating room departments could be reviewed with an expanded scope that includes downstream units, like the post-anesthesia care unit, and accounts for employees that could work in both areas. Moreover, the consideration of stochasticity could be conductive to verify the current insights for practical settings. As research benefits from diverse perspectives, prospective work could apply alternative modeling approaches for the issue, like e.g., vehicle routing problem formulations. It is imaginable that alternative formulations open up space for further consideration of flexible options. Considering diagnosis estimations by means of machine learning, similar methods to the COVIDAL classifier could be set up regarding other virus diseases that tend to be predictable from blood parameters. Beyond the planning tasks discussed in detail within this thesis, flexibility could also be researched for further hospital tasks referring to medical or resource capacity planning, especially. In any of the planning processes, it is important that decision makers always consider a trade-off between optimal resource utilization, practical viability, and resilience.

References

Bartenschlager, Christina C.; Ebel, Stefanie S.; Kling, Sebastian; Vehreschild, Janne; Zabel, Lutz T.; Spinner, Christoph D.; Schuler, Andreas; Heller, Axel R.; Borgmann, Stefan; Hoffmann, Reinhard; Rieg, Siegbert; Messmann, Helmut; Hower, Martin; Brunner, Jens O.; Hanses, Frank and Römmele, Christoph (2023). "COVIDAL: A Machine Learning Classifier for Digital COVID-19 Diagnosis in German Hospitals". In: *ACM Transactions on Management Information Systems* 14.2, pp. 1–16. DOI: 10.1145/3567431.

Bölt, Ute (2024). "Statistische Krankenhausdaten - Grunddaten der Krankenhäuser 2022". In: *Krankenhaus-Report 2024*. Ed. by Jürgen Klauber; Jürgen Wasem; Andreas Beivers; Carina Mostert and David Scheller-Kreinsen. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 437–464. ISBN: 978-3-662-68791-8.

Cambridge University Press, ed. (2024). *Flexibility*. Cambridge. URL: https://dictionary.cambridge.org/dictionary/english/flexibility (visited on 14/08/2024).

Ebel, Stefanie S.; Brunner, Jens O. and Schimmelpfeng, Katja (2024a). "Analyzing the value of flexibility in integrated surgery and staff scheduling on a daily planning level applying a column generation heuristic". In: *Working Paper*.

Ebel, Stefanie S.; Grieger, Milena; Erhard, Melanie and Brunner, Jens O. (2024b). "Analyzing staffing decisions subject to flexible scheduling and demand volatility in healthcare". In: *Working Paper*.

Fuchs, Michaela and Weyh, Antje (2023). "Die Arbeitsmarktsituation in Krankenhäusern". In: *Krankenhaus-Report 2023*. Ed. by Jürgen Klauber; Jürgen Wasem; Andreas Beivers and Carina Mostert. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 33–47. ISBN: 978-3-662-66880-1.

Gupta, Shivam; Modgil, Sachin; Bhattacharyya, Samadrita and Bose, Indranil (2022). "Artificial intelligence for decision support systems in the field of operations research: review and future scope of research". In: *Annals of Operations Research* 308.1-2, pp. 215–274. ISSN: 0254-5330. DOI: 10.1007/s10479-020-03856-6.

Hans, Erwin W.; van Houdenhoven, Mark and Hulshof, Peter J. H. (2012). "A Framework for Healthcare Planning and Control". In: *Handbook of healthcare system scheduling*. Ed. by Randolph W. Hall. International Series in Operations Research & Management Science. Boston, MA: Springer Science+Business Media LLC, pp. 303–320. ISBN: 978-1-4614-1734-7. DOI: 10.1007/978-1-4614-1734-7{\textunderscore}12.

Harris, Sean and Claudio, David (2022). "Current Trends in Operating Room Scheduling 2015 to 2020: a Literature Review". In: *Operations Research Forum* 3.1. DOI: 10.1007/s43069-022-00134-y.

Hutfluss, Roman; Oppitz, Corinna; Ebel, Stefanie S.; Schüller, Markus and Brunner, Jens O. (2024). "Herausforderungen bei der Personaleinsatzplanung von Ärzten im Krankenhaus: Automatisierte Planung mit quantitativen Methoden unter Berücksichtigung von Fairness und Mitarbeiterwünschen". In: *Working Paper*.

Nickl-Weller, Christine; Stefanie Matthys; Tanja Eichenauer and Cor Wagenaar, eds. (2024). *Fast and Flexible: Planning for Adaptive Health Care Systems*. 1st ed. Vol. v.9. Health Care of the Future Series. Berlin: MWV Medizinisch Wissenschaftliche Verlag. ISBN: 978-3-95466-885-4.

Osterloh, Falk (2022). "Krankenhäuser: Die Herausforderungen sind groß". In: *Deutsches Ärzteblatt* 119.26, pp. 1169–1170.

Osterloh, Falk (2024). "Krankenhaus Rating Report: Lage der Häuser bleibt angespannt". In: *Deutsches Ärzteblatt* 121.14, pp. 926–928.

PwC Deutschland GmbH, ed. (2022). *Fachkräftemangel im deutschen Gesundheitswesen 2022*. Frankfurt am Main. URL: https://www.pwc.de/de/

gesundheitswesen-und-pharma/fachkraeftemangel-im-deutschen-gesundheitswesen-2022.html (visited on 14/08/2024).

Rahimi, Iman and Gandomi, Amir H. (2020). "A Comprehensive Review and Analysis of Operating Room and Surgery Scheduling". In: *Archives of Computational Methods in Engineering*. DOI: 10.1007/s11831-020-09432-2.

Shehadeh, Karmel S. and Padman, Rema (2022). "Stochastic optimization approaches for elective surgery scheduling with downstream capacity constraints: Models, challenges, and opportunities". In: *Computers & Operations Research* 137, p. 105523. ISSN: 03050548. DOI: 10.1016/j.cor.2021.105523.

Statistisches Bundesamt, ed. (2023a). *Grunddaten der Krankenhäuser 2022*. Wiesbaden. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankenhaeuser/_inhalt.html (visited on 14/08/2024).

Statistisches Bundesamt, ed. (2023b). *Operationen und Prozeduren der vollstationären Patientinnen und Patienten in Krankenhäusern 2022*. Wiesbaden. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankenhaeuser/Publikationen/Downloads-Krankenhaeuser/statistischer-bericht-operationen-prozeduren-5231401227015.html (visited on 14/08/2024).

Verband der Hochschullehrerinnen und Hochschullehrer für Betriebswirtschaft e.V., ed. (2024). *VHB-Rating 2024 für Publikationsmedien*. Göttingen. URL: https://www.vhbonline.org/service/vhb-rating-2024 (visited on 14/08/2024).

Appendix A.

Analyzing the value of flexibility

Ebel SS, Brunner JO, Schimmelpfeng K (2024): **Analyzing the value of flexibility in integrated surgery and staff scheduling on a daily planning level applying a column generation heuristic.**

Status: to be submitted soon to INFORMS Manufacturing & Service Operations Management; Category A.

Original Research Paper

Analyzing the value of flexibility in integrated surgery and staff scheduling on a daily planning level applying a column generation heuristic

Stefanie S. Ebel[1], Jens O. Brunner[1,2,3], and Katja Schimmelpfeng[4]

[1]Health Care Operations / Health Information Management,
Faculty of Business and Economics, Faculty of Medicine, University of Augsburg,
Universitätsstraße 16, 86159 Augsburg, Germany
[2]Department of Technology, Management and Economics,
Technical University of Denmark, Kongens Lyngby, Denmark
[3]Next Generation Technology, Region Zealand, Denmark
[4]Department of Procurement and Production, University of Hohenheim
Schwerzstraße 40, 70599 Stuttgart, Germany

Correspondence: jotbr@dtu.dk

State: 29 August, 2024

Analyzing the value of flexibility in integrated surgery and staff scheduling on a daily planning level applying a column generation heuristic

Abstract

Highly qualified medical specialists, as well as operating theatres, are amongst the most crucial, valuable, and scarce resources in hospitals. On a daily planning level, the main decisions in allocation scheduling concern the sequencing of patients' treatments in the operating room, along with a meaningful assignment of eligible medical staff members to incurring tasks. To solve both scheduling problems simultaneously, we provide a basic mixed-integer program, as well as decomposition-based model reformulations for a column generation algorithm. The heuristic creates and matches schedules for surgeries and for operating room teams. To validate the approach, we use data from a maximum care provider hospital in the south of Germany to generate test instances. All the instances are tested in a variety of scenarios, addressing different scopes of flexibility. The column generation heuristic provides optimal results in more than 76% of all runs. Our solution approach leads to results of high quality and good performance, and enables us to provide meaningful insights regarding the value of planning flexibility. Especially, we can deduce that flexibility in surgery sequencing is of higher impact for the integrated planning problem than the scope of planning staff flexible.

Keywords

operating room planning, allocation scheduling, flexible scheduling, column generation, healthcare, mixed-integer programming

Acknowledgment

This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and Grant no. 438507036.

1. Introduction

Throughout the last decades, the healthcare industry progressively faced challenges like increasing demand and staff shortages which both affect hospital costs and the quality of care [5, 13]. Since about 60% of hospital inpatients need at least one surgery and the operating room unit is responsible for about 40% of a hospital's cost, effective and efficient scheduling of surgical interventions, as well as an appropriate allocation of surgical staff, is a crucial task in operating room management [11, 24, 33, 34].

In general, operating room (OR) planning can be researched with regard to different planning levels. On the strategic level the goal of the so-called case-mix-planning problem is to identify an optimal composition of surgical cases to be executed by the hospital [18]. In tactical planning, weekly OR capacities are distributed among a hospital's medical specialties by means of the so-called master surgery schedule (MSS) [28]. However, the focus of our work is on the operational planning level where decisions can be separated in two phases: first, in advance scheduling a patient's surgery is assigned to a certain day within the upcoming planning horizon of usually one or a few weeks. Second, in allocation scheduling surgeries of a considered day are sequenced and assigned to rooms and times [26, 37]. Our aim is to schedule surgeries respective to time and room at a certain day and, additionally, to assign required human resources appropriately.

Our contribution is twofold: First, we develop an integrated mixed-integer program (MIP) to combine both allocation scheduling and staff task assignment. The MIP is decomposed and implemented as a problem-specific column generation heuristic. Second, we analyse the value of planning flexibility. Especially, we evaluate the possibility of room changes vs. room fixations for the surgical procedures as well as for the staff members. Moreover, the opening hours are considered. The numerical study is executed with 108 entities and a total of 1620 instances. The column generation heuristic solves more than 76% of all instances optimally. Overall, our computational experiments lead to high-quality and high-performance solutions that enable us to derive meaningful insights. We present key findings about the value and interrelationships of these factors. As one major outcome, our results confirm considering extra flexibility in surgery sequencing and allocation is of higher impact than extra flexibility in staff assignment.

The remainder of this paper is structured as follows: Section 2 provides an overview of the most relevant literature regarding integrated OR and staff scheduling. In section 3, we describe the underlying optimization problem in detail and provide a generic MIP model. In section 4, we adapt the model to develop adjusted models for solving the problem by means of a column generation approach. Subsequently, we also define the solution process. In section 5, the computational setup of our study regarding the value of flexibility is explained and experiments as well as their outcomes are presented. Finally, section 6 concludes our insights and provides remarks for future research.

2. Literature Review

OR planning received significant attention over the decades. Rahimi and Gandomi [26] survey decision levels, performance criteria, and solution approaches. Additionally, Harris and Claudio [15], Zhu et al. [37], and Gür and Eren [14] examine patient types and scheduling strategies. Samudra et al. [30] review surgical resource assignments on the operational planning level. On the other hand, also staff scheduling is well-studied. Burke et al. [8], Benazzouz et al. [7], and Ngoo et al. [23] outline optimization problems, assignment decisions, goals, constraints, and solution approaches in nurse scheduling. Erhard et al. [10] describe features, decision levels, modeling, objectives, and methodologies for physician scheduling. The study at hand considers both fields concurrently and reviews relevant literature in three steps: first, general integrated surgery and staff scheduling; second, the application of column generation algorithms; and third, the consideration of staff members as surgical teams.

Integrated surgery and staff planning According to Rachuba et al. [25], operating theatres, physicians, and nurses are amongst the most prominent factors of integrated planning in hospitals. Within the following review, the focus is on shared human resources in the OR department, i.e., anesthetists and nurses. Pham and Klinkert [24] first explored the integrated planning of surgeries and resources, like nurses, technologists, anesthetists, or equipment. They treat surgery scheduling as a job shop problem in a MIP. Testing 10 instances of 16 to 25 patients they detect bottlenecks and provide feasible solutions for small to medium cases. Similarly, Xiang et al. [36] consider surgeons, nurses, and anesthetists within an ant colony optimization approach and discrete event simulation. Tested on 5 instances with 10 to 30 patients, 2 to 6 operating rooms, and various algorithmic parameters, they prove their approach shortens the makespan and solves the problem efficiently. Belkhamsa et al. [6] seek to improve Xiang et al.'s [36] results by devising additional heuristics, like a genetic algorithm and iterated local search. They sequence surgeries considering human resources, assuming the staff-to-surgery assignments are predetermined. Tested on 10 instances with 40 to 50 patients and 6 to 10 operating rooms, their study reveals improved performance over benchmark results. Van Huele and Vanhouke [33] propose priority rules for physician and surgery scheduling, and embed them into various construction heuristics. They evaluate these heuristics across 27 instances, with 3 to 15 operating rooms, finding their approach applicable to most practical scenarios. Bandi and Gupta [4] research OR staffing and surgical case scheduling considering staff preferences like back-to-back and same-room scheduling. They provide a robust MIP. Testing their models with a real-world dataset they are able to reduce OR costs and overtime. In Latorre-Núñez et al. [19], surgery sequencing is treated as resource-constrained hybrid flow shop problem. The demands of anesthetists, nurses, and materials are summed up and limited by availability. Using a constructive heuristic and a genetic algorithm, they solve 30 instances with 15 to 40 patients, reducing bottlenecks and finding near-optimal solutions. Rath et al. [27] focus on sequencing and assigning critical resources, namely anesthetists, using a robust optimization approach to tackle uncertainty. They apply a model-based heuristic based on Kelley's algorithm and a sample average approximation. With reasonable results for 35 instances involving 10 to 65 patients and 2 to 23 operating rooms, they validate their model against real-world data and develop a practical decision support system.

Integrated surgery and staff planning applying column generation Column generation approaches are deployed in several planning problems of human resources and/or working tasks. Beliën and Demeulemeester [5] introduced the earliest branch-and-price approach for integrating nurse and surgery scheduling, managing nurse roster variability by linking it with OR schedules. Hence, they decompose the problem into two subproblems for nurse rosters and surgery schedules, and integrate them in a master problem. Considering advance scheduling, they adapt OR capacities of surgeons and assign nurses to shifts. Their approach, tested on 96 instances with 3 to 12 operating rooms, adequately handles the two-sided problem dimension. Hashemi Doulabi et al. [16, 17] research advance scheduling and allocation scheduling concurrently. Their subproblem generates daily surgical sequences by operating room, while the master problem selects columns over a weekly planning horizon and ensures human resource capacities. Branching is simplified using a constraint programming model for the subproblem. The branch-and-price approach is tested with 65 instances of 40 to 120 patients across 5 days and 6 operating rooms. Ultimately, the authors show improvements from their prior [16] to their subsequent approach [17] and highlight the superiority of branch-and-price compared to a MIP. Similarly, Naderi et al. [22] sequence surgeries and assign anesthetists for each day of a week. Using logic-based Benders decomposition, they explore diverse model decomposition options. Testing on 45 instances with 25 patients per week and 3 rooms, they find promising results adaptable to various hospital and industry settings. Lim and Mobasher [20] assume predetermined surgical sequences for a day, and focus on nurse assignment in surgery execution. Extending their earlier optimization study [21], they develop a column generation approach, tested against a MIP model on 3 instances. The tests aim for nurse assignments to 3 to 40 surgeries across 2 to 26 rooms, with the algorithm found to work effectively. Akbarzadeh et al. [1] develop a column generation approach for surgery scheduling and nurse duty assignment, integrating a diving heuristic. Tested on 320 instances with 10 to 30 patients scheduled across 2 to 4 rooms, the approach yields near-optimal solutions, addressing allocation scheduling and nurse re-rostering to maximize profit and meet actual demand. Tsang et al. [32] present a stochastic programming and distributionally robust optimization approach akin to Rath et al. [27]. They focus on column and constraint generation to reformulate their robust optimization model, enhancing it with symmetry-breaking constraints. Analyzing 12 instances with 100 scenarios, scheduling 15 to 80 patients to 6 to 32 rooms, they demonstrate the efficiency and practicality of their models and solution methods. Volland et al. [35] research daily and weekly task assignment for hospital logistic assistants, separating the task and staff scheduling components. They propose two column generation approaches with different decomposition strategies for task and shift scheduling. Evaluating with 48 instances of 130 to 500 tasks per week, they achieve optimal results, suggesting adaptability beyond healthcare to other industries.

Integrated surgery and surgical team planning Scheduling surgical teams instead of each resource separately is promising since team members' familiarity influences their productivity as evaluated by Avgerinos and Gokpinar [3]. Silva et al. [31] create an integer programming model for surgery scheduling and human resource integration. They differentiate renewable and non-renewable resources, with a focus on anesthetists, however, the model is open to various resource types. They apply two heuristics, validated in 144 instances with 40 to 90 patients, demonstrating near-optimal solutions. Guido and Conforti [12] aim to find pareto-optimal solutions for combined tactical and operational planning by assigning surgical teams to OR blocks and patients. Using a genetic algorithm, they solve 32 instances with 6 types, ranging from 130 to 1000 patients, achieving balanced solutions for multiple goals. Askarizadeh and Aziz [2] purpose to apply a genetic algorithm, data envelopment analysis, and simulation to improve OR scheduling. They assign surgical groups to operating rooms and time blocks, planning an in-depth case study at a partner hospital. Riise et al. [29] adopt an approach similar to Pham and Klinkert [24], breaking down patient treatment into activities and defining modes comprising necessary resources. These modes include both human and material resources and may be parameterized differently according to the planning level. Daily planning applies surgical teams of surgical nurses, anesthesia nurses, and anesthetists. Using an iterative search algorithm, they test their model on 10 daily instances (10 patients), 20 weekly instances (40 to 80 patients), and 10 monthly instances (728 patients), achieving satisfactory results across different planning levels. Calegari et al. [9] define teams as comprising surgeons, nurses, and anesthetists. They develop a five-step heuristic to schedule surgeries, while balancing resource availability and surgery volume. Testing on six instances with 10 to 20 patients and doing a pilot project, they conclude that their approach surpasses real-world scheduling methods.

Our work uniquely addresses integrated allocation scheduling and staff assignment with differently qualified surgical teams using a column generation approach. Compared to other decomposition-based approaches, we consider all shared human resources in our formulation instead of a single resource type. Unlike some prior works, our calculations are based on real-world data. Table 1 summarizes the differences in resource scope, planning horizon, methodology, and data utilization among the discussed papers.

		Staff					Methodology			
Author	**Year**	**Teams**	**Anesthetists**	**Nurses**	**Further**	**Planning horizon**	**Approach**[1]	**MIP**	**Decomposition based approach**[2]	**Real-world data**[3]
Akbarzadeh et al.	2020			✗		2 days	H	✗	CG+	✗
Askarizadeh and Aziz	2018	✗	✗	✗		1 day	H,S	✗		✗
Bandi and Gupta	2020				✗	1 day	E	✗		✗
Beliën and Demeulemeester	2008			✗		4 weeks	E		BP	
Belkhamsa et al.	2018		✗	✗		1 day	H			(✗)
Calegari et al.	2020	✗	✗	✗		1 day	H			✗
Guido and Conforti	2017	✗	✗	✗		1 week	H	✗		✗
Hashemi Doulabi et al.	2014				✗	1 week	E		CG	
Hashemi Doulabi et al.	2016				✗	1 week	E		BP+	
Latorre-Núñez et al.	2016		✗	✗		1 day	H	✗		
Lim and Mobasher	2012			✗		1 day	H	✗	CG	✗
Naderi et al.	2021		✗			1 week	E	✗	BD	✗
Pham and Klinkert	2005	✗	✗	✗	✗	1 day	E	✗		(✗)
Rath et al.	2017		✗			1 day	H	✗		✗
Riise et al.	2016	✗	✗	✗	✗	diverse	H			✗
Silva et al.	2015	✗	✗	✗	✗	2 days	H	✗		✗
Tsang et al.	2024		✗			1 day	H	✗	CG+	✗
van Huele and Vanhoucke	2015		✗			1 week	H	✗		
Volland et al.	2017				✗	1 week	H	✗	CG	✗
Xiang et al.	2015		✗	✗		1 day	H,S	✗		
our approach		✗	✗	✗		1 day	H	✗	CG	✗

[1] E: exact solution approach, H: heuristic solution approach, S: simulation
[2] BD: benders decomposition, BP: branch-and-price, CG: column generation, +: additional extension
[3] when in brackets the authors used data from previous literature

Table 1: Literature Overview

3. Problem description and formulation

The focus of this work is the so-called allocation scheduling which aspires the following situation: Patients who need surgery have been assigned to a specific day upfront. Next, the management of the OR determines the starting time and room for each intervention. In this context, the planner needs to consider practical requirements of the treatments, like staff experience, room equipment, time-related aspects, and many more. We schedule a set of inpatient surgeries $s \in S$ belonging to the same medical specialty, but differing in requested qualification levels $q \in Q$ necessary to conduct the treatment. The surgeries may be assigned to rooms (locations) $l \in L$ at different time periods $t \in T$ of one day. Additionally, surgical teams (working teams) $w \in W$ consisting of anesthetists, anesthesia nurses, and surgical nurses, are scheduled to work in a specific room at a certain time slot. The composition of each team is assumed to be predetermined. A team has a qualification level $q \in Q$ that emphasizes their experience, e.g., $q = 1$ means low experience and $q = 2$ means high experience. The teams' qualifications need to be matched against the surgeries' requirements where downgrading is allowed, i.e., also a highly qualified team is able to do a surgery requiring only low qualification.

To induce meaningful decisions, penalty factors C^{miss} for surgery cancellation, C_q^{out} for potential staff undercoverage, and team weights C_w addressing the team's qualification level are used. The staff pool contains enough surgical teams to fulfill all treatments and teams are available for the whole day. However, taking staff shortages and unforeseen events into account, they should work as efficient as possible, i.e., highly qualified teams are given higher C_w. For example, if team $w1$ has a low qualification and team $w2$ has a high qualification, the weights are set as $C_{w1} < C_{w2}$. Therefore, highly qualified teams are only considered to work in the OR if a team with lower qualification cannot fulfill the tasks. Binary parameters $\hat{A}_{wqlt}$ denote how teams may be assigned, i.e., encode their qualification and capability. For consistency reasons, surgical teams are required to stay in the same room for a minimum number of periods T^{min} after starting work there. Respectively, L^{max} restricts the number of room changes which are allowed for each team throughout the day. The overall capacity of operating rooms is considered to be sufficient to conduct all the surgeries. The deterministic duration D_s denotes the number of time periods to complete surgery s. Considering the surgery's required qualification level q, the staff demand, namely one team per surgery, is expressed by the binary parameter H_{sq}. Additionally, binary parameters $\hat{X}_{slt}^{\text{start}}$ summarize the information when and where surgeries may start, i.e., a time period's viability to conduct the surgery.

The decisions regard an adequate matching of personnel demand and personnel supply. At first, binary variables x_s^{miss} capture if a surgery s is conducted or cancelled, and binary variables y_w^{work} state if team w is working in the OR department or not. If surgery s takes place, one of its binary variables x_{slt}^{start} denotes that surgery s starts in room l in time period t. Subsequently, binary variables x_{slt}^{on} denote whether surgery s is running in room l during time period t. Based on this, a non-negative integer variable n_{qlt} captures the resulting personnel demand for teams experienced with at least qualification level q in room l in time period t. In contrast, if team w is working, binary variables y_{wlt}^{start} indicate if the team is starting a service in room l in time period t. Based on the start, binary variables y_{wlt}^{on} denote that team w is working in room l during time period t, and more precisely, binary variables a_{wqlt} indicate that team w has at least qualification level q and is serving in room l in time period t. To match supply and demand, the variables a_{wqlt} and n_{qlt} are compared to each other. To ensure the optimization problem's feasibility, additional non-negative integer variables u_{qlt} capture potential staff shortages if the demand for qualification level q in room l in time period t cannot be met. An overview of all indices, parameters, and variables is given below.

Indices and sets

$s \in S$	set of surgeries
$w \in W$	set of surgical (working) teams
$q \in Q$	set of qualification levels
$l \in L$	set of operating rooms (locations)
$t \in T$	set of time periods per day

Parameters

C^{miss}	weight factor for cancelled surgeries
C_w	weight factor for deploying team w in the OR
C_q^{out}	weight factor for understaffing
$\hat{A}_{wqlt}$	1 $\forall l \in L$ if team w has at least qualification level q, 0 otherwise
$T^{\min}$	minimum duration working in the same room
$L^{\max}$	maximum number of allowed room changes for teams
$\hat{X}_{slt}^{\text{start}}$	1 $\forall l \in L$ if period t is part of the starting time window of surgery s, 0 otherwise
D_s	duration of surgery s
H_{sq}	demand of surgery s for qualification level q

Decision variables

x_s^{miss}	$\in \{0;1\}$	1 if surgery s is cancelled, 0 otherwise
x_{slt}^{start}	$\in \{0;1\}$	1 if surgery s starts in room l in time period t, 0 otherwise
x_{slt}^{on}	$\in \{0;1\}$	1 if surgery s is running in room l during time period t, 0 otherwise
n_{qlt}	$\in \mathbb{N}_0$	demand for qualification level q in room l in time period t
y_w^{work}	$\in \{0;1\}$	1 if team w is working in the OR, 0 otherwise
y_{wlt}^{start}	$\in \{0;1\}$	1 if team w starts working in room l in time period t, 0 otherwise
y_{wlt}^{on}	$\in \{0;1\}$	1 if team w is working in room l in time period t
a_{wqlt}	$\in \{0;1\}$	supply of team w for qualification level q in room l in time period t
u_{qlt}	$\in \mathbb{N}_0$	staff undercoverage for qualification level q in room l in time period t

$$
\begin{aligned}
\min \quad & \sum_{w \in W} C_w \cdot y_w^{\text{work}} + \sum_{s \in S} C^{\text{miss}} \cdot D_s \cdot x_s^{\text{miss}} + \sum_{q \in Q} \sum_{l \in L} \sum_{t \in T} C_q^{\text{out}} \cdot u_{qlt} && (1.1)\\
\text{s.t.} \quad & \sum_{w \in W} a_{wqlt} + u_{qlt} \geq n_{qlt}, && \forall q \in Q, l \in L, t \in T, && (1.2)\\
& \sum_{l \in L} \sum_{t \in T} x_{slt}^{\text{start}} + x_s^{\text{miss}} = 1, && \forall s \in S, && (1.3)\\
& \sum_{s \in S} x_{slt}^{\text{on}} \leq 1, && \forall l \in L, t \in T, && (1.4)\\
& \sum_{\zeta = \max(1, t - D_s + 1)}^{t} x_{sl\zeta}^{\text{start}} = x_{slt}^{\text{on}}, && \forall s \in S, l \in L, t \in T, && (1.5)\\
& \sum_{s \in S} H_{sq} \cdot x_{slt}^{\text{on}} = n_{qlt}, && \forall q \in Q, l \in L, t \in T, && (1.6)\\
& \sum_{l \in L} y_{wlt}^{\text{start}} \leq y_w^{\text{work}}, && \forall w \in W, t \in T, && (1.7)\\
& \sum_{l \in L} y_{wlt}^{\text{on}} \leq y_w^{\text{work}}, && \forall w \in W, t \in T, && (1.8)\\
& T^{\min} \cdot y_{wlt}^{\text{start}} \leq \sum_{\zeta = t}^{t + T^{\min} - 1} y_{wl\zeta}^{\text{on}}, && \forall w \in W, l \in L, t \in \{1, \ldots, |T| - T^{\min} + 1\}, && (1.9)\\
& \sum_{l \in L} \sum_{t \in T} y_{wlt}^{\text{start}} \leq 1 + L^{\max}, && && (1.10)\\
& \hat{A}_{wqlt} \cdot y_{wlt}^{\text{on}} = a_{wqlt}, && \forall w \in W, l \in L, t \in T, && (1.11)\\
& y_{wlt}^{\text{on}} - y_{wl,t-1}^{\text{on}} \leq y_{wlt}^{\text{start}}, && \forall w \in W, l \in L, t \in T \setminus \{1\}, && (1.12)\\
& y_{wlt}^{\text{on}} \leq y_{wlt}^{\text{start}}, && \forall w \in W, l \in L, t = 1, && (1.13)\\
& x_s^{\text{miss}} \in \mathbb{B}, && \forall s \in S, && (1.14)\\
& x_{slt}^{\text{start}} \leq \hat{X}_{slt}^{\text{start}}, x_{slt}^{\text{start}}, x_{slt}^{\text{on}} \in \mathbb{B} && \forall s \in S, l \in L, t \in T, && (1.15)\\
& y_w^{\text{work}} \in \mathbb{B} && \forall w \in W, && (1.16)\\
& y_{wlt}^{\text{start}}, y_{wlt}^{\text{on}} \in \mathbb{B} && \forall w \in W, l \in L, t \in T, && (1.17)\\
& a_{wqlt} \in \mathbb{B} && \forall w \in W, q \in Q, l \in L, t \in T, && (1.18)\\
& u_{qlt}, n_{qlt} \in \mathbb{N}_0, && \forall q \in Q, l \in L, t \in T && (1.19)
\end{aligned}
$$

The objective (1.1) minimizes the weighted value resulting from surgical sequencing, team assignments, and understaffing. Long surgeries lead to a higher penalty if they are cancelled as they take more effort if they need to be re-scheduled. Constraints (1.2) ensure that work supply exceeds work demand, possibly by means of a respective value for staff shortage. The latter are the key constraints to link surgery scheduling and staff scheduling. Constraints (1.3) state that each surgery starts exactly once or is cancelled. In constraints (1.4), an overlapping of multiple surgeries at the same room and time is prohibited. This also ensures that the time-related OR capacity is respected and operating theatres are not overbooked. Next, constraints (1.5) and (1.6) serve as variable linkage for surgery scheduling. Constraints (1.5) state that surgery s is running in room l at time period t, when it starts within a respective backwards time window. Once a surgery is running, constraints (1.6) calculate the resulting human resource demand. Afterwards, the constraints regarding staff scheduling are pointed out. Constraints (1.7) ensure that team w can only start a service in time period t in any room l if the team is working at all. Comparably, constraints (1.8) limit team w to work on only one service at a time and only if the team is working at all. Moreover, constraints (1.9) guarantee that once a team w starts working in room l, they stay working in this room for at least $T^{\min}$ time periods. Additionally, constraints (1.10) restrict the number of room changes for each team w to at most $L^{\max}$ changes. Constraints (1.11) to (1.13) serve as variable linkage for staff scheduling. Constraints (1.11) state that if team w is working in room l at time period t the team serves every qualification level q they are qualified for. Since only one surgery takes place in each room at each time period, this is a valid interpretation of downgrading work assignments. Finally, constraints (1.12) capture the start of a service if team w is working in room l at time period t and has not been working there in the previous time period. Likewise, constraints (1.13) detect a start in the first time period. The remaining constraints serve as variables' domain restrictions. In this context, the constraints (1.14) to (1.18) define the domains of binary decision variables, and constraints (1.19) define the domains of non-negative, integer variables. Additionally, (1.15) ensures that surgeries only start in valid time periods.

4. Solution approach

In order to solve the problem efficiently, a reformulation based on a Dantzig–Wolfe decomposition is developed. This leads to one master problem and two subproblems which are solved by means of a column generation procedure. The first subproblem is called task-sequencing subproblem (TSSP). It is used to determine promising surgical sequences (addressing (1.3) to (1.6), (1.14), (1.15), and partly (1.19)). In contrast, the so-called task-assignment subproblem (TASP) creates service assignment plans for aggregated groups of OR teams (referring to 1.7 to 1.13, 1.16 to 1.18). Finally, the restricted master problem (RMP) does not include all columns and gets absent columns, i.e., columns created by the subproblems, and matches work supply with work demand to minimize the resulting objective (applying reformulated 1.2 and partly 1.19).

Aiming to reduce symmetry throughout the algorithm, the teams $w \in W$ are aggregated. To manage multiple teams of the same qualification level q, teams are organized into groups G. Inversely, each group $g \in G$ consists of a number of W_g identically qualified teams. Groups are disjunctive so that each team belongs to exactly one group. Hence, the TASP is not solved per team but only per group. Moreover, the decision in the master problem switches from assigning specific teams to assigning a number of teams per group. Due to the possibility of downgrading assignments for a team of group g, a differentiated notation of sets g and q is still reasonable. In the following sections 4.1 and 4.2, the models for the RMP, TSSP, TASP are explained, and in section 4.3, the column generation procedure is pointed out. In addition to the previously stated indices and sets the models also apply the following sets: The set of teams $w \in W$ is replaced by a set of groups $g \in G$. The TSSP creates sequences of surgeries, namely task-sequence plans $p \in P$. On the other hand, the TASP generates so-called task-assignment plans $j \in J^g$ in which teams of group g are assigned to work specific services. These sets are not disjunctive, since an assignment plan is applicable for the group it is originally created for, and also for groups with a higher qualification level.

Additional indices and sets

$g \in G$	set of groups of surgical teams
$p \in P$	set of task-sequence plans
$j \in J^g$	set of task-assignment plans per work group with $\bigcup_{g \in G} J^g = J$

4.1. Restricted master problem (RMP)

The RMP serves as matching problem of task-sequence plans and task-assignment plans, i.e., work demand and work supply. Decomposing the model from section 3, the parameters and decision variables change slightly, as follows: The team weight factor C_w becomes C_g due to the aggregation of teams into groups. The original decisions regarding n_{qlt} and a_{gqlt} along with their objective value effects are transferred into the subproblems. Thus, the generated columns are associated with new weights, parameters, and decision variables for proper and wise selection in the RMP.Therefore, C_j^g weights the task-assignment columns $j \in J^g$ from the TASP and A_{jqlt}^g stores their staff assignments. An integer variable λ_j^g serves to select how many teams work this plan. Respectively, task-sequence columns $p \in P$ from the TSSP are used with C_p and their demand is captured in N_{pqlt}. A binary variable ρ_p captures whether to choose a particular column.

Additional parameters

C_j^g	weight factor of task-assignment plan j for a single team of group g
C_p	weight factor of task-sequence plan p
A_{jqlt}^g	1 if task-assignment plan j is valid for group g and enables a service with qualification level q in room l in time period t, 0 otherwise
N_{pqlt}	demand of task-sequence p regarding qualification level q in room l in time period t
W_g	number of teams in group g

Additional decision variables

ρ_p	$\in \{0;1\}$	1 if task-sequence plan p is chosen, 0 otherwise
λ_j^g	$\in \mathbb{N}_0$	number of teams of group g working task-assignment plan j

$$\min \quad \sum_{g \in G} \sum_{j \in J} C_j^g \cdot \lambda_j^g + \sum_{p \in P} C_p \cdot \rho_p + \sum_{q \in Q} \sum_{l \in L} \sum_{t \in T} C_q^{\text{out}} \cdot u_{qlt} \tag{2.1}$$

$$\text{s.t.} \quad \sum_{g \in G} \sum_{j \in J^g} A_{jqlt}^g \cdot \lambda_j^g + u_{qlt} \geq \sum_{p \in P} N_{pqlt} \cdot \rho_p, \quad \forall q \in Q, l \in L, t \in T, \tag{2.2}$$

$$\sum_{p \in P} \rho_p = 1, \tag{2.3}$$

$$\sum_{g \in G} \sum_{j \in J^g} \lambda_j^g \leq W_g, \quad \forall g \in G, \tag{2.4}$$

$$\rho_p \in \{0;1\}, \quad \forall p \in P, \tag{2.5}$$

$$\lambda_j^g \in \mathbb{N}_0, \quad \forall g \in G, j \in J^g, \tag{2.6}$$

$$u_{qlt} \in \mathbb{N}_0, \quad \forall q \in Q, l \in L, t \in T \tag{2.7}$$

The objective (2.1) still minimizes the weighted value resulting from surgery sequencing, staff assignments, and understaffing. Constraints (2.2) equivalently match work supply with work demand in the restricted master problem. The dual variables for these constraints are denoted as $\delta_{qlt} \geq 0$. The constraint (2.3) serves as convexity constraint for the task-sequencing plans with dual variable $\mu \in \mathbb{R}$. Similarly, constraints 2.4 serve as convexity constraints for the task-assignment plans with dual variable $\pi_g \leq 0$. Finally, constraints (2.5) to (2.7) address the decision variables' domains.

Consideration of an analytical lower bound Basically, column generation terminates once no further columns with negative reduced costs are identified, i.e., LP-optimality of the master problem is achieved. However, this process suffers from the so-called tailing-off effect where a multiplicity of iterations is necessary to terminate the algorithm. Therefore, to speed up the solution process and to facilitate the identification of valid integer solutions an additional time-related lower bound is embedded into the RMP. It is possible to pre-calculate the minimum number W_q^{Min} of surgical teams with at least qualification level q. Constraints 2.8 address the minimum demand of teams and are related to the dual variable $\tau_q \geq 0$.

$$\sum_{g \in G: q_g \geq q} \sum_{j \in J^g} \lambda_j^g \geq W_q^{\text{Min}} \quad \forall q \in Q \tag{2.8}$$

In an example the model includes two qualification levels, namely q_1 ="low" and q_2 ="high". The surgeries along with their respective demands and duration are known. E.g., there are three rooms and the opening hours per room are 10h. The set of surgeries comprises five surgeries requiring high-qualified staff and taking 14h. Another five surgeries require low-qualified staff and take 15h. From this information, it is possible to sum up the capacity requirement for each qualification level. To determine the minimum demand $W^{\text{Min}}_{\text{high}}$, the duration of all surgeries requiring highly qualified staff is summed up, divided by the overall opening hours $|T|$, and rounded up to be an integer value, i.e., $W^{\text{Min}}_{\text{high}} = \lceil \frac{14}{10} \rceil = 2$. As the lower bound relates to the minimum number of teams having *at least* qualification level q, the value of $W^{\text{Min}}_{\text{low}}$ sums up the total duration of all surgeries, i.e., those requiring either highly qualified staff or lower qualified staff, which results in $W^{\text{Min}}_{\text{low}} = \lceil \frac{14+15}{10} \rceil = 3$.

4.2. Subproblems for task-sequencing and task-assignment

In the following, the setup of the two subproblems is outlined.

Task-sequencing subproblem (TSSP) The TSSP is used to create meaningful surgery sequences, i.e., it schedules the surgical procedures across all the rooms. Along with the surgery sequence this subproblem also determines the resulting human resource demand to provide the representation of sequencing columns in the master problem. The objective function determines the weight factor of the intended column and considers the dual values of the RMP as shown in Equation 3.1. The first term of the subproblem's objective aims to avoid cancellations of surgeries and refers to the respective part of objective 1.1. A new surgery schedule is generated if the reduced costs given in 3.1 reach a negative value. The full model with all constraints is given in section A.1.

$$\min \ \bar{C}_p = \sum_{s \in S} C^{\text{miss}} \cdot D_s \cdot x^{\text{miss}}_s + \sum_{q \in Q} \sum_{l \in L} \sum_{t \in T} \delta_{qlt} \cdot n_{qlt} - \mu \tag{3.1}$$

Task-assignment subproblem (TASP) The TASP is solved per group $g \in G$ and determines promising working patterns, i.e., schedules of service assignments. These columns may be allocated to the teams in group g and further teams which hold the necessary qualification levels. The objective function determines the weighted value of the intended column for one team of group g. Along with this it applies the dual values from the RMP to receive respective reduced costs as presented in Equation 4.1. If the objective 4.1 provides negative reduced costs a promising new column is identified. The full model is given in section A.2.

$$\min \ \bar{C}^g_j = C^g - \sum_{q \in Q} \sum_{l \in L} \sum_{t \in T} \delta_{qlt} \cdot a^g_{qlt} - \pi_g - \sum_{q \in Q : q \leq q_g} \tau_q \tag{4.1}$$

4.3. Column generation heuristic

The models presented for the RMP, TSSP and TASP are embedded into a column generation algorithm. Before the heuristic starts initial columns are generated. The easiest way to initialize the problem is an empty column for the medical specialty, i.e., a column cancelling all surgeries with an inadequately high objective weight so that better columns will be generated subsequently. However, we used the original surgical sequence from the underlying data sample to initialize the algorithm.

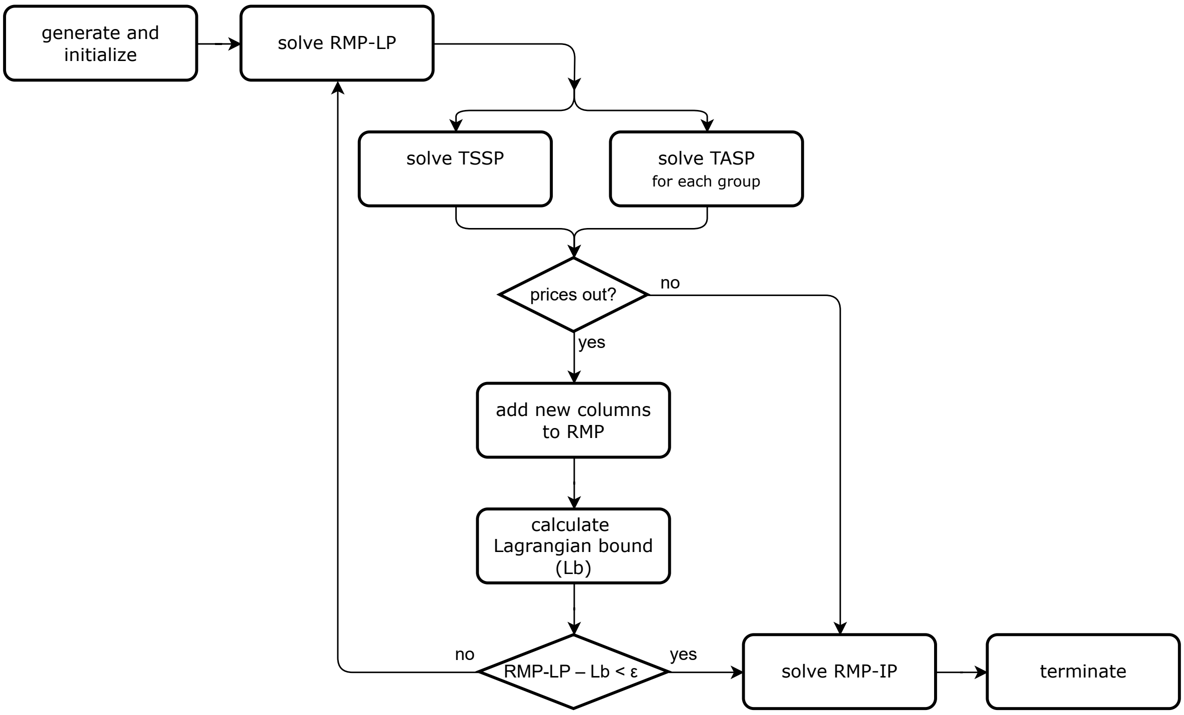


Figure 1: Flowchart of the column generation procedure

Afterwards, the column generation follows the process depicted in Figure 1. Once the algorithm is set up and initialized, it starts looping through its iterative behaviour. In the beginning of an iteration the RMP is solved as a linear program (RMP-LP) and its objective as well as the dual values are stored. After the master problem has been handled, the subproblems are evaluated: the TSSP gets solved once and the TASP is solved for all groups. All subproblems are solved optimally. If any subproblem prices out, the resulting new columns are added to the sets of schedules P and J, respectively, to be used in the next iteration. The objective values of all optimization problems in an iteration are further evaluated and used to calculate a Lagrangian bound for the algorithm. The iterative process stops, if no further columns with negative reduced costs were identified or if the difference between the RMP-LP objective and the Lagrangian dual bound is significantly small. Afterwards, the algorithm determines a valid solution for the original optimization problem by solving the RMP as an integer program (RMP-IP).

5. Computational study

The column generation was implemented as a Java project employing CPLEX 20.1.0 and has been run on a Windows 10 computer with 16 GB random access memory and an Intel® Xeon® Gold 2.3GHz quadcore processor. Some analyses of the dataset are summarized in section 5.1 to define data samples for the computational study. Section 5.2 points out the setup of the study and its instances. The outcomes are presented in section 5.3 regarding the performance of the methodology. Lastly, different scopes of flexibility are evaluated based on the results to provide managerial insights in section 5.4.

5.1. Data analyses and creation of data pool

The algorithm is tested based on a real-world dataset from a maximum care provider hospital in the south of Germany. The detailed analysis is pointed out in Appendix B. To create a data pool for our study, we aim to extract single schedule samples out of the dataset. Therefore, the following criteria are deduced and applied on the dataset:

- extract only schedules containing a surgical load between 6:45 hours to 10:00 hours,
- extract only schedules containing at least two surgeries,
- calculate the mean and the variation coefficient of surgery duration in each schedule,
- classify schedules to contain on average “short” surgeries, “medium lasting” surgeries, or “long” surgeries.

These steps result in a data pool consisting of more than 650 schedule samples.

5.2. Generation of instances

The main properties influencing flexibility in OR planning regard the expected duration and complexity of surgeries, the opening hours, and the possibility of room changes for both teams and surgeries. Therefore, we systematically want to align the instances with these drivers. To evaluate the surplus of room changes, the focus is given on a scope of three operating rooms. Figure 2 outlines the overall setup process of our instances: First, surgical programs are selected from the data pool as base cases.

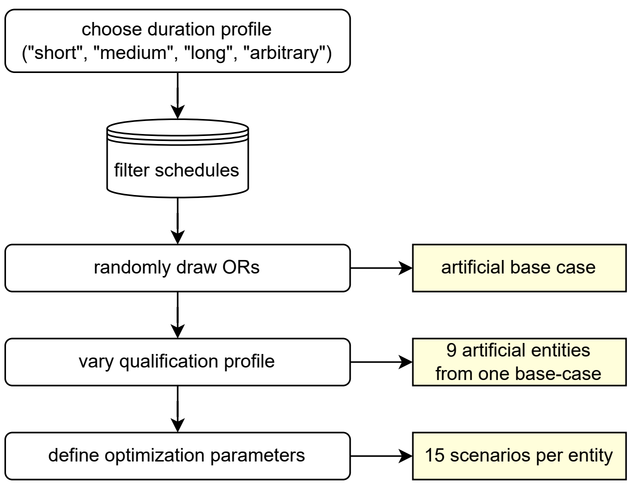
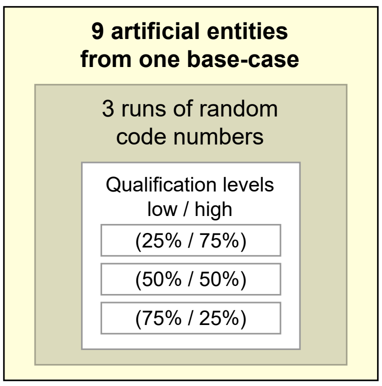


Figure 2: Process of generation of artificial instances

Second, the complexity of the surgeries is varied systematically to receive multiple entities out of one base case. Third, for each entity, scenarios with different setups of parameters are created. In the end, study instances are formed from the combination of entities and scenarios.

In a first step, surgical programs with different duration profiles are created. Depending on the choice, the schedule samples are filtered according to their labels as outlined in section B.5. Afterwards, three schedule realizations are randomly drawn from the filtered data pool. The resulting surgery program is considered as a base case. Within our base cases the number of patients varies between six patients up to 21 patients. The base case is varied within the following steps to create diversified instances.

In total, nine entities are created out of one base case and differ in the qualification profiles that the surgeries require. Figure 3a depicts the creation scheme of the nine entities, divided into three runs of applying three qualification settings each. More precisely, in each of the three runs, every surgery receives a random code number between 0 and 1. Then, for each run, the surgery codes are matched with three different qualification settings so that three entities are created in each of the runs. All of the entities generated are meant to be evaluated with respect to 15 scenarios as depicted in Figure 3b. The scenarios refer to the algorithm's parameter setting. They differ in the scope of fixations vs. optimization potentials for both surgeries and teams. The first set of scenarios applies the original sequence of the surgeries as given in the real-world dataset, i.e., the room and time where the surgery takes place is fixed. In the second set of scenarios, only the room where the surgery has to take place is fixed, whereas the starting time may be suitably set within a 10 hour time span. In the third set of scenarios, the surgeries do not underlie any fixations and can be placed to any room and any time within the 10h opening hours. For all of these sets, the parameter of maximum room changes per day for the teams is varied between 0 and 2. Scenario 4 and scenario 5 share similarities with the second and third scenario, respectively. However,



(a) Creation of entities out of a basement surgery program

15 scenarios per entity

	Opening hours 10h	Opening hours 11.5h
Original sequence (surgery room and time fixed)	1.0) staff w/o changes 1.1) staff max. 1 change 1.2) staff max. 2 changes	- - -
Only staff schedules (surgery room fixed)	2.0) staff w/o changes 2.1) staff max. 1 change 2.2) staff max. 2 changes	4.0) staff w/o changes 4.1) staff max. 1 change 4.2) staff max. 2 changes
Surgery and staff schedules	3.0) staff w/o changes 3.1) staff max. 1 change 3.2) staff max. 2 changes	5.0) staff w/o changes 5.1) staff max. 1 change 5.2) staff max. 2 changes

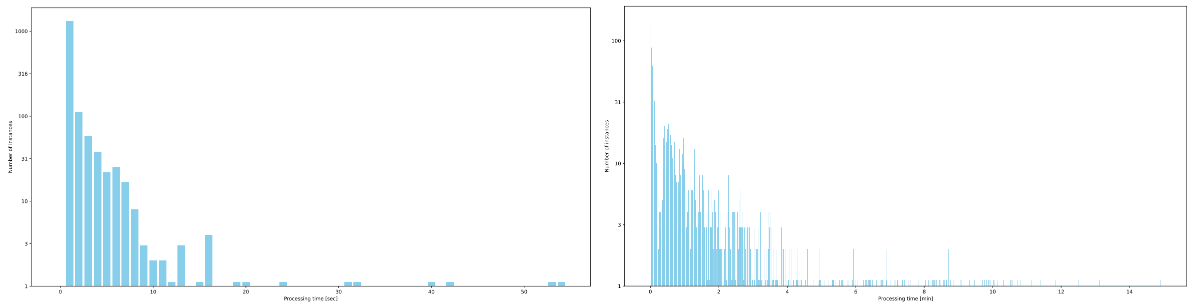
(b) Creation of optimization scenarios for entities

Figure 3: Overview on the setup of entities and scenarios

the scenarios 4 and 5 apply opening hours of 11.5 hours instead of 10 hours. For the overall computational study three separate base cases for each instance type are chosen. This results in twelve base cases which lead to 108 entities. In total, 1620 instances of the algorithm are evaluated and compared.

5.3. Computational performance

Running the 1620 instances with and without the consideration of an analytical lower bound, all of the runs terminated successfully. Both approaches led to gratifying results, however, the runs applying the lower bound from constraints 2.8 turned out to reach significantly better processing times. The distributions of processing times are represented in Figure 4a and Figure 4b for the runs with and without lower bound, respectively. A detailed description of the charts is given in Appendix C.



(a) including analytical lower bounds

(b) plain column generation without lower bounds

Figure 4: Distribution of solution times for 1620 instances

With the lower bound included, the minimum processing time takes 1.066 seconds and the maximum processing time takes 54.0 seconds. Without a lower bound, the processing times take between 1.09 seconds and 14:55.119 minutes, i.e., the x-axis of Figure 4b applies a wider scope. A detailed comparison shows that more than 90% of the instances terminate faster if the lower bound is included. This leads to an average time saving of 1:24.088 minutes and an average reduction of 180 iterations. For 84.9% of the instances, the final integer solutions (IP) with and without lower bound turn out to be identical. The average percentaged optimality gap between the final RMP solution (LP) and IP is 13.55% with and 64.42% without including the lower bound, respectively. However, this consideration is biased as our objective function is a weighted sum of an underlying ordinal scale. Comparing the values of the analytical lower bound (LB) against LP and IP for each instance, the observations from Figure 11a and Figure 11b are made. Accordingly, the column generation heuristic identifies optimal solutions for 76.1% of all instances with lower bound and for 84.6% of the instances without lower bound, respectively. The aggregated number of teams employed is identical between IP and LB in 94.9% of all instances regardless of the embedding of the lower bound into the solution process. Further details about the differences in the staff mix between IP and LB are presented in Table 5 in Appendix C. In summary, embedding the analytical lower bound into the column generation procedure leads to significant time reduction. The heuristic slightly loses in the quality of results but still reaches optimality in the majority of cases. Therefore, the focus is given to the algorithm's calculation with lower bounds in the following examinations.

5.4. Insights on planning different types of flexibility

A first impression about the outcomes for various duration profiles, qualifications profiles, and optimization scenarios is provided in Figure 5. The rows show results of instances per duration profile. The columns refer to instances with qualification profiles (75%/25%), (50%/50%), and (25%/75%), i.e., ratio of surgeries requiring *low/high* qualification. Each diagram depicts the average behaviour of instances of different optimization scenarios. The boxplots show the minimum, maximum, and mean number of low- and high-qualified teams over an average of nine instances each. Comparing the diagrams within one row from left to right, the single boxplots shrink in size and move towards the edges of the diagram when the treatments' complexity increases. Thus, an increasing ratio of complex surgeries requires more highly qualified teams and teams with lower qualification are replaced by highly qualified staff. A top to bottom comparison within one diagram shows that with an increasing scope of flexibility, the markers for mean values move towards the center of the diagram. Hence, increasing flexibility enables the replacement of highly qualified teams with lower qualified teams. This behaviour occurs independent from the duration profile as it can be observed in all rows of Figure 5. The results permit further evaluation regarding specific questions in the context of planning flexibility. Therefore, the results for different scenarios are compared and analyzed.

	opening hours 10h	**opening hours 11.5h**
completely fixed surgeries	• change beneficial in ~ 30% of entities • ∅ replace 1 high-qualified by low-qualified	
partly fixed surgeries	• change beneficial in ~ 9% of entities • ∅ replace 1 high-qualified by low-qualified	• change beneficial in ~ 11% of entities • ∅ replace 1 high-qualified by low-qualified
flexible surgeries	changes without impact	changes without impact

Table 2: Comparison of effects of room changing options for surgical teams

Is it beneficial for the planning process to allow room changes for the surgical teams? The surplus of flexibility for teams is assessed by comparing the results of the single scenarios per entity. For every kind of surgery fixation and opening hours, the scenarios differ in their setting of parameter $L^{\max}$. The comparison in Table 2 focuses on $L^{\max} = 0$ and $L^{\max} \geq 0$, since no significant differences can be observed between $L^{\max} = 1$ and $L^{\max} = 2$. The row label complete fixation of surgeries means that the surgery sequences stay as initialized, i.e., they cannot be rescheduled to another room or time period as given in scenarios 1.0, 1.1, and 1.2. A partly fixation means that surgeries can be rescheduled to another time period but need to stay in the same operating room. Scenarios 2.0, 2.1, and 2.2 apply this for 10h opening hours and scenarios 4.0, 4.1, and 4.2 for 11.5h opening hours. The term flexible surgeries means that surgeries may be rescheduled to any other room and time period. This is the case in scenarios 3.0, 3.1, and 3.2 for a 10h time span and in scenarios 5.0, 5.1, and 5.2 for a 11.5h time span. The evaluation shows that first, an increasing flexibility in surgery planning decreases the value of staff flexibility. With complete surgery fixation, about 30% of the 108 entities benefit from staff flexibility by replacing one highly qualified team by one lower qualified team on average. With a partial surgery fixation, only about 9% and 11% of the entities reach benefits from staff flexibility and if surgeries are fully flexible, the staff flexibility has no more impact. Second, expanded opening hours slightly increase the value of staff flexibility. This is shown for the partial fixation of surgeries where staff flexibility leads to benefits for about 9% of the entities if opening hours of 10h are applied and for about 11% of the entities if longer opening hours are applied.

	opening hours 10h	**opening hours 11.5h**
staff without room changes	• change beneficial in $\sim$ 68% of entities • ∅ replace 1.2 high-qualified by low-qualified	• change beneficial in $\sim$ 78% of entities • ∅ replace 1.2 high-qualified by low-qualified
staff with room changes	• change beneficial in $\sim$ 60% of entities • ∅ replace 1 high-qualified by low-qualified	• change beneficial in $\sim$ 77% of entities • ∅ replace 1 high-qualified by low-qualified

Table 3: Comparison of effects of room changing options for surgical interventions

Is it beneficial for the planning process to allow room changes for the surgeries? For this analysis, the results of scenarios per entity are compared for the settings $L^{\max} = 0$, $L^{\max} \geq 0$, and for different opening hours as stated in Table 3. It can be observed that an increasing flexibility in staff planning slightly decreases the value of surgery flexibility. Without room changes for the teams, about 68% of the 108 entities benefit from flexible shifting of surgeries in 10h opening hours and about 78% of the entities with longer opening hours. For these entities it is possible to replace on average 1.2 highly qualified teams by lower qualified teams. If teams are allowed to change rooms, these values slightly decrease to about 60% and about 77% and a fewer average for replacing highly qualified by lower qualified teams. These results demonstrate that more entities can profit by surgery flexibility if the opening hours are enlarged. In other words, expanded opening hours increase the value of surgery flexibility.

	partly fixed surgeries	**flexible surgeries**
staff without room changes	extension without impact	• change beneficial in ∼ 21% of entities • ∅ replace 0.5 high-qualified by low-qualified
staff with room changes	• change beneficial in ∼ 8% of entities • ∅ replace 1 high-qualified by low-qualified	• change beneficial in ∼ 21% of entities • ∅ replace 0.5 high-qualified by low-qualified

Table 4: Comparison of effects of expanding the opening hours

Is it profitable to allow expanded opening hours which may provide the opportunity to better distribute treatments? The last comparison of scenarios per entity regards the evaluation of time-related impact. Therefore, the settings $L^{\max} = 0$ and $L^{\max} \geq 0$ and the fixation vs. flexibility of surgeries is compared as given in Table 4. If no flexibility is allowed for either surgeries or surgical teams, also the extension of the opening hours has no impact, i.e., at least one of the other drivers needs to be flexible to reach benefits. Comparing the results of Table 4 vertically and horizontally, there is no surplus of enabling room changes for teams if surgeries may be placed flexible. If surgeries may be planned flexible about 21% of the entities benefit from flexibility and can replace an average of 0.5 highly qualified teams with lower qualified teams. This effect is not influenced by the degree of staff flexibility. However, if surgeries are partly fixed, i.e., are only allowed to be shifted in time but are not allowed to be placed in other rooms, staff flexibility reaches a benefit for about 8% of the entities by replacing an average of one highly qualified team with a lower qualified team. In summary, it can be concluded that increasing surgery flexibility also increases the value of expanding the opening hours. The impact of staff flexibility on the value of expanded opening hours can be supposed but cannot be clearly validated from the experimental results.

Which situations make room changing possibilities of surgeries or teams beneficial? Combining the results of the former evaluations, it can be concluded that surgery flexibility is more useful if teams are fixed and vice versa. Grouping surgeries of similar complexity so that all of them are conducted in the same room is especially useful if room changes for the surgical teams are not allowed. This is verified as surgery flexibility is more useful if staff members cannot change their rooms and reaches its highest impact if the opening hours are enlarged. The conclusion is further confirmed from the observation that staff flexibility is of no surplus if surgeries may be placed flexibly. Since enlarged opening hours profit more from surgery flexibility than from staff flexibility and staff flexibility becomes dispensable once surgeries may be shifted between rooms, it can be inferred that planning surgery flexibility has higher impact in efficient scheduling of highly qualified teams than extra staff flexibility.

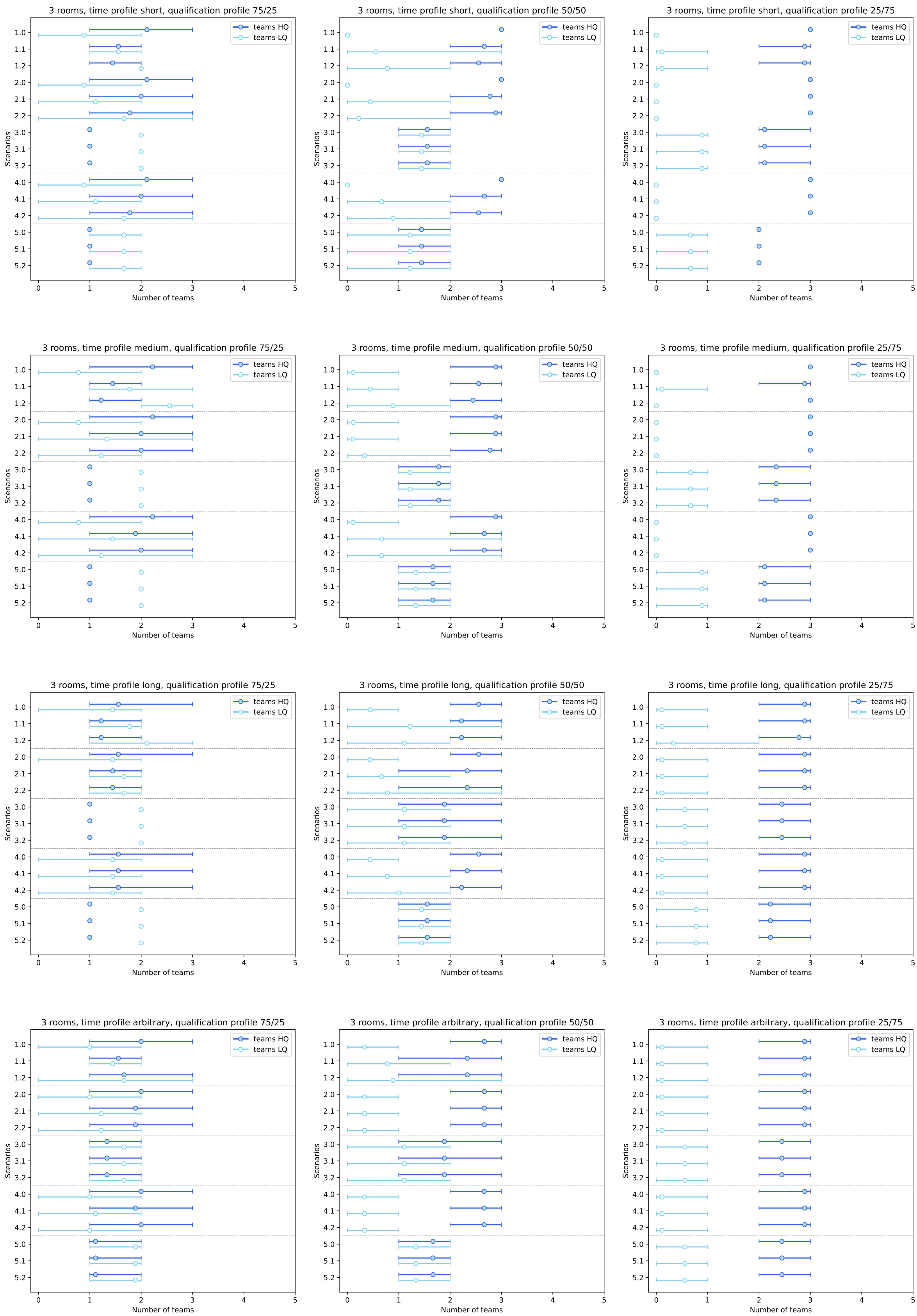


Figure 5: Boxplots of average minimum, maximum, and mean of required teams

6. Conclusion

In this paper, we provide a MIP model for simultaneous allocation scheduling and surgical team assignment. Our goal is to evaluate the surplus of enabling the planning of flexibility for different factors of this problem. In order to solve the optimization problem efficiently, we decompose the MIP and provide respective model reformulations. The reformulation addresses as well symmetry issues as an option for process speed-up by consideration of wise lower bounds. We embed the models within a column generation procedure and generate a multitude of test instances based on real-world surgical schedules. The solution approach reaches optimal solutions for 76.1% of our experimental instances and is characterized by a desirable runtime from both a theoretical and a practical point of view.

Analyzing the value of planning flexible, we compare a variety of scenarios. The focus of our evaluations lies on room changing opportunities for surgical teams, room shifts for surgeries and an expansion of the OR opening hours. Comparing the average behaviour in different scenarios across all our instances, we present practical insights and sum them up as key findings. An expansion of opening hours increases both the benefits from flexible surgery planning and flexible staff planning. However, contrasting staff flexibility with surgery flexibility, the latter turns out to be of higher impact on efficient scheduling in the integrated OR planning problem.

Further enhancements of our approach in future research could regard model extensions as well as methodological aspects. For a better fit towards the practical setting, there would be options to include break assignments for surgical teams or to consider downstream units capacities. Regarding the surgery scheduling process, the results could be refined by the consideration of stochastic durations of the interventions. From a methodological point of view, our solution approach offers the possibility to reformulate the staff subproblem to formulations that might reduce their computational complexity like e.g., a shortest path problem formulation. With some further adjustments also the embedding into a branch-and-price solution approach is possible.

References

[1] Akbarzadeh, Babak, Moslehi, Ghasem, Reisi-Nafchi, Mohammad and Maenhout, Broos: A diving heuristic for planning and scheduling surgical cases in the operating room department with nurse re-rostering. In: *Journal of Scheduling*, Volume 23(2):S. 265–288, 2020. ISSN 1094-6136. doi:10.1007/s10951-020-00639-6.

[2] Askarizadeh, Mehdi and Aziz, Faieza Abdul: Planning and scheduling operating rooms with mix integer programming and meta-heuristic method. In: *Journal of Computational and Theoretical Nanoscience*, Volume 15(3):S. 961–965, 2018. doi:10.1166/jctn.2018.7182.

[3] Avgerinos, Emmanouil and Gokpinar, Bilal: Team familiarity and productivity in cardiac surgery operations: The effect of dispersion, bottlenecks, and task complexity. In: *Manufacturing & Service Operations Management*, Volume 19(1):S. 19–35, 2017. doi:10.1287/msom.2016.0597.

[4] Bandi, Chaithanya and Gupta, Diwakar: Operating room staffing and scheduling. In: *Manufacturing & Service Operations Management*, Volume 22(5):S. 958–974, 2020. doi:10.1287/msom.2019.0781.

[5] Beliën, Jeroen and Demeulemeester, Erik: A branch-and-price approach for integrating nurse and surgery scheduling. In: *European Journal of Operational Research*, Volume 189(3):S. 652–668, 2008. ISSN 03772217. doi:10.1016/j.ejor.2006.10.060.

[6] Belkhamsa, Manel, Jarboui, Bassem and Masmoudi, Malek: Two metaheuristics for solving no-wait operating room surgery scheduling problem under various resource constraints. In: *Computers & Industrial Engineering*, Volume 126:S. 494–506, 2018. ISSN 0360-8352. doi:10.1016/j.cie.2018.10.017.

[7] Benazzouz, Touria, Echchatbi, Abdelwahed and Bellabdaoui, Adil: A literature review on the nurses' planning problems. In: *International Journal of Mathematics and Computational Science*, Volume 1:S. 268–274, 2015.

[8] Burke, Edmund K., de Causmaecker, Patrick, Berghe, Greet Vanden and van Landeghem, Hendrik: The state of the art of nurse rostering. In: *Journal of Scheduling*, Volume 7(6):S. 441–499, 2004. ISSN 1094-6136. doi:10.1023/B:JOSH.0000046076.75950.0b.

[9] Calegari, Rafael, Fogliatto, Flavio S., Lucini, Filipe R., Anzanello, Michel J. and Schaan, Beatriz D.: Surgery scheduling heuristic considering or downstream and upstream facilities and resources. In: *BMC health services research*, Volume 20(1):S. 684, 2020. doi:10.1186/s12913-020-05555-1.

[10] Erhard, Melanie Johanna Patricia, Schoenfelder, Jan, Fügener, Andreas and Brunner, Jens Otto: State of the art in physician scheduling. In: *European Journal of Operational Research*, Volume 265(1):S. 1–18, 2018. ISSN 03772217. doi:10.1016/j.ejor.2017.06.037.

[11] Guerriero, Francesca and Guido, Rosita: Operational research in the management of the operating theatre: A survey. In: *Health Care Management Science*, Volume 14(1):S. 89–114, 2011. ISSN 1572-9389. doi:10.1007/s10729-010-9143-6.

[12] Guido, Rosita and Conforti, Domenico: A hybrid genetic approach for solving an integrated multi-objective operating room planning and scheduling problem. In: *Computers & Operations Research*, Volume 87:S. 270–282, 2017. ISSN 03050548. doi:10.1016/j.cor.2016.11.009.

[13] Guinet, Alain and Chaabane, Sondes: Operating theatre planning. In: *International Journal of Production Economics*, Volume 85(1):S. 69–81, 2003. doi:10.1016/S0925-5273(03)00087-2.

[14] Gür, Şeyda and Eren, Tamer: Application of operational research techniques in operating room scheduling problems: Literature overview. In: *Journal of healthcare engineering*, Volume 2018:S. 5341394, 2018. ISSN 2040-2295. doi:10.1155/2018/5341394.

[15] Harris, Sean and Claudio, David: Current trends in operating room scheduling 2015 to 2020: a literature review. In: *Operations Research Forum*, Volume 3(1), 2022. doi:10.1007/s43069-022-00134-y.

[16] Hashemi Doulabi, Seyed Hossein, Rousseau, Louis-Martin and Pesant, Gilles: A constraint programming-based column generation approach for operating room planning and scheduling. In: Simonis, Helmut (Ed.) *Integration of AI and OR Techniques in Constraint Programming*. Springer International Publishing, Cham, 2014, Volume 8451, S. 455–463. doi:10.1007/978-3-319-07046-9{\textunderscore}32.

[17] Hashemi Doulabi, Seyed Hossein, Rousseau, Louis-Martin and Pesant, Gilles: A constraint-programming-based branch-and-price-and-cut approach for operating room planning and scheduling. In: *INFORMS Journal on Computing*, Volume 28(3):S. 432–448, 2016. ISSN 1091-9856. doi:10.1287/ijoc.2015.0686.

[18] Hof, Sebastian, Fügener, Andreas, Schoenfelder, Jan and Brunner, Jens Otto: Case mix planning in hospitals: a review and future agenda. In: *Health Care Management Science*, Volume 20(2):S. 207–220, 2017. ISSN 1572-9389. doi:10.1007/s10729-015-9342-2.

[19] Latorre-Núñez, Guillermo, Lüer-Villagra, Armin, Marianov, Vladimir, Obreque, Carlos, Ramis, Francisco and Neriz, Liliana: Scheduling operating rooms with consideration of all resources, post anesthesia beds and emergency surgeries. In: *Computers & Industrial Engineering*, Volume 97:S. 248–257, 2016. ISSN 0360-8352. doi:10.1016/j.cie.2016.05.016.

[20] Lim, Gino and Mobasher, Arezou: Operating suite nurse scheduling problem: A column generation approach. In: *IIE Annual Conference Proceedings*, S. 1–10, 2012. URL `https://search.proquest.com/docview/1151086099?accountid=8429`.

[21] Mobasher, Arezou, Lim, Gino, Bard, Jonathan F. and Jordan, Victoria: Daily scheduling of nurses in operating suites. In: *IIE Transactions on Healthcare Systems Engineering*, Volume 1(4):S. 232–246, 2011. ISSN 1948-8300. doi:10.1080/19488300.2011.631097.

[22] Naderi, Bahman, Roshanaei, Vahid, Begen, Mehmet A., Aleman, Dionne M. and Urbach, David R.: Increased surgical capacity without additional resources: Generalized operating room planning and scheduling. In: *Production and Operations Management*, Volume 30(8):S. 2608–2635, 2021. doi:10.1111/poms.13397.

[23] Ngoo, Chong Man, Goh, Say Leng, Sze, San Nah, Sabar, Nasser R., Abdullah, Salwani and Kendall, Graham: A survey of the nurse rostering solution methodologies: The state-of-the-art and emerging trends. In: *IEEE Access*, Volume 10:S. 56504–56524, 2022. doi:10.1109/ACCESS.2022.3177280.

[24] Pham, Dinh-Nguyen and Klinkert, Andreas: Surgical case scheduling as a generalized job shop scheduling problem. In: *European Journal of Operational Research*, Volume 185(3):S. 1011–1025, 2008. ISSN 03772217. doi:10.1016/j.ejor.2006.03.059.

[25] Rachuba, Sebastian, Reuter-Oppermann, Melanie and Thielen, Clemens: Integrated planning in hospitals: A review. In: *arXiv*, 2023. doi:10.48550/arXiv.2307.05258.

[26] Rahimi, Iman and Gandomi, Amir H.: A comprehensive review and analysis of operating room and surgery scheduling. In: *Archives of Computational Methods in Engineering*, 2020. doi:10.1007/s11831-020-09432-2.

[27] Rath, Sandeep, Rajaram, Kumar and Mahajan, Aman: Integrated anesthesiologist and room scheduling for surgeries: Methodology and application. In: *Operations Research*, Volume 65(6):S. 1460–1478, 2017. ISSN 0030-364X. doi:10.1287/opre.2017.1634.

[28] Razali, Mohamad Khairulamirin Md, Rahman, Abdul Hadi Abd, Ayob, Masri, Jarmin, Razman, Qamar, Faizan and Kendall, Graham: Research trends in the op-

timization of the master surgery scheduling problem. In: *IEEE Access*, Volume 10:S. 91466–91480, 2022. doi:10.1109/ACCESS.2022.3202546.

[29] Riise, Atle, Mannino, Carlo and Burke, Edmund K.: Modelling and solving generalised operational surgery scheduling problems. In: *Computers & Operations Research*, Volume 66:S. 1–11, 2016. ISSN 03050548. doi:10.1016/j.cor.2015.07.003.

[30] Samudra, Michael, van Riet, Carla, Demeulemeester, Erik, Cardoen, Brecht, Vansteenkiste, Nancy and Rademakers, Frank E.: Scheduling operating rooms: achievements, challenges and pitfalls. In: *Journal of Scheduling*, Volume 19(5):S. 493–525, 2016. ISSN 1094-6136. doi:10.1007/s10951-016-0489-6.

[31] Silva, Thiago A.O., de Souza, Mauricio C., Saldanha, Rodney R. and Burke, Edmund K.: Surgical scheduling with simultaneous employment of specialised human resources. In: *European Journal of Operational Research*, Volume 245(3):S. 719–730, 2015. ISSN 03772217. doi:10.1016/j.ejor.2015.04.008.

[32] Tsang, Man Yiu, Shehadeh, Karmel S., Curtis, Frank E., Hochman, Beth R. and Brentjens, Tricia E.: Stochastic optimization approaches for an operating room and anesthesiologist scheduling problem. In: *Operations Research*, 2024. ISSN 0030-364X. doi:10.1287/opre.2022.0258. URL `https://arxiv.org/abs/2204.11374`.

[33] van Huele, Christophe and Vanhoucke, Mario: Decomposition-based heuristics for the integrated physician rostering and surgery scheduling problem. In: *Health Systems*, Volume 4(3):S. 159–175, 2015. ISSN 2047-6965. doi:10.1057/hs.2014.27.

[34] Villarreal, Monica C. and Keskinocak, Pinar: Staff planning for operating rooms with different surgical services lines. In: *Health Care Management Science*, Volume 19(2):S. 144–169, 2016. ISSN 1572-9389. doi:10.1007/s10729-014-9307-x.

[35] Volland, Jonas, Fügener, Andreas and Brunner, Jens Otto: A column generation approach for the integrated shift and task scheduling problem of logistics assistants in hospitals. In: *European Journal of Operational Research*, Volume 260(1):S. 316–334, 2017. ISSN 03772217. doi:10.1016/j.ejor.2016.12.026.

[36] Xiang, Wei, Yin, Jiao and Lim, Gino: An ant colony optimization approach for solving an operating room surgery scheduling problem. In: *Computers & Industrial Engineering*, Volume 85:S. 335–345, 2015. ISSN 0360-8352. doi:10.1016/j.cie.2015.04.010.

[37] Zhu, Shuwan, Fan, Wenjuan, Yang, Shanlin, Pei, Jun and Pardalos, Panos M.: Operating room planning and surgical case scheduling: a review of literature. In: *Journal of Combinatorial Optimization*, Volume 37(3):S. 757–805, 2019. doi:10.1007/s10878-018-0322-6.

A. Models of subproblems

Subsequently, the models of the subproblems from section 4.2 are given in detail. All of their constraints are outlined and related to the generic MIP model.

A.1. Task-sequencing subproblem

The model applies the same sets, parameters, and decision variables as the model presented in section 3 and in addition the dual variables δ_{qlt} and μ from the RMP.

$$\bar{C}_p = \min \quad \sum_{s\in S} C^{\text{miss}} \cdot D_s \cdot x_s^{\text{miss}} + \sum_{q\in Q}\sum_{l\in L}\sum_{t\in T} \delta_{qlt} \cdot n_{qlt} - \mu \tag{3.1}$$

$$\text{s.t.} \quad \sum_{l\in L}\sum_{t\in T} x_{slt}^{\text{start}} + x_s^{\text{miss}} = 1, \qquad \forall s \in S, \tag{3.2}$$

$$\sum_{s\in S} x_{slt}^{\text{on}} \leq 1, \qquad \forall l \in L, t \in T, \tag{3.3}$$

$$\sum_{\zeta=\max(1,t-D_s+1)}^{t} x_{sl\zeta}^{\text{start}} = x_{slt}^{\text{on}}, \qquad \forall s \in S, l \in L, t \in T, \tag{3.4}$$

$$\sum_{s\in S} H_{sq} \cdot x_{slt}^{\text{on}} = n_{qlt}, \qquad \forall q \in Q, l \in L, t \in T, \tag{3.5}$$

$$\sum_{s\in S}\sum_{l\in L}\sum_{t\in T} x_{slt}^{\text{start}} \geq 1, \tag{3.6}$$

$$x_s^{\text{miss}} \in \mathbb{B}, \qquad \forall s \in S, \tag{3.7}$$

$$x_{slt}^{\text{start}} \leq \hat{X}_{slt}^{\text{start}}, x_{slt}^{\text{start}}, x_{olt}^{\text{on}} \in \mathbb{B}, \quad \forall s \in S, l \in L, t \in T, \tag{3.8}$$

$$n_{qlt} \in \mathbb{N}_0, \qquad \forall q \in Q, l \in L, t \in T \tag{3.9}$$

Constraints (3.2) to (3.5) correspond to the constraints (1.3) to (1.6) of the MIP model. The same holds true for the constraints (3.7), (3.8), and (3.9) which correspond to the constraints (1.14), (1.15), and (1.19). Constraint (3.6) was added to ensure that no empty columns are created.

A.2. Task-assignment subproblem

This model also applies the same notation as presented in section 3. An exception is index w which is replaced by g as explained in section 4. The model uses the dual variables δ_{qlt}, π_g, and τ_q from the RMP. Since the subproblem is solved per group, the value of g is fixed and the index moves into the superscript of parameters and variables.

$$\bar{C}_j^g = \min \quad C^g - \sum_{q\in Q}\sum_{l\in L}\sum_{t\in T} \delta_{qlt} \cdot a_{qlt}^g - \pi^g - \sum_{q\in Q: q \leq q_g} \tau_q \tag{4.1}$$

$$\text{s.t.} \quad \sum_{l\in L} y_{lt}^{\text{on},g} \leq 1, \qquad \forall t \in T, \tag{4.2}$$

$$y_{lt}^{\text{start},g} \cdot T^{\min} \leq \sum_{\zeta=t}^{t+T^{\min}-1} y_{l\zeta}^{\text{on},g}, \quad \forall l \in L, t \in \{1, \dots, |T| - T^{\min} + 1\}, \tag{4.3}$$

$$\sum_{l\in L}\sum_{t\in T} y_{lt}^{\text{start},g} \leq 1 + L^{\max}, \tag{4.4}$$

$$y_{lt}^{\text{on},g} \cdot \hat{A}_{qlt}^g = a_{qlt}^g, \qquad \forall q \in Q, l \in L, t \in T, \tag{4.5}$$

$$y_{lt}^{\text{on},g} - y_{l,t-1}^{\text{on},g} \leq y_{lt}^{\text{start},g}, \qquad \forall l \in L, t \in T \setminus \{1\}, \tag{4.6}$$

$$y_{lt}^{\text{on},g} \leq y_{lt}^{\text{start},g}, \qquad \forall l \in L, t = 1, \tag{4.7}$$

$$y_{lt}^{\text{start},g}, y_{lt}^{\text{on},g} \in \mathbb{B}, \qquad \forall l \in L, t \in T, \tag{4.8}$$

$$a_{qlt}^g \in \mathbb{B} \qquad \forall q \in Q, l \in L, t \in T \tag{4.9}$$

Constraints (4.2) to (4.7) corresponds to constraints (1.8) to (1.13) of the MIP, respectively. The same holds for constraints (4.8) to (4.9) which are the TASP equivalent to constraints (1.17) to (1.18) of the MIP.

B. Insights of the real-world data analyses

This appendix section points out the analyses of a real-world dataset originating from a maximum care provider hospital in the south of Germany. The dataset comprises information from the central surgery department over a three month time horizon. The raw data got extracted and analysed in order to identify surgery characteristics. In this context, we focus on single realizations of OR occupancies, i.e., we assess each schedule in each room and at each day separately without consideration of medical specialties or dates. Afterwards, the information is applied to filter and label the raw-data and to create data samples.

B.1. Opening hours

In order to determine the usual opening hours of the OR department, the starting times of surgeries are grouped into one hour time slots and counted. Plotting the results, Figure 6 shows the number of surgeries starting at different times of a day. The x-axis gives the hours of the day, from 5 AM to 10 PM applying a 24 hour time format. The y-axis gives the number of surgeries that are starting within the respective time window.

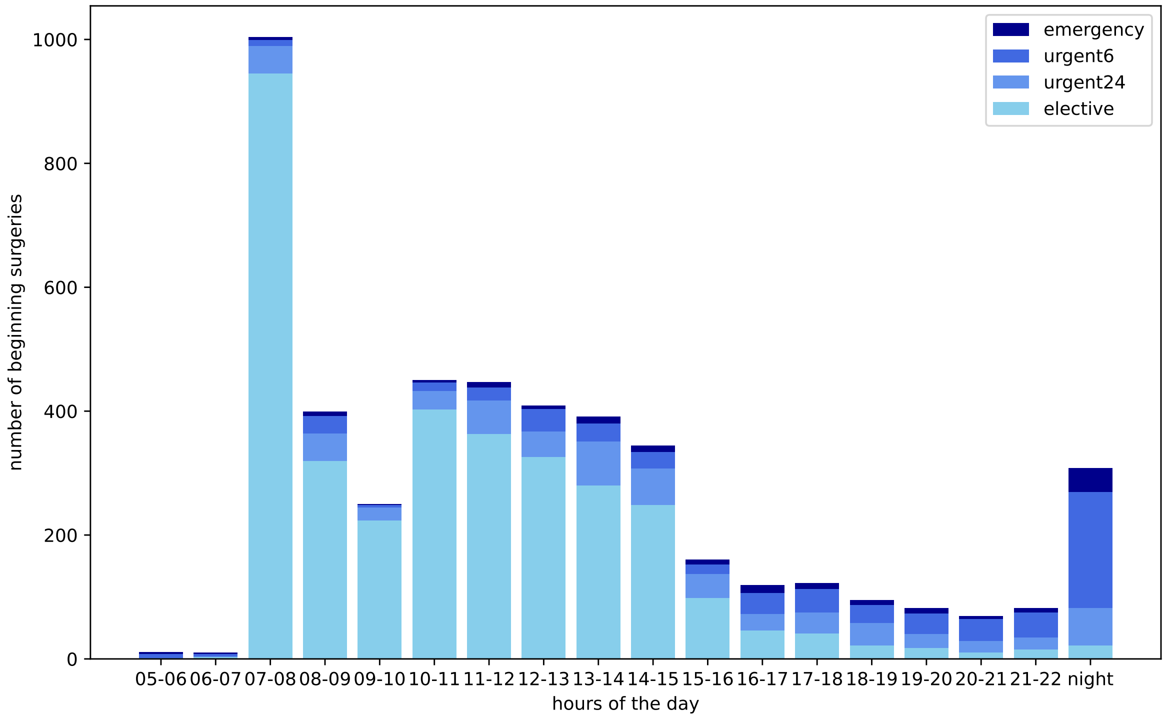


Figure 6: Number of surgery starting times differentiated by urgency

The stacked bars represent the number of starting surgeries respecting their urgencies. "Elective" means patients that have a regular assignment for their surgeries and were known multiple days or weeks before the surgery date. The chart shows that this kind of surgery is the majority of all surgical cases and usually starts in the morning, forenoon, or in the early afternoon. By the end of the day their occurrence shrinks. The label "urgent24" means patients whose treatment is urgent but not life-threatening so that they should undergo surgery within the next 24 hours. These patients can be said to be known one day before the surgery. Therefore, they have similar starting times as elective surgeries, however, they also occur during the night if they cannot wait until the next morning. Similarly, the label "urgent6" comprises surgeries which are more urgent and should undergo surgery within the next six hours. Finally, the label "emergency" addresses patients whose health condition is life-threatening and they need to undergo surgery as soon as possible. The two latest categories are processed throughout the whole day and represent a majority of the cases in the late evening or during the night.

Recapitulating this information, the data demonstrates that about 95 % of elective surgeries are starting between 7:30 AM and 4:30 PM. Considering all treatments regardless of their urgency, about 82 % of all surgeries are starting within this time window. Due to this, a nine hour time horizon is used for further analyses. This includes the assumption that all rooms of the OR unit apply identical opening hours.

B.2. Operating room utilization

As a next step, the typical load of ORs is determined. From Figure 6 we deduce usual opening hours of nine hours and calculate the utilization of operating theatres based on this time span. Therefore, we focus on single schedules per room and day, as mentioned before. For each operating room, the sum of it's surgical time per day gets summed up and the total amount of time for each room and day is divided by a nine hour time span which leads to results between 0.0 (exclusively) and 1.7.

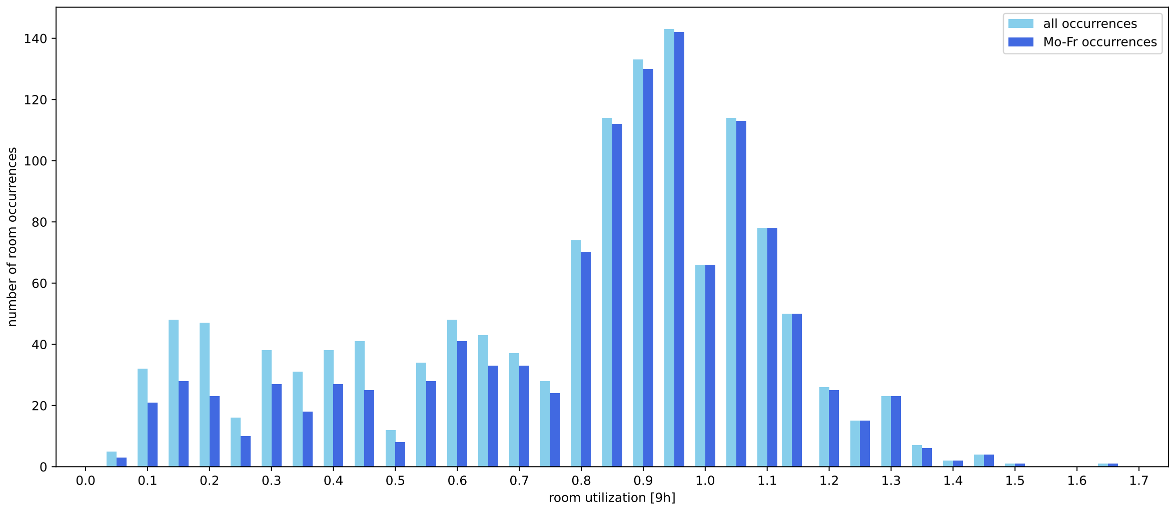


Figure 7: Operating room utilization based on nine hour opening hours

The x-axis in Figure 7 depicts different levels of utilization and the y-axis counts how often the respective utilization occurs within the dataset. In this context, e.g., the bar labeled 0.8 means that the utilization of a room counted within this bar is in the interval $(0.8, 0.85]$. The most frequent utilization rates show up to be between 0.8 and 1.1. This refers to surgical times between 7:15 hours to 10:20 hours, which comprises more than 55% of the active operating theatres. Another 10% of the active operating theatres are allocated between 10:20 hours to 11:15 hours. For high utilization rates, the overall sum of occurences does not differ significantly from the occurences of only weekdays (Mo-Fr). However, for low utilization rates, the quantity shows greater differences which indicates that a low utilization of an operating theatre is more likely at weekend days. This observation comes along with the common practice that only urgent and emergency surgeries are conducted during the weekend or holidays.

B.3. Daily room capacity per medical specialty

Most commonly, a medical specialty is given one operating theatre per day where they conduct surgeries. However, there are also medical specialties who are allowed to occupy more than one room per day.

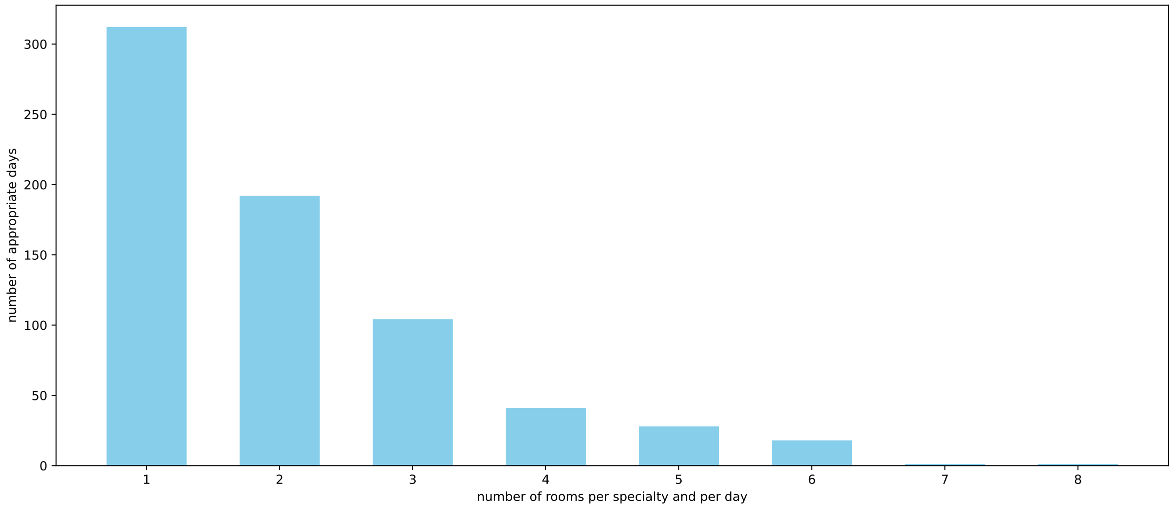


Figure 8: Frequency of number of rooms per specialty and day

Figure 8 shows the number of rooms a specialty is occupying per day. The x-axis depicts the number of operating theatres and the y-axis gives the frequency how often, i.e., at how many days, a specialty uses the respective number of operating theatres per day. As shown in the chart, a number of more than six operating rooms for one medical specialty at one day is a rare exception. Overall, in more than 85% of all cases a specialty uses at most three operating rooms per day.

B.4. Surgery duration

To address surgeries adequately, the aim is to classify surgery duration by means of the labels "short", "medium", and "long". Therefore, Figure 9 shows aggregated information about surgery duration. The duration in hours is given on the x-axis with a maximum value at 12.5. More specifically, the bar at position 1h counts all surgeries with a duration between 1:00 (exclusively) to 1:15. The primary y-axis depicts the number of surgeries which have the respective duration, represented by the bars. The secondary y-axis points out the accumulated ratio of surgeries that have a lower or equal duration to the value on the x-axis. This is represented by the lines.

In a first step, it is necessary to determine an upper bound for surgery duration to be considered within the analyses. The maximum surgery duration within the dataset took a total of 12:30 hours. However, as shown in Figure 9, the majority of surgeries take a less amount of time. According to the lower blue line, 94 % of all the surgeries are

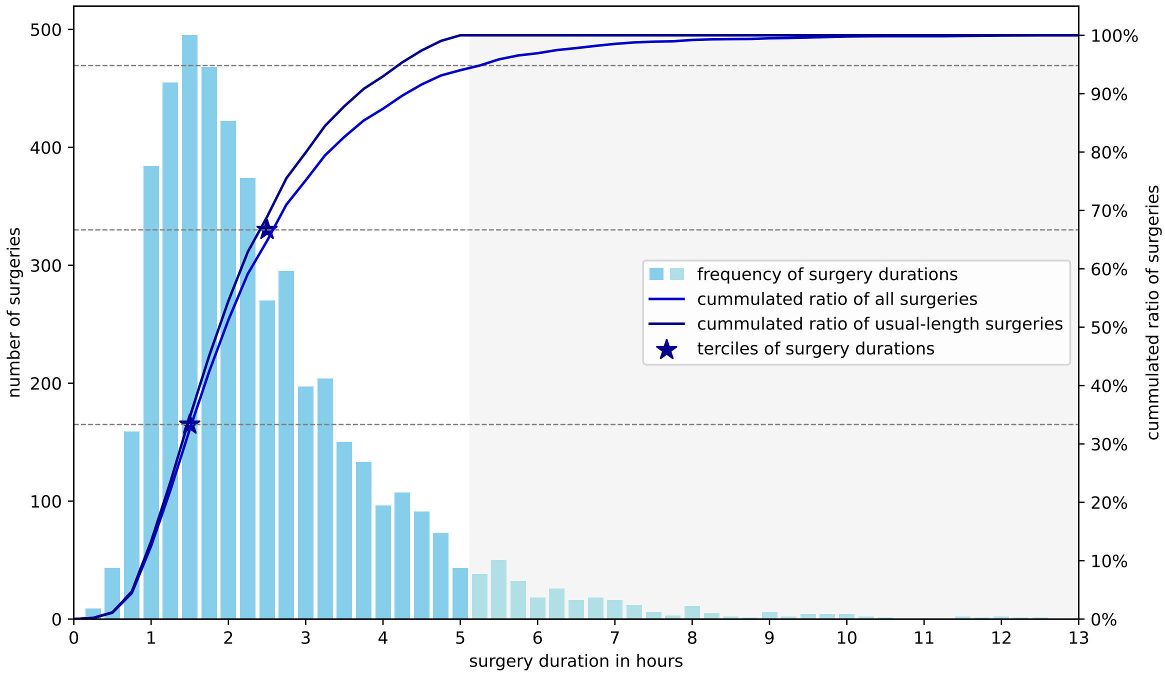


Figure 9: Number of surgeries related to their duration

shorter than five hours. Thus, five hours are used and marked as a cutting point, so that surgeries taking longer than five hours are ignored in further inspections.

The darker blue line shows the accumulated ratio of surgeries for only those treatments with a duration of less than five hours. For the remaining surgeries, the terciles of their time consumption are calculated. The terciles are marked by two star symbols and match exactly the two lines of accumulated ratios. Finally, interpreting the terciles leads to the result that surgeries shorter than 1:30 hours are labeled as short surgeries. Surgeries between 1:30 hour and 2:30 hours are labeled as medium, and surgeries which take longer than 2:30 hours are said to have a long duration.

B.5. Setup of data pool and samples

In order to create an adequate data pool for the construction of artificial instances, the raw data of the real-world dataset is filtered to provide useful samples. This means that the single realizations of operating theatre allocations are filtered, assessed, and labeled due to the surgical program they contain.

First, only schedules with a utilization between 75% and 110%, i.e., a runtime between 6:45 hours to 10:00 hours, remain in the sample set. Second, schedules need to contain at least two surgeries. These steps ensure that the evaluation of different drivers of flexibility won't be biased from underutilization or infeasability. Finally, a set of more than 650 samples is used as a basis for the further instance creation process.

Third, all the remaining samples are labeled regarding the duration of their treatments. Therefore, for each schedule, the average duration of it's surgical program is calculated

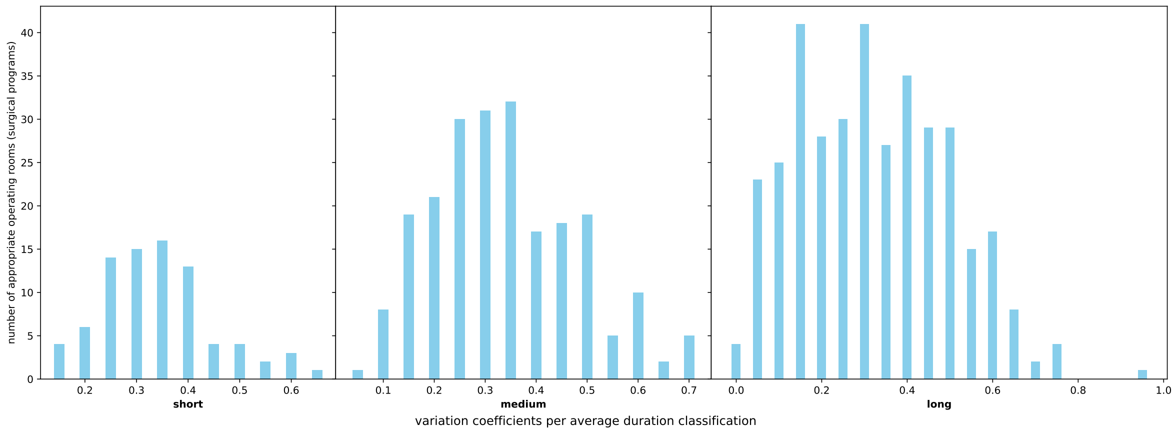


Figure 10: Frequency of variation coefficients for different surgery duration's classification

and labeled as "short", "medium", or "long". Based on the mean duration, also the variance, standard deviation, and the variation coefficient for the surgeries in this schedule are calculated. Figure 10 represents how many rooms of each duration classification occur with respect to different variation coefficients. The x-axis depicts the different duration classifications and variation coefficients. In this context, a label of e.g. 0.3 states that the variation coefficient of the rooms counted by the respective bar is within the interval $(0.3, 0.35]$. It can be observed that for rooms labeled with longer duration there occurs more variability in surgery duration.

In the end, this pool of real-world samples provides the opportunity to arbitrarily combine schedule realizations as new artificial instances. Moreover, the labeling provides the option to e.g. create new instances with mainly short surgeries. For this example, it would be appropriate to filter only those schedule samples that are addressed in the section of "short" average duration and have a variation coefficient of at most $CV \leq 0.4$.

C. Plots of the column generation results

The following presents the outcomes of 3240 runs of the column generation heuristic in a graphical format. We focus on the comparison of executing the algorithm with and without consideration of the problem-specific lower bound. I.e., first, the 1620 instances have been executed with consideration of the lower bound, and afterwards, they have been run in a plain column generation without additional lower bound.

C.1. Solution times

All of the 1620 instances terminated within a reasonable amount of time regardless of whether the lower bound is included or not. To assess the performance, we count the frequency of solution time realizations for time intervals of one second. The following barcharts show the distribution of the algorithm's processing time with and without consideration of pre-calculated analytical lower bounds. Note, that the y-axes apply a logarithmic scale. Moreover, the two charts apply a different scale for their x-axes.

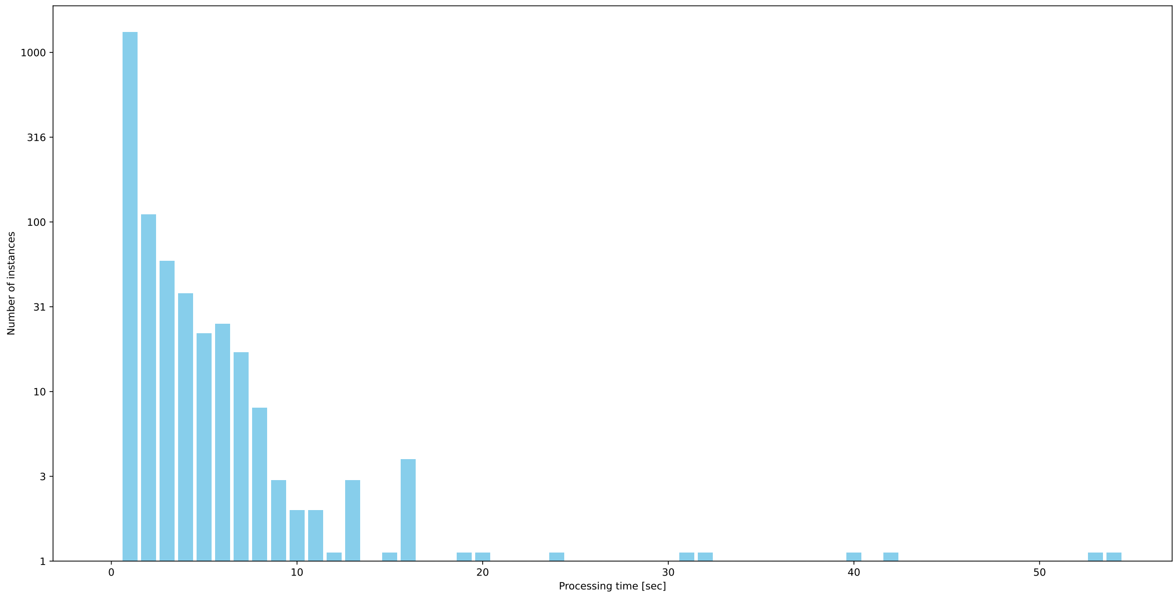


Figure 4a: Distribution of processing times including lower bounds

Figure 4a shows the distribution of processing times with the lower bound included into the column generation heuristic. The runtime ranges between a minimum of 1.066 seconds and a maximum of 54.0 seconds. Hence, the x-axis of Figure 4a depicts a scope up to one minute. The bar at position 1 sec. counts all instances that showed a processing time in the interval $[1\text{sec}, 2\text{sec})$. As the chart shows, this is the case for the majority of instances. The median takes a value of 1.362 seconds and the average runtime is 2.115 seconds. Overall, including the lower bound, more than 81% of the instances terminate in less than two seconds and more than 98% of the instances terminate in less than ten seconds.

Figure 4b shows the distribution of processing times of the plain column generation heuristic without a lower bound. The runtime ranges between 1.09 seconds and 14:55.119 minutes. The bar notation is identical to the definition before, however, the x-axis is enlarged due to longer processing times. The median takes a value of 36.296 seconds and the average time is 01:18.182 minutes. About 33% of the instances still terminate in less than ten seconds. More than 63% of the instances terminate in less than one minute and more than 92% terminate in less than four minutes.

Contrasting the processing times for each instance, it turns out that more than 86% of the instances loop through less iterations if the lower bound is included, further 7% deploy the identical number of iterations. This leads to more than 90% of the instances which benefit from reduced time consumption if the lower bound is considered. On average, the instances can save a time amount of 1:24.088 minutes processing time and the calculatory steps are reduced by a mean of 180 iterations.

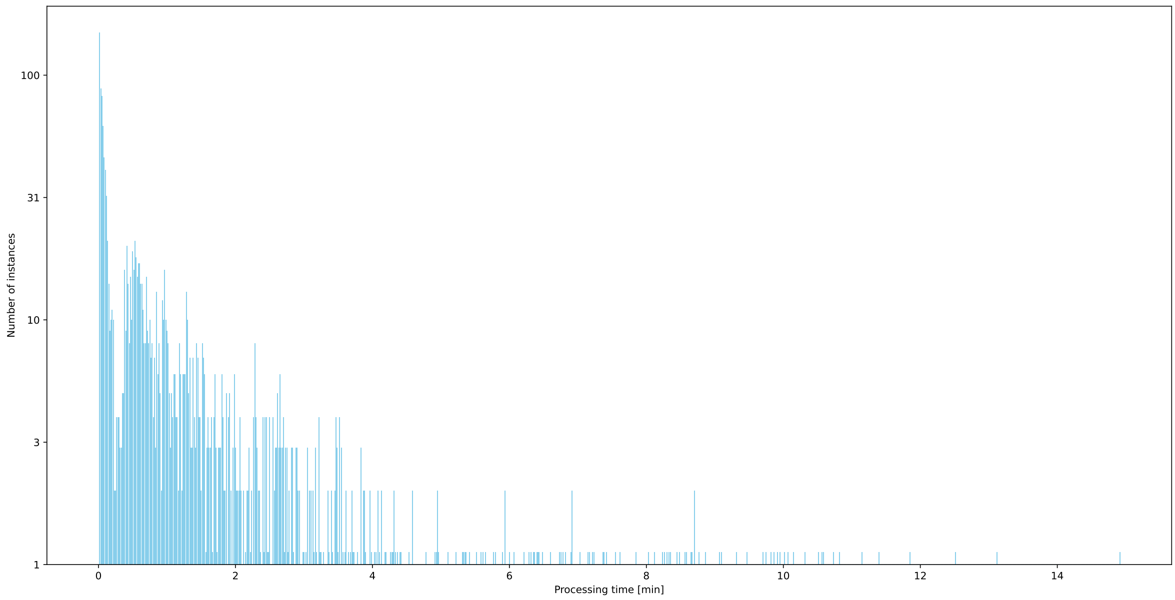
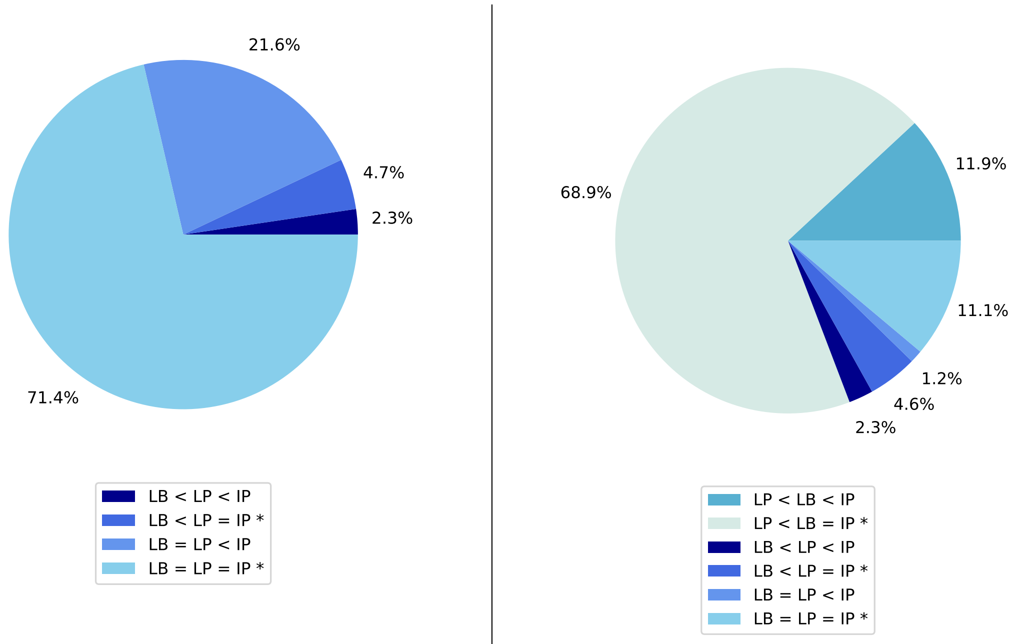


Figure 4b: Distribution of processing times without lower bounds

C.2. Optimality

To assess the quality of results, the problem-specific analytical lower bound (LB) is compared against the column generation's final RMP solution as a linear program (LP), and the final integer value (IP) for each instance. The occurence of their relations is counted and depicted in the pie charts in Figure 11a and in Figure 11b. The cases in which the heuristic reaches optimality are marked with a star symbol in the legend. Hence, the column generation heuristic provides optimal results in 76.1% of all instances if the lower bound is considered. For the calculation of plain column generation without lower bounds, the heuristic reaches optimal results in 84.6% of all instances. The higher ratio of optimality when neglecting the lower bounds can be explained as column generation has a longer runtime and discovers more columns, in total. Overall, IP is identical for 84.9% of the instances with and without consideration of the lower bound.

Evaluating the relative gap between the LP value and the IP value, the column generation heuristic reaches a mean optimality gap of 13.55% when lower bounds are included and of 64.42% without consideration of lower bounds. However, the objective function used for our calculations is a weighted sum of an underlying ordinal scale (number of highly qualified and lower qualified teams). Therefore, the assessment of optimality gaps requires another than the percentage view. To interpret gap values properly, the absolute difference regarding the number of highly qualified teams and lower qualified teams is analysed. In this context, the results of IP are compared against LB. As stated before, the optimal staff mix is identified in 76.1% of all instances if the lower bound is embedded into the algorithm. In further 18.8% ,the aggregated number of all workers is identical for LB and IP, i.e., the overall staff demand is the same but



(a) including lower bounds into column generation

(b) plain column generation without lower bounds

Figure 11: Distribution of objective value relations

the qualification mix is differently. Exemplarily, 16.8% of the instances employ one highly qualified team in the IP solution that is likely to be replaced by a lower qualified team in the optimal solution according to the LB. This means that in 94.9% of all instances the aggregated number of required staff is identified optimally when the lower bound is applied. A precise overview of these results is provided in the fourth column of Table 5. If the lower bound is not considered within the column generation heuristic, the optimal staff mix is identified in 84.6% of cases. For further 10.3% of the instances, the aggregated number of all workers is identical for LB and IP. I.e. in total, for 94.9% of all instances the optimal number of required staff is identified as well.

Staff mix difference			**Occurence accross 1620 runs**	
HQ[1]	**LQ**[1]	**Total**[2]	**with LB**	**without LB**
0	0	0	76.1%	84.6%
1	-1	0	16.8%	10.2%
2	-2	0	2.0%	0.1%
0	1	1	1.5%	3.7%
1	0	1	3.0%	1.3%
2	-1	1	0.05%	-
1	1	2	0.5%	0.05%
1	2	3	0.05%	-

[1] surplus of number of highly (HQ) / lower (LQ) qualified teams
[2] difference regarding aggregated number of teams in staff mix

Table 5: Gap analyses considering the absolute difference of number of workers between LB and IP

Appendix B.

Analyzing staffing decisions

Ebel SS, Grieger M, Erhard M, Brunner JO (2024): **Analyzing staffing decisions subject to flexible scheduling and demand volatility in healthcare.**

Status: to be submitted soon to Springer Journal of Scheduling; Category A.

Original Research Paper

Analyzing staffing decisions subject to flexible scheduling and demand volatility in healthcare

Stefanie S. Ebel[1], Milena Grieger[1], Melanie Erhard[1], and Jens O. Brunner[1,2,3]

[1]Health Care Operations / Health Information Management,
Faculty of Business and Economics, Faculty of Medicine, University of Augsburg,
Universitätsstraße 16, 86159 Augsburg, Germany
[2]Department of Technology, Management and Economics,
Technical University of Denmark, Kongens Lyngby, Denmark
[3]Next Generation Technology, Region Zealand, Denmark

Correspondence: jotbr@dtu.dk

State: 29 August, 2024

Analyzing staffing decisions subject to flexible scheduling and demand volatility in healthcare

Abstract

Medical staff generates the largest and most relevant cost volume in hospitals. Thus, personnel staffing and scheduling are of crucial importance. We focus on staffing decisions subject to a flexible shift system considering break assignments and multiple demand profiles. The number of required employees is minimized subject to essential constraints such as the coverage of forecasted demand in each planning period and legal regulations. We formulate the problem as an integer programming model and provide decomposed model reformulations to apply a column generation heuristic. In our computations, we use random data based on various settings which apply several volatile demand profiles and shift profiles. Solving 1,600 instances, the column generation heuristic leads to near-optimal results within reasonable runtime. We extensively study the effects of flexibility and volatility on staffing decisions. Our analysis reveals a significant reduction of the workforce size as flexibility increases in the problem setting. For several settings, the workforce can by reduced by more than 20% if flexible shift planning is induced. Additionally, it outlines the behalf of considering break assignment in flexible scheduling. Finally, it provides key factors for staffing estimations in different healthcare settings.

Keywords

flexible scheduling, column generation, healthcare, mixed-integer programming

Acknowledgment

This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) and Grant no. 438507036.

1 Introduction

Throughout the last 30 years, the number of hospital patients in Germany increased by more than 30%. During the same time, the number of hospitals decreased by more than 20% [4, 27]. Concurrently, hospitals have to cope with increased workload and further challenges, like staff shortages due to burnout and vacant positions [13, 24, 30]. Moreover, since medical staff accounts for more than 60% of hospitals' costs [28], efficient staffing and scheduling decisions are crucial for hospital management. Generally, physician and nurse scheduling problems are difficult to solve, but are often handled manually in hospitals, making them error-prone [15]. The literature on healthcare staff scheduling mostly neglects break assignments in flexible shift models [8, 10, 12]. Consequently, schedules frequently lack viability regarding staff satisfaction [16], overtime [14] and break assignments. Overall, the consideration of flexible shift systems including break assignments and its effect on staffing decisions is crucial for practice. Addressing break assignments is important since healthcare workers are frequently not able to take their breaks on an operational level, even though it is mandatory by law [31]. To tackle this issue on the strategic level, staffing models need to be extended by incorporating break assignment decisions.

The purpose of this paper is to study the trade-off between staffing decisions and shift flexibility with special consideration of break assignments. Our contribution is as follows. First, we develop an integer program (IP) which provides full flexibility regarding shift systems (i.e., starting time, length, and break assignment). Our main objective is to determine the optimal workforce size subject to demand coverage and general labor regulations, like rest periods and working hours. Particularly, we extend the break assignment modeling idea by Bechtold and Jacobs [2] allowing a complete overlap of shifts. Due to the problem's complexity, the IP is decomposed and solved by means of a problem-specific column generation heuristic. Second, we study the trade-off decision with randomly generated demand based on eight different patterns and eight types of flexible shift systems. The stylized profiles resemble real-world operations in healthcare. We receive high-quality results by solving 1,600 instances and deduce valuable insights. The analyses show that flexible systems lower the workforce size, the staffing estimation, and overstaffing. Workforce savings between 15% to 35% can be observed contrasting rigid with flexible shift systems. Additionally, we measure the benefit of considering break assignments, which leads to more precise staffing from a practical perspective. These information can be utilized by medical decision makers in staffing processes, hospital's department planning, or union contract setup.

The paper is structured in the following way: In Section 2, we review the relevant literature on healthcare staffing and scheduling. Afterwards, we provide a precise problem definition followed by the mathematical model (see Section 3). Next, we present the model decomposition and column generation solution approach in Section 4. In Section 5, we describe our study settings and evaluate the effects of staffing decisions with various demand profiles and types of flexible shift systems. Eventually, we conclude the paper in Section 6 and give ideas for further research.

2 Literature review

Hospital staff scheduling has received increasing attention throughout the last decades, especially nurse scheduling [3, 8, 10, 25] and physician scheduling [12]. Erhard et al. [12] review features and models for physician scheduling, concluding a lack of research in flexible shift types and break planning. This is also verified in Xu and Hall [32] who survey fatigue in personnel planning with respect to shift and rest scheduling. Subsequently, we review literature related to healthcare shift scheduling in general, afterwards we focus on flexible shift and break scheduling in both healthcare and non-healthcare environments.

General healthcare shift scheduling Using predefined shifts Camiat et al. [9] introduce a productivity index for physician scheduling in an emergency department. Comparing the results of their mixed-integer program (MIP) to a partner hospital, the schedules' fairness remains, staff undercoverage and department's overcrowding reduce, and the quality of care improves. Guo and Bard [17] consider working day patterns and planned overtime in monthly nurse shift scheduling using column generation. By testing artificial instances, they receive desirable results from both a methodological and practical perspective.

Flexible shift scheduling in healthcare Flexibility in healthcare staff scheduling is frequently limited to shifts. Fügener and Brunner [14] address overtime and working hours in physician scheduling. They find that flexible shift extensions reduce unplanned overtime and decrease staffing levels. Isken and Aydas [18] determine a set of flexible shifts matching variable demand. Solving their MIP in two phases, they conclude that flexibility improves demand matching and lowers costs by 10% to 20%. Böðvarsdóttir et al. [5] develop a MIP and a flexible framework for nurse scheduling with additional requirements, met by e.g., defining shifts as needed. They receive near-optimal results, desirable and valuable from a practical perspective. Flexible breaks are barely found in healthcare literature and are first considered by Lim et al. [21]. They solve the nurse-to-surgery assignment using column generation and a swapping heuristic. Afterwards, they apply a break scheduling MIP to the solution and conclude that the outcomes are efficient and admissible in most of their test instances.

Flexible shift and break scheduling in healthcare Brunner et al. [6, 7] first investigate flexible shift and break scheduling in physician scheduling. First, solving a MIP and a heuristic approach [6], and later a branch-and-price algorithm [7], both approaches provide high-quality schedules that outperform the current practice. Moreover, branch-and-price copes with several real-world options and solves large instances efficiently. To the best of our knowledge similar approaches are only studied by Stolletz and Brunner [29], Erhard [11], and Kraul et al. [20]. Each of them applies shift matrices containing all eligible shifts, which outperforms implicit shift modeling as shown in Stolletz and Brunner [29]. Erhard [11] analyses input effects on physician staffing levels over a multi-week horizon. They investigate the number of allowed working days, shift lengths, the planning horizon's length or high vs. low demand, and point out resulting differences in the performance. Kraul et al. [20] take operational planning into account and ensure the viability of rest assignments considering different educational levels. Their approach includes probability factors and safety levels, and provides high-quality results from both an algorithmic and practical point of view.

Flexible shift and break scheduling in non-healthcare settings In further personnel planning literature, Pandey et al. [26] determine the optimal workforce size considering flexible shift and break assignments in a three-step heuristic approach. Presenting near-optimal results, they reduce under- and overstaffing with raising problem size. Additional studies include varying demand profiles with flexible scheduling: Álvarez et al. [1] and Kiermaier et al. [19] research the effects of multiple break profiles and conclude that their approaches lower under-/overstaffing and labour costs. Mattia et al. [23] consider breaks as a random component that employees decide for themselves. Presenting a multi solution search heuristic, they conclude to bridge the gap between scheduling breaks upfront and replanning them.

Our work uniquely addresses effects on the workforce size considering various demand profiles and diverse options for flexible shift and break scheduling. Table 1 surveys the problem settings, methodologies, and scopes of flexibility in the literature. Our aim is to research optimal staffing in the healthcare sector. Therefore, we present an IP model offering fully flexible shifts in a predefined shift matrix, as well as flexible break scheduling. We test our approach with multiple demand profiles and apply column generation as a solution approach.

		Context			Flexibility				Methodology		
Author	**Year**	**Healthcare**	**Staffing insights**	**Planning horizon**	**Demand profiling**	**Predefined shifts**	**Flexible shifts**	**Flexible breaks**	**Approach[1]**	**MIP**	**Decomposition based approach[2]**
Álvarez et al.	2020			1 day	✗	✗	✗	✗	H	✗	
Böðvarsdóttir et al.	2022	✗		4 weeks		✗	✗		E	✗	
Brunner et al.	2009	✗		2 weeks			✗	✗	H	✗	
Brunner et al.	2010	✗		2-6 weeks	✗		✗	✗	E	✗	BP
Camiat et al.	2019	✗		3 months		✗			E	✗	
Erhard	2021	✗	✗	1-6 weeks	(✗)	✗	✗	✗	H	✗	CG
Fügener and Brunner	2019	✗	✗	1 week			✗		H	✗	CG
Guo and Bard	2022	✗		4 weeks		✗			H	✗	CG
Isken and Aydas	2022	✗		4 weeks			✗		E	✗	
Kiermaier et al.	2020			1 week	✗	✗	✗	✗	H	✗	
Kraul et al.	2024	✗		1-4 weeks		✗	✗	✗	E	✗	BP
Lim et al.	2016	✗		1 day		✗		✗	H		CG
Mattia et al.	2023			1 day	✗	✗		✗	H	✗	
Pandey et al.	2021		✗	1 week			(✗)	✗	H	✗	
Stolletz and Brunner	2012	✗		2 weeks		✗	✗	✗	H	✗	
This work		✗	✗	1 week	✗	✗	✗	✗	H	✗	CG

[1] E: exact solution approach, H: heuristic solution approach, S: simulation
[2] BP: branch-and-price, CG: column generation

Table 1: Literature overview

3 Problem description

The problem examined in our work covers a planning horizon consisting of D days (e.g., one week), each divided into T periods (e.g., 1-hour intervals). The number of planning periods per (planning) day can be less than or equal to 24 hours and may overlap two calendar days. For instance, a planning day might start at 5 AM and end at 1 AM the next real day, resulting in $|T| = 20$ periods, each 1-hour long. We consider a homogeneous set of employees E, who are to be scheduled while respecting a minimum (maximum) amount of working hours $H^{\min}$ ($H^{\max}$) over the planning horizon. This regulation is either stated in the labor contract or based on special agreements between the concerned parties.

We introduce two kinds of flexibility in our model: shift flexibility and break flexibility. The set of flexible shifts $s \in S$ is predefined according to the following parameters. Each shift s is defined by a flexible starting period T_s^{first} and ending period T_s^{last}, as well as a flexible shift length W_s, which ranges between a minimum length $S^{\min}$ and a maximum length $S^{\max}$, e.g., between seven and 13 periods. A shift matrix is generated providing all the available shift types. This matrix consists of binary parameters A_{st}, which indicate a working period t for a specific shift s (equal to one) or not (equal to zero). Consequently, each shift starting on a (planning) day ends on the same (planning) day. A minimum off-duty rest time R must be ensured between two consecutive shift assignments. To maintain continuity throughout an employee's week, it may be desirable to limit the number of shift types S^{num} used per employee per week. Breaks may be placed flexibly or may be omitted altogether. If breaks are considered, each shift includes a single break that can be assigned flexibly within a predefined break window. This window is determined by a minimum number of working periods W^{pre} before the break and a minimum number of working periods W^{post} after the break. To guarantee a certain quality and quantity of service and care, the generated schedule has to ensure that the (forecasted) demand N_{dt} is covered in each period t of every day d within the planning horizon. If breaks are neglected, all shifts are reduced by one period, i.e., $S^{\min}$ and $S^{\max}$ are reduced by one. Note that these schedules are generated to ensure the feasibility of our staffing decision and are not intended to be in use on an operational level.

We formulate the problem as an IP model. In preliminary work, we tested different modeling approaches for flexible shift systems with breaks. Breaks can be included when constructing the shift matrix [29], which displaces some constraints but considerably increases the size of the shift matrix. Alternatively, the shift matrix contains flexible shifts without breaks, requiring specific constraints to schedule breaks appropriately. In this context, breaks may be either assigned to individual employees, shifts, or periods, extending the approach of Bechtold and Jacobs [2]. After evaluating these formulations in terms of solution quality and solution times, we found that assigning breaks to periods is the superior approach. Consequently, in our formulation, breaks are modeled implicitly and must be assigned to each individual in a post-processing step which is computationally simple [2]. Our main objective is to minimize the total workforce size while ensuring appropriate levels of care, i.e., forecasted demand coverage, and adherence of common labor regulations. The following notation is used to formulate the model.

Indices and sets

$e \in E$	set of employees with index e
$d \in D$	set of days with index d
$t \in T$	set of daily time periods with index t
$s \in S$	set of shifts with index s
$t \in L^{\text{start}} \subseteq T$	set of all possible starting periods for a (lunch) break
$t \in L^{\text{end}} \subseteq T$	set of all possible ending periods for a (lunch) break

Parameters

$S^{\min}$	minimum shift length
$S^{\max}$	maximum shift length
S^{num}	maximum number of different shifts per employee and week
A_{st}	1 if shift s covers period t, 0 otherwise
T_s^{first}	first working period in shift s
T_s^{last}	last working period in shift s
W_s	number of working periods in shift s
W^{pre}	number of working periods before the break is allowed to start
W^{post}	number of working periods after the break has ended
R	minimum number of rest periods between consecutive shifts
$H^{\min}$	minimum number of regular working periods (working hours) per employee and week
$H^{\max}$	maximum number of regular working periods (working hours) per employee and week
N_{dt}	demand in period t of day d
L^{first}	earliest possible period for a (lunch) break for each shift type per planning day
L^{last}	latest possible start for the beginning of a (lunch) break for each shift type per day
V	auxiliary shortcut for overlapping shifts' break time windows defined as $V = (S^{\max} - W^{\text{pre}} - W^{\text{post}}) - 1$

Decision variables

y_e	$\in \{0; 1\}$	1 if employee e is working in the planning horizon, 0 otherwise
x_{esd}	$\in \{0; 1\}$	1 if employee e is assigned to shift s on day d, 0 otherwise
z_{es}	$\in \{0; 1\}$	1 if shift s is used for employee e, 0 otherwise
l_{dt}	$\in \mathbb{N}_0$	number of (lunch) break assignments in period t on day d

$$\min \quad \sum_{e\in E} y_e \tag{1.1}$$

$$\text{s.t.} \quad \sum_{s\in S} x_{esd} \leq 1, \qquad \forall e \in E, d \in D, \tag{1.2}$$

$$x_{es_1d} + x_{es_2(d+1)} \leq 1, \qquad \forall e \in E, s_1, s_2 \in S :, \; |T|-T^{\text{last}}_{s_1}+T^{\text{first}}_{s_2}-1<R, d \in D\backslash|D|, \tag{1.3}$$

$$H^{\min}y_e \leq \sum_{s\in S}\sum_{d\in D} W_s x_{esd} \leq H^{\max}y_e, \qquad \forall e \in E, \tag{1.4}$$

$$x_{isd} \leq z_{es}, \qquad \forall e \in E, s \in S, d \in D, \tag{1.5}$$

$$\sum_{s\in S} z_{es} \leq S^{\text{num}} \qquad \forall e \in E, \tag{1.6}$$

$$\sum_{e\in E}\sum_{s\in S} A_{st}x_{esd} - l_{dt} \geq N_{dt}, \qquad \forall d \in D, t \in T, \tag{1.7}$$

$$\sum_{e\in E}\sum_{s\in S} x_{esd} = \sum_{t\in T} l_{dt}, \qquad \forall d \in D, \tag{1.8}$$

$$\sum_{\tau=L^{\text{first}}}^{t} l_{d\tau} - \sum_{e\in E} \sum_{\substack{s\in S:\\ T^{\text{last}}_s-W^{\text{post}}\leq t}} x_{esd} \geq 0, \qquad \forall d \in D, t \in L^{\text{end}}\backslash L^{\text{last}}, \tag{1.9}$$

$$\sum_{\tau=t}^{L^{\text{last}}} l_{d\tau} - \sum_{e\in E} \sum_{\substack{s\in S:\\ t\leq T^{\text{first}}_s+W^{\text{pre}}}} x_{esd} \geq 0, \qquad \forall d \in D, t \in L^{\text{start}}\backslash L^{\text{first}}, \tag{1.10}$$

$$\sum_{\tau=t}^{\theta} l_{d\tau} - \sum_{e\in E} \sum_{\substack{s\in S:\\ T^{\text{first}}_s+W^{\text{pre}}\geq t\\ \theta\geq T^{\text{last}}_s-W^{\text{post}}}} x_{esd} \geq 0, \qquad \begin{array}{l}\forall d \in D, t \in \{L^{\text{first}}, \ldots, L^{\text{last}}\},\\ \theta\in\{q, \ldots, \min(t+V, L^{\text{last}})\}\end{array}, \tag{1.11}$$

$$y_e, x_{esd}, z_{is} \in \{0, 1\}, \qquad \forall e \in E, s \in S, d \in D, \tag{1.12}$$

$$l_{dt} \in \mathbb{N}_0, \qquad \forall d \in D, t \in T \tag{1.13}$$

The objective function (1.1) minimizes the total number of employees required over the planning horizon to cover the forecasted demand. The first block of constraints (1.2) to (1.6) models feasible shift assignments. Specifically, constraints (1.2) guarantee that every employee e is assigned to at most one shift per day from a set of flexible shift types. Due to working regulations, certain two-shift combinations are forbidden for a single employee on consecutive working days. Constraints (1.3) enforce a minimum off-duty rest time R between two shift assignments on consecutive working days d and $d+1$. Constraints (1.4) handle the weekly working hours for each employee e according to the allowed range from $H^{\min}$ to $H^{\max}$. Constraints (1.5) and (1.6) are optional and can be included if the model should limit the maximum number of shifts per employee. Specifically, constraints (1.5) ensure that employees are only assigned to shifts that are selected for them, and constraints (1.6) sum the number of active shifts per employee, ensuring the total does not exceed the maximum S^{num}. Constraints (1.7) assure (forecasted) demand coverage subject to break assignments in each period t on day d. The number of employees on duty is reduced by those taking a break, ensuring a minimum number of employees are always working. The final set of constraints (1.8) to (1.11) determines the flexible break assignments. We fundamentally extend the modeling approach presented by Bechtold and Jacobs [2] because their method does not account for overlapping shifts, where the break window of one shift can be completely within the break window of another shift. Constraints (1.8) assure that the number of shift assignments on any day d is equal to the number of break assignments on the same day. In addition, the appropriate location of the break according to the labor regulations must be assured. Constraints (1.9) are called forward-passing constraints [2]. They force the number of breaks assigned from the first possible break period up to each possible ending period t to be bigger than or equal to the number of shift assignments with break window ending by no later than t. In contrast, constraints (1.10) are called backward-passing constraints and work in the opposite direction. They ensure that the number of break assignments in and after period t is at least as great as the number of shift assignments with break windows starting from t onward. While these constraints are sufficient for assigning breaks when the break window of any shift type is not a subset of another shift type [2], our model's flexibility allows for overlapping shift types. This can result in a break window of one shift becoming a subset of the break window of another, necessitating an extension of the mathematical formulation. Constraints (1.11) ensure that every shift s gets a break within its break window assigned, even though this break window is part of the break window of another shift. To illustrate the need for these constraints, consider two shifts: s_1, with a break window from period two to period seven, and s_2, with a break window from period three to period five. A feasible assignment without constraints (1.11) could assign a break in period two and another break in period six. The newly introduced constraints forbid this assignment where shift s_1 gets two breaks while shift s_2 gets no break. For this, constraints (1.11) consider each break window explicitly. In theory, a huge number of constraints might be added. However, only some of them would be tight. For practical instances, the number of constraints seems to be moderate as shown in our computations where we have not encountered any difficulties with the extension slowing down the solution process. Finally, variable definitions are given in (1.12) and in (1.13).

Note, this way of assigning breaks requires an additional post-processing mechanism to create final schedules with appropriate break allocations. However, post-processing has almost no computational burden since a greedy procedure is all that is needed. For each day, the greedy heuristic considers each period (in ascending order) and assigns the breaks to the employee first whose break window has the smallest number of periods left [2]. Since we focus on staffing, the individual shift schedules including appropriate break assignments are of minor interest.

4 Solution approach

Since solving the generic formulation optimally is computationally intensive, we develop a reformulation based on Dantzig-Wolfe decomposition consisting of one restricted master problem (RMP) and one pricing subproblem (SP). We decompose the problem by employee, i.e., constraints (1.2) to (1.6) move into the subproblem generating weekly shift assignment plans for employees. The resulting plans are denoted by set $j \in J$ and characterized by the following parameters. The information whether plan j covers period t on day d is stored in a binary parameter A_{jdt}. Likewise, parameter X_{jsd} stores the information whether plan j applies shift s on day d.

The RMP receives the rosters $j \in J$ as input but lacks a complete enumeration of all possible rosters. It applies them setting integer variables λ_j in order to cover demand and minimize the workforce. Additionally, the RMP schedules breaks as outlined before. To ensure feasibility throughout the solution process, we add integer variables u_{dt} capturing understaffing in period t on day d, weighted with penalty parameter C^{under}. Note that while A_{jdt} and X_{jsd} are parameters in the RMP, they reflect the values of decision variables in the SP.

Additional indices and sets

$j \in J$	set of shift plans with index j

Additional parameters

C^{under}	penalty weight for understaffing per period
A_{jdt}	1 if plan j covers period t on day d, 0 otherwise
X_{jsd}	1 if plan j applies shift s on day d, 0 otherwise

Additional decision variables

λ_j	$\in \mathbb{N}_0$	number of employees working shift plan j
u_{dt}	$\in \mathbb{N}_0$	number of understaffed employees in period t on day d

$$\min \quad \sum_{j\in J} \lambda_j + C^{\text{under}} \cdot \sum_{d\in D} \sum_{t\in T} u_{dt} \tag{2.1}$$

$$\text{s.t.} \quad \sum_{j\in J} A_{jdt}\lambda_j - l_{dt} + u_{dt} \geq N_{dt}, \qquad \forall d \in D, t \in T, \tag{2.2}$$

$$\sum_{j\in J} \sum_{s\in S} X_{jsd}\lambda_j = \sum_{t\in T} l_{dt}, \qquad \forall d \in D, \tag{2.3}$$

$$\sum_{\tau=L^{\text{first}}}^{t} l_{d\tau} - \sum_{j\in J} \sum_{\substack{s\in S:\\ T_s^{\text{last}}-W^{\text{post}}\leq t}} X_{jsd}\lambda_j \geq 0, \quad \forall d \in D, t \in L^{\text{end}} \backslash L^{\text{last}}, \tag{2.4}$$

$$\sum_{\tau=t}^{L^{\text{last}}} l_{d\tau} - \sum_{j\in J} \sum_{\substack{s\in S:\\ t\leq T_s^{\text{first}}+W^{\text{pre}}}} X_{jsd}\lambda_j \geq 0, \qquad \forall d \in D, t \in L^{\text{start}} \backslash L^{\text{first}}, \tag{2.5}$$

$$\sum_{\tau=t}^{\theta} l_{d\tau} - \sum_{j\in J} \sum_{\substack{s\in S:\\ T_s^{\text{first}}+W^{\text{pre}}\geq t\\ \theta\geq T_s^{\text{last}}-W^{\text{post}}}} X_{jsd}\lambda_j \geq 0, \qquad \begin{array}{l} \forall d \in D, t \in \{L^{\text{first}}, \ldots, L^{\text{last}}\}, \\ \theta \in \{t, \ldots, \min(t+V, L^{\text{last}})\} \end{array}, \tag{2.6}$$

$$l_{dt}, u_{dt} \in \mathbb{N}_0, \qquad \forall d \in D, t \in T, \tag{2.7}$$

$$\lambda_j \in \mathbb{N}_0, \qquad \forall j \in J \tag{2.8}$$

The RMP's objective (2.1) minimizes the number of applied rosters, i.e., the workforce. Moreover, it minimizes penalties resulting from understaffing. Constraints (2.2) is an amendment of (1.7) that includes demand undercoverage variables and is linked with dual variables $\pi_{dt} \geq 0$. Regarding break scheduling, constraints (2.3) to (2.6) correspond to (1.8) to (1.11), with dual variables $\mu_d \in \mathbb{R}$, $\gamma^1_{dt} \geq 0$, $\gamma^2_{dt} \geq 0$, and $\gamma^3_{dt\theta} \geq 0$, respectively. Finally, constraints (2.7) to (2.8) ensure the domains of the binary and integer variables.

As set J of the RMP includes only a subset of all possible weekly rosters, promising schedules need to be generated on demand. Therefore, the information of the dual variables is handed to the pricing subproblem which creates convenient weekly shift schedules for the employees. The binary decision variables x_{sd} store whether shift s is assigned on day d, and accordingly, binary variables a_{dt} capture whether period t of day d is a working period. This refers to the parameters X_{jsd} and A_{jdt} of the RMP.

$$
\begin{aligned}
\min \quad & 1 - \sum_{d \in D} \sum_{t \in T} \pi_{dt} a_{dt} + \sum_{d \in D} \sum_{s \in S} \mu_d x_{sd} + \sum_{d \in D} \sum_{t \in L^{\text{end}} \setminus L^{\text{last}}} \sum_{\substack{s \in S: \\ T_s^{\text{last}} - W^{\text{post}} \leq t}} \gamma^1_{dt} x_{sd} \qquad (3.1) \\
& + \sum_{d \in D} \sum_{t \in L^{\text{start}} \setminus L^{\text{first}}} \sum_{\substack{s \in S: \\ t \leq T_s^{\text{first}} + W^{\text{pre}}}} \gamma^2_{dt} x_{sd} \\
& + \sum_{d \in D} \sum_{t = L^{\text{first}}}^{L^{\text{last}}} \sum_{\theta = t}^{\min(t+V, L^{\text{last}})} \sum_{\substack{s \in S: \\ T_s^{\text{first}} + W^{\text{pre}} \geq t \\ \theta \geq T_s^{\text{last}} - W^{\text{post}}}} \gamma^3_{dt\theta} x_{sd} \\
\text{s.t.} \quad & \sum_{s \in S} x_{sd} \leq 1, \qquad \forall d \in D, \qquad (3.2) \\
& x_{s_1 d} + x_{s_2 (d+1)} \leq 1, \qquad \forall s_1, s_2 \in S : |T| - T_{s_1}^{\text{last}} + T_{s_2}^{\text{first}} - 1 < R, \\
& \qquad d \in D \setminus |D|, \qquad (3.3) \\
& H^{\min} \leq \sum_{s \in S} \sum_{d \in D} W_s x_{sd} \leq H^{\max}, \qquad (3.4) \\
& x_{sd} \leq z_s, \qquad \forall s \in S, d \in D, \qquad (3.5) \\
& \sum_{s \in S} z_s \leq S^{\text{num}}, \qquad (3.6) \\
& \sum_{s \in S} A_{st} x_{sd} = a_{dt}, \qquad \forall d \in D, t \in T, \qquad (3.7) \\
& x_{sd}, z_s \in \{0, 1\}, \qquad \forall s \in S, d \in D, \qquad (3.8) \\
& a_{dt} \in \{0, 1\}, \qquad \forall d \in D, t \in T \qquad (3.9)
\end{aligned}
$$

The objective of the SP (3.1) minimizes the reduced cost of absent columns. A new promising column for the RMP is identified if the reduced cost is negative. Constraints (3.2) to (3.6) are equivalent formulations of (1.2) to (1.6), however, index e has been dropped due to the decomposition strategy by employee. Constraints (3.7) serves to match variables x_{sd} and a_{dt}, while constraints (3.8) and (3.9) ensure the domain of the decision variables.

The models of RMP and SP are embedded into a column generation heuristic depicted in Figure 1. At first, the algorithm is set up and initialized. We start the procedure with an empty initialization, meaning that set J does not contain any rosters, resulting in the u_{dp} variables leading to the maximum possible objective value. From this starting point the solution is improved stepwise. An iteration starts solving the RMP in its linear programming form, referred to as RMP-LP. Once the RMP-LP is solved, the column generation heuristic solves the SP optimally to identify a potential new column. Next, the procedure checks whether the solution of the SP prices out, i.e., results in a column with negative reduced costs. If the solution provides a column with negative reduced costs, it is added to the RMP. Moreover, a Lagrangian bound (LB) is calculated from the objectives of RMP and SP, as outlined in Lübbecke and Desrosiers [22]. The algorithm then checks for convergence by comparing the difference between the RMP-LP solution and the Lagrangian bound against a predefined threshold ϵ. If the difference is higher than the threshold the algorithm starts a new iteration by solving the RMP-LP with the newly added columns. Once the difference between RMP-LP's objective and LB is smaller than the threshold, indicating sufficient convergence, or no adequate column is found, the algorithm stops it's iterative procedure. In the final stage, the RMP is solved as an IP, denoted by RMP-IP, to obtain a valid integer solution. Upon completing this step, the heuristic terminates, having successfully found an efficient solution.

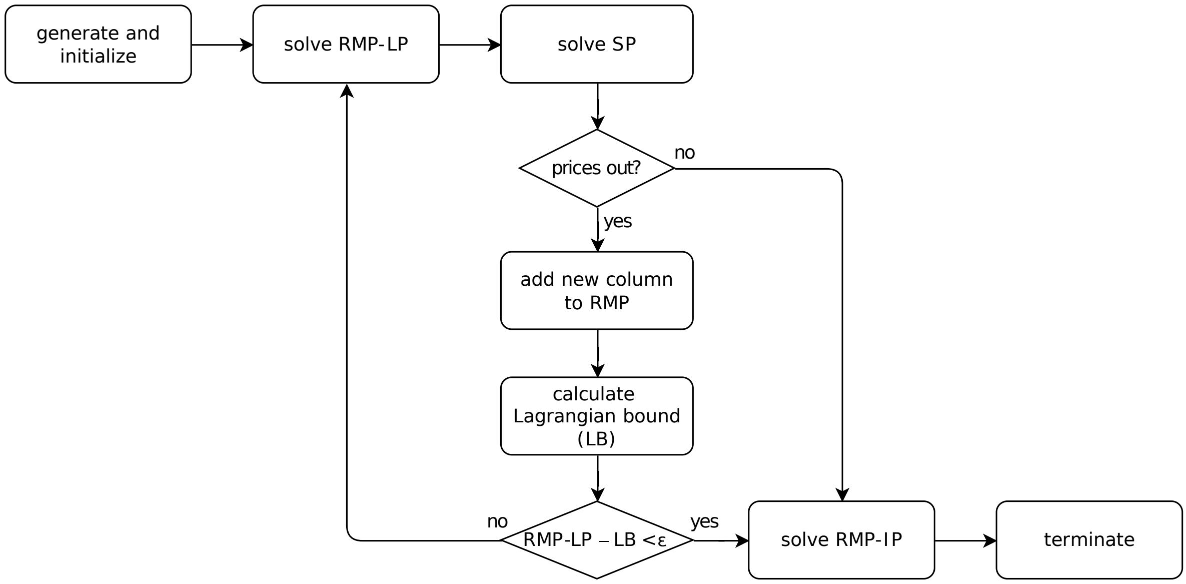


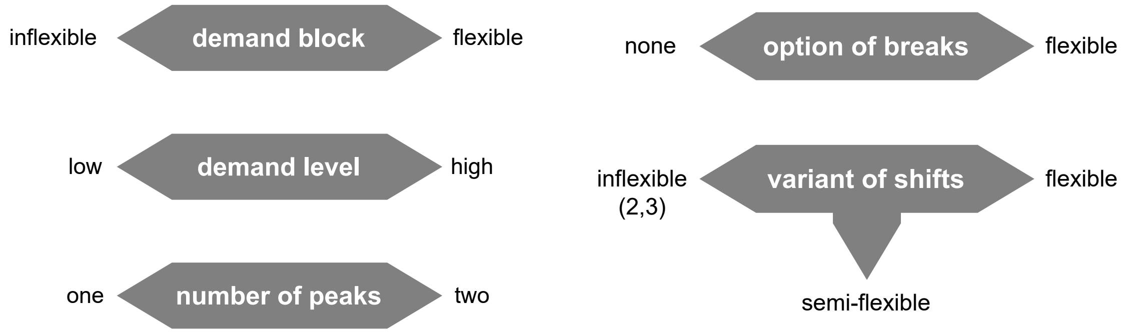
Figure 1: Process of the column generation heuristic

5 Computational experiments

Our study aims to gain insights into the impact of varying levels of flexibility in both the demand structure and the model configuration. Therefore, instances are set up according to different demand and shift profiles as described in Section 5.1. The column generation algorithm was implemented as a Java project employing CPLEX 20.1.0 and experiments are run on Windows 10 computer with 16 GB random access memory and an Intel® Core™ i5-8265U 1.8 Ghz quadcore processor. The performance of the solution approach is discussed in Section 5.2. Afterwards, Section 5.3 evaluates the benefits of flexible shift and break planning and concludes insights for healthcare planning processes.

5.1 Study setup for different flexibility settings

The settings of our instances are based on two different types of flexibility profiles, i.e., demand and shift profiles, that are varied systematically. The former influence the demand levels when instances are generated, the latter influence the constraints that are included when executing the column generation.



(a) Demand profiling - options for flexibility in demand generation **(b)** Shift profiling - options for flexibility in algorithm configuration

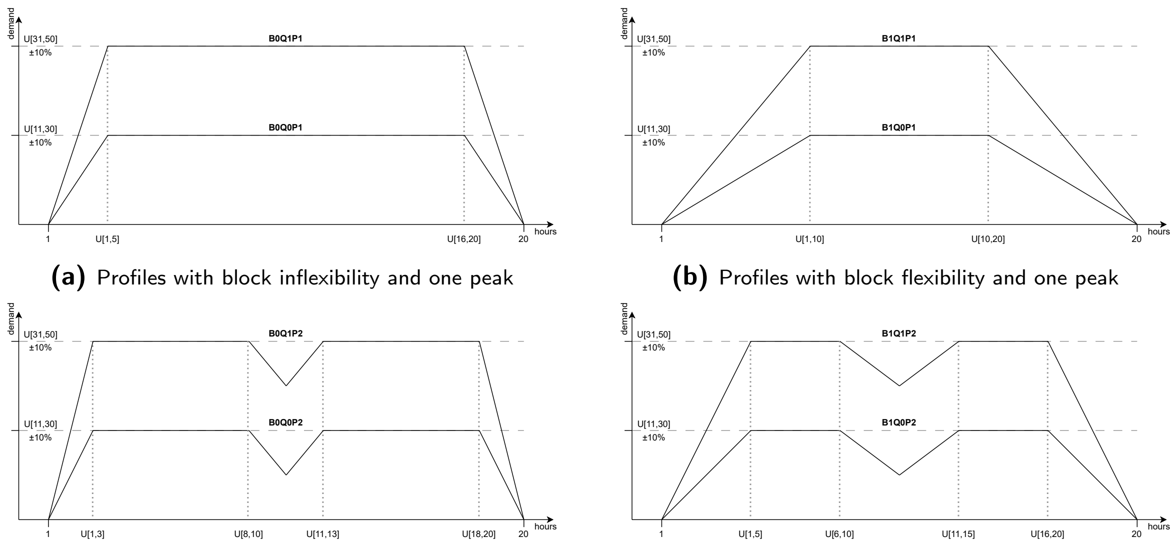
Figure 2: Overview on different types of flexibility

All demand profiles follow a similar pattern: demand rises in the beginning of the day, reaches a constant level (block region), and decreases by the end of the day. However, the demand profiles differ in three factors stated in Figure 2a. First, the block of demand can be either flexible (*B1*) or inflexible (*B0*). This means the length of the block region is adjustable when generating the instance, or it needs to fit within stricter time windows regarding its start and end. Secondly, either a low (*Q0*) or high (*Q1*) number of employees is required (demand quantity), where low categorizes a demand between 11 to 30 employees, and high means 31 to 50 employees. Third, there occur either one (*P1*) or two (*P2*) peaks, i.e., one or two block regions. The combination of these factors leads to eight demand profiles as presented in Table 3. A day is divided into 20 time periods.

Demand profile	Block flexibility	Demand quantity	Peaks
B0Q0P1	inflexible	low	1 peak
B0Q1P1	inflexible	high	1 peak
B1Q0P1	flexible	low	1 peak
B1Q1P1	flexible	high	1 peak
B0Q0P2	inflexible	low	2 peaks
B0Q1P2	inflexible	high	2 peaks
B1Q0P2	flexible	low	2 peaks
B1Q1P2	flexible	high	2 peaks

Table 3: Parameter setting of different demand profiles

For visualization, Figure 3 sketches the demand profiles. Exemplarily, to create an instance according to profile *B0Q0P1* (see Figure 3a) two time periods are randomly drawn from the uniform distributions $U[1,5]$ and $U[16,20]$ to state the beginning and the ending of the block, respectively. A block demand is drawn from a uniform distribution $U[11,30]$. For periods of increasing or decreasing demand we draw the actual realizations from a normal distribution. The mean μ corresponds to the linearly interpolated value of the respective line segment (either increasing or decreasing), with the standard deviation set to $\sigma = 3$. This procedure follows the logic described in Brunner et al. [7] and is based on real-world assumptions. To extend the procedure, we vary the block demand throughout the block's time periods by $\pm 10\%$, e.g., if a demand of 20 is drawn from the uniform distribution, the actual demand varies between 18 to 22 in every block time period. An exemplary instance of demand profile *B0Q0P1* is shown in Figure 4. A complete list of all instances is provided in the supplementary material. These demand profiles address several demand structures occurring in real-world hospital departments. Moreover, the stochastic components in the demand generation process concern demand volatility in healthcare settings. Thus, the demand generation process is aligned with results' robustness.



(a) Profiles with block inflexibility and one peak

(b) Profiles with block flexibility and one peak

(c) Profiles with block inflexibility and two peaks

(d) Profiles with block flexibility and two peaks

Figure 3: Design of different demand profiles

Shift profiles tackle the idea of different scopes of flexibility that are allowed in the solution process. From a practical perspective, they address different scheduling policies that may occur in hospital departments, such as rigid shift systems consisting of only a few shifts. They differ in two factors presented in Figure 2b. First, there is the option to plan breaks flexibly or to ignore break planning. Second, shifts may either be scheduled fully flexible, semi-flexible, or inflexible. If shifts are scheduled flexibly, there are all shifts available that can be generated respecting $S^{\min}$ and $S^{\max}$, so that $|S| = 77$ if breaks

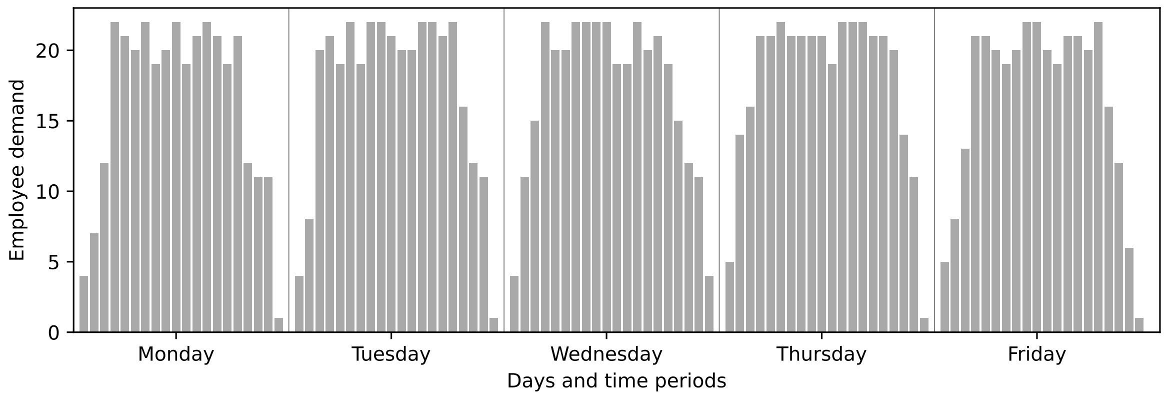


Figure 4: An exemplary instance of demand profile *B0Q0P1*

are scheduled and $|S| = 84$ if breaks are ignored. If the shift option is inflexible, set S is reduced to either two 12-hours shifts or three 8-hour shifts. With the semi-flexible option all of the shifts are available in set S, but a single column $j \in J$ may apply only one shift type which needs to be the same for all working days. Table 4 denotes all shift profiles resulting from the combination of these options. The shortcuts *O0* and *O1* indicate whether break planning (off-time) is included or excluded, as shown in the second column. The shortcuts F, M, and I encode the variant of the shift as flexible (F), semi-flexible (M), or inflexible (I) with the respective number of shifts being added as specified in the third column. The last two columns depict the effects for the column generation and show which constraints are (not) included in which shift profile.

Shift profile	**Break option**	**Shift variant**	$\lvert S \rvert$	**Constraints** (3.5)-(3.6)	**Constraints** (2.3)-(2.6)
O0F	none	flexible	84		
O0M	none	semi-flexible	84	✓	
O0I3	none	inflexible	3		
O0I2	none	inflexible	2		
O1F	flexible	flexible	77		✓
O1M	flexible	semi-flexible	77	✓	✓
O1I3	flexible	inflexible	3		✓
O1I2	flexible	inflexible	2		✓

Table 4: Model configuration per shift profile

Finally, the combination of eight demand profiles and 8 shift profiles gives us a total of 64 computational settings. For all of these settings, we solve 25 instances each, which results in an overall number of 1,600 instances solved.

5.2 Computational performance

In internal tests we found the MIP to lead to undesirable solution times as several instances were not solvable within a time limit of two hours. Therefore, the 1,600 instances are solved by means of the column generation heuristic. For most settings the algorithm terminates within seconds as given in Figure 5. The left part provides average solution times when break planning is neglected whereas the right part shows average solution times with embedded flexible break planning. Limiting the pool of shifts to only two or three shifts (*I2* and *I3*) the algorithm takes less than one second. For the other shift profiles (F and M) the algorithm takes a slightly higher, but still reasonable, amount of time.

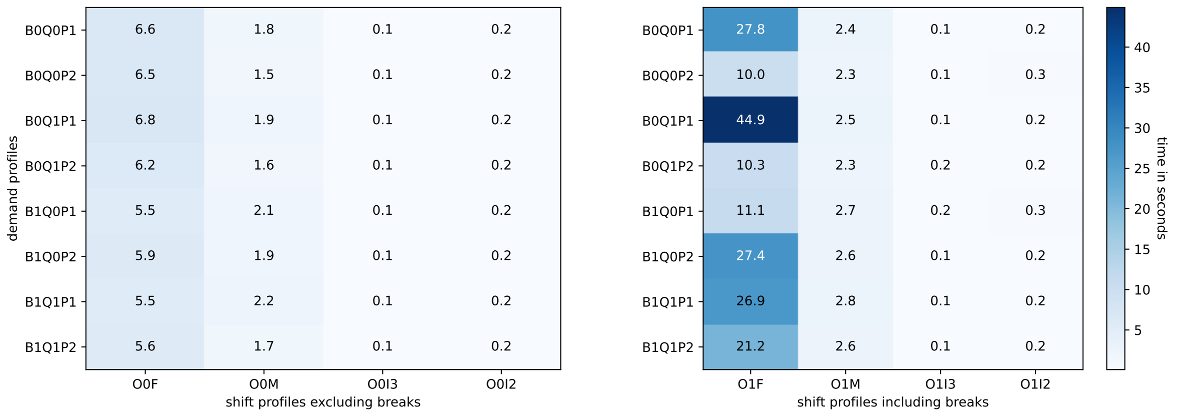


Figure 5: Average solution times in seconds per setting (each corresponds to the average of 25 instances)

Additionally, the column generation heuristic leads to near-optimal results as given in Figure 6. For the calculation of gaps, we evaluated the difference between the final RMP-LP's solution rounded up to the next integer value and the solution of RMP-IP. Taking a closer look into the results, each cell in Figure 6 depicts the average gap across 25 instances. Exemplarily, the average number of employees in setting *B0Q0P1-O0F* is 54.6 employees, with an average optimality gap of 1.0%. This means the solutions identified for the instances of this settings differs on average by 0.55 employees from the optimal solution. The highest mean gap of 2.3% is reached for setting *B1Q0P1-O1F*, requiring an average number of 38.1 employees. Consequently, the solutions differ on average by 0.88 employees from optimality.

5.3 Insights on staffing with flexible shift planning

The results for different settings are compared regarding the resulting workforce size and ratio between instances' demand and workforce power. These analyses permit staffing insights related to different settings of flexibility.

In a first step, we review the workforce required to cover demand in each setting of the study. The grouped bars in Figure 7a and Figure 7b show the average number of

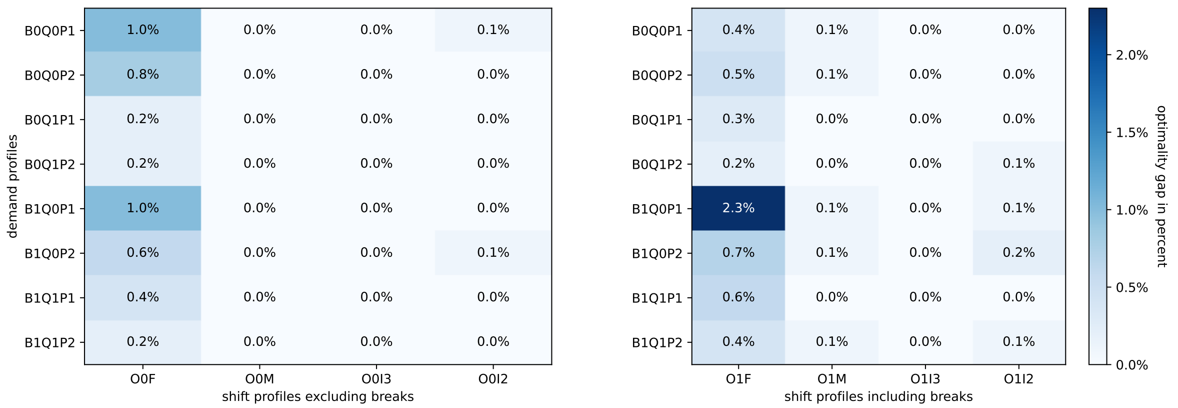
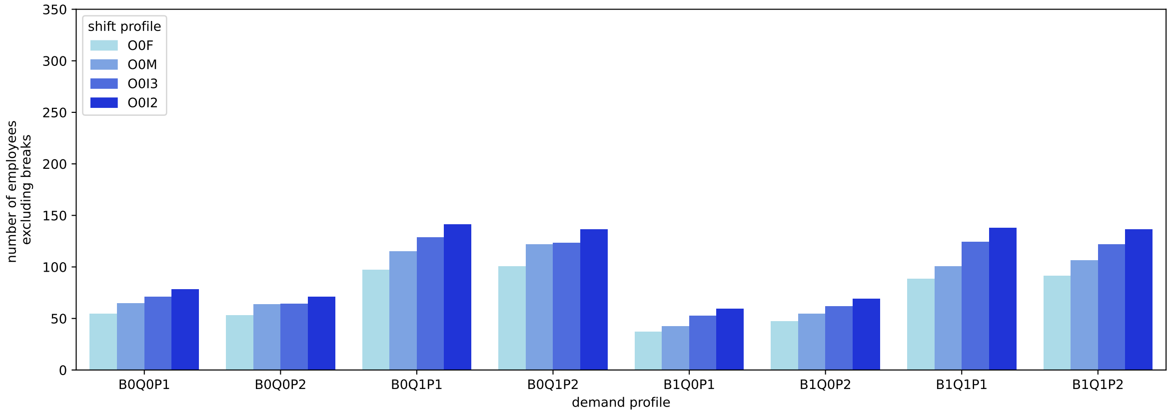


Figure 6: Average optimality gaps of the column generation heuristic per setting (each corresponds to the average of 25 instances)

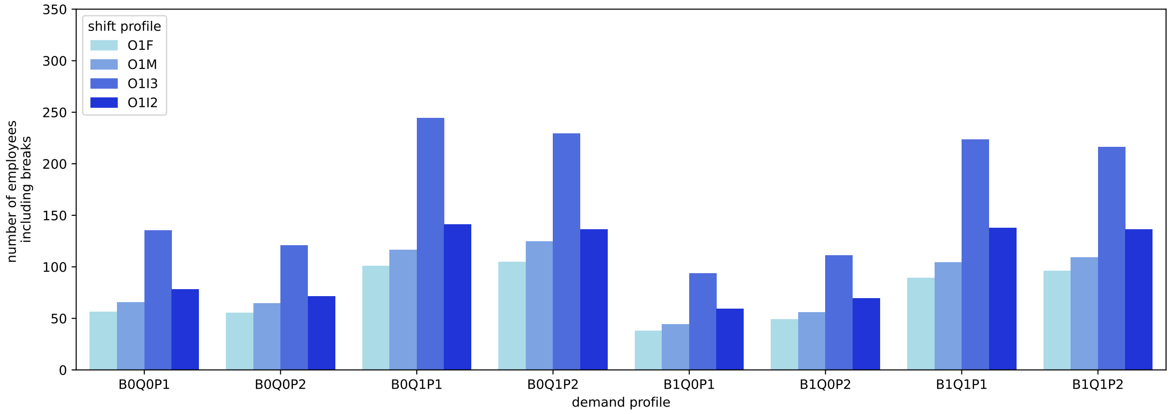
employees needed per setting, i.e., per demand and shift profile. As expected, instances with an underlying demand profile having high demand (*Q1*) require more staff on average, than instances of demand profiles with low demand (*Q0*). The number of peaks is of minor impact as instances require a similar number of employees regardless of the number of peaks (comparing *P1* vs. *P2* side by side). However, comparing settings with the same demand intensity and same number of peaks, instances with flexible blocks (*B1*) require slightly less staff than instances with inflexible blocks (*B0*).

Focusing on shift profiles, the charts show the highest demand of employees for inflexible shift systems where only two (*I2*) or three (*I3*) shifts are in the shift set. The more flexible the shift profile gets, the lower becomes the number of required employees. Semi-flexible profiles generally require fewer employees than inflexible shift profiles but more than flexible profiles. Comparing profiles excluding breaks (*O0*, Figure 7a) with profiles including breaks (*O1*, Figure 7b), the results' structure is similar, however with differing workforce sizes. Contrasting the effects of demand profiles with the effects of shift profiles, demonstrates that shift profiles with various scope of flexibility gravely affect the required workforce size. On the other hand, the occurrence of multiple demand blocks or the block's flexibility has less impact on the workforce size.

To densify the information about shift profiles from Figure 7 and further point out the prospects of flexible planning, Figure 8 presents the average percentage workforce reduction per demand profile. Therefore, the required number of employees in inflexible settings (*I2* and *I3*) is consolidated and the minimum value of both is taken as a reference. Then, the required average number of employees in inflexible settings is contrasted with the average workforce size in semi-flexible and flexible settings. Hence, enlarging the shift pool but limiting individual rosters to one shift type per employee and week, a workforce reduction by 1% up to 19% can be achieved if break planning is neglected as shown in Figure 8a. Including break planning these values increase and semi-flexible settings require 8% to 25% less staff than inflexible settings as depicted in



(a) settings excluding break planning

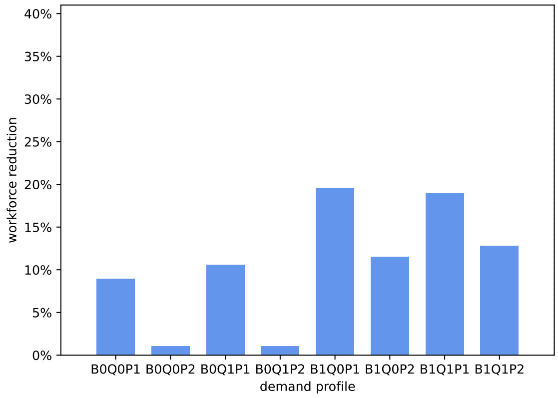


(b) settings including break planning

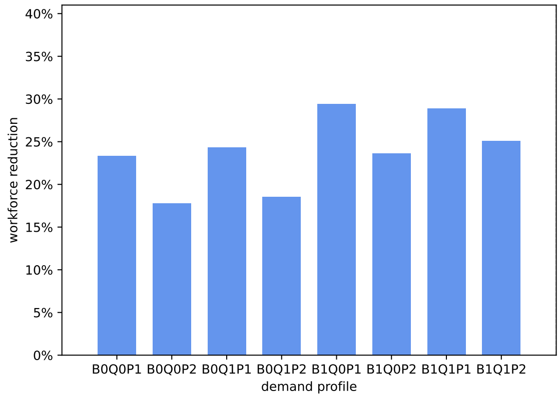
Figure 7: Average number of employees per setting (each bar corresponds to the average of 25 instances)

Figure 8c. Further possible savings are reached when considering completely flexible settings. Benchmarking inflexible settings against flexible settings, the workforce can be reduced by 17% to 29% without consideration of breaks, as given in Figure 8b. Finally, comparing inflexible and flexible settings with break planning in Figure 8d, the workforce lowers by 22% to 35% utilizing flexible options. A remarkable pattern in each plot of Figure 8 is that every second bar displays a similar value. These bars refer to demand profiles with the same time-related structure but with different demand intensity. Accordingly, the potential to save employees by enabling flexible shift systems is interrelated with the underlying demand profile. Across all plots, the lowest possible workforce reduction is identified with inflexible demand blocks and two peaks. In contrast, flexible shift scheduling options have the highest impact when the underlying flexible demand block has one peak.

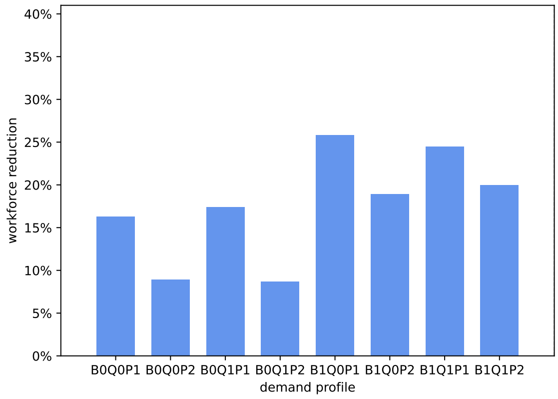
Figure 9 shows the workforce discrepancies solving instances with and without consideration of breaks. Accordingly, workforce sizes are underestimated if break assignments are neglected in the planning process, even with flexible shift systems. This may lead to challenges in covering demand adequately. The effect is especially serious for inflexible shift profiles with three shifts (*I3*), where significant staffing differences occur. A reason



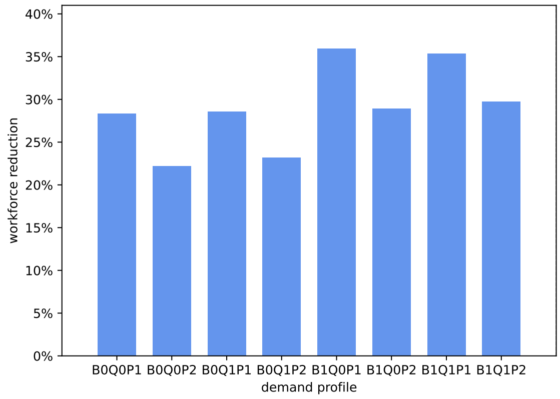
(a) inflexible vs. semi-flexible settings excluding break planning



(b) inflexible vs. flexible settings excluding break planning



(c) inflexible vs. semi-flexible settings including break planning



(d) inflexible vs. flexible settings including break planning

Figure 8: Percentage workforce reduction contrasting inflexible settings with flexibility options

for this might be the combination of two aspects: first, the time periods in which multiple shifts on the shift set overlap, and second, the values of parameters W^{pre} and W^{post}. The precise interdependencies of these factors open up a field for further follow-up studies.

Figure 10a and Figure 10b depict the average surplus ratio between the staffed employees' working hours and the instances' demand, i.e., $\frac{\sum_{i \in I} H^{\max} y_i}{\sum_{d \in D} \sum_{p \in P} N_{dp}}$. This metric provides further insights into the staffing levels and serves as an estimate for workforce needs in different settings. A ratio close to 1.0 (marked with dotted line) indicates that the staffing levels are closely aligned with the demand, implying optimal utilization of the workforce. Higher ratios may suggest potential inefficiencies from supposable overstaffing. The results show a similar structure across all demand profiles: the ratio is the lowest for flexible shift profiles, i.e., the most efficient staffing is reached with flexible shift scheduling, reaching ratios between 1.1 to 1.2. If breaks are ignored in the planning process, inflexible shift profiles reach an average factor of 1.3 up to 1.8 to cover demand in all periods. Including break assignments, these values increase, resulting in workforce size ratios ranging from 1.5 to 2.9. From a managerial perspective, the analysis of the metric provides useful information for staffing decisions, shift design, and scheduling

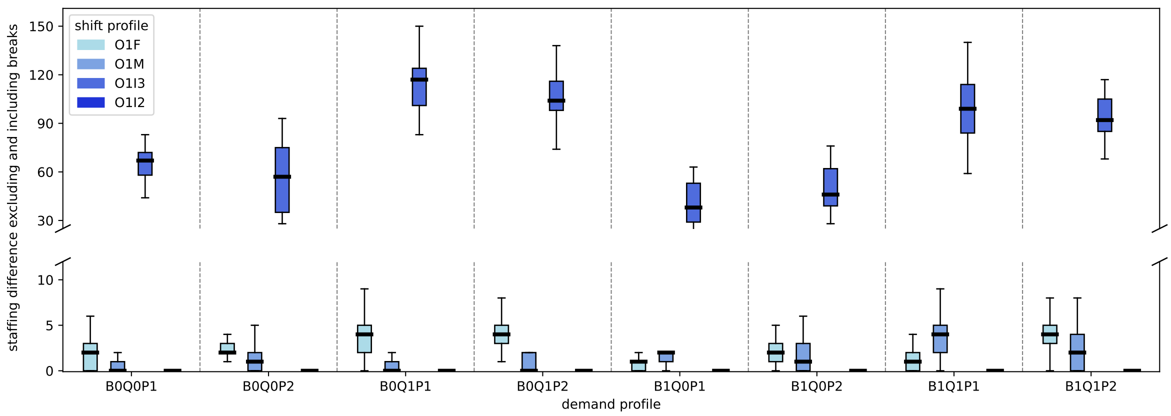


Figure 9: Discrepancy of staffing levels with consideration or neglect of breaks

policies. Multiplying the overall demand with the ratio provides an applicable staffing level for the respective department.

The evaluation of the study's results presents relevant information for hospital management. Different demand profiles have a slight impact on the workforce size, however, they affect the potential workforce reduction that may be achieved by utilizing flexible shift profiles. The analyses of resulting workforce sizes reveal that flexible shift profiles lower staffing requirements. Opposing inflexible to flexible shift profiles the workforce can be lowered by 17% to 35%, depending on the respective setting.

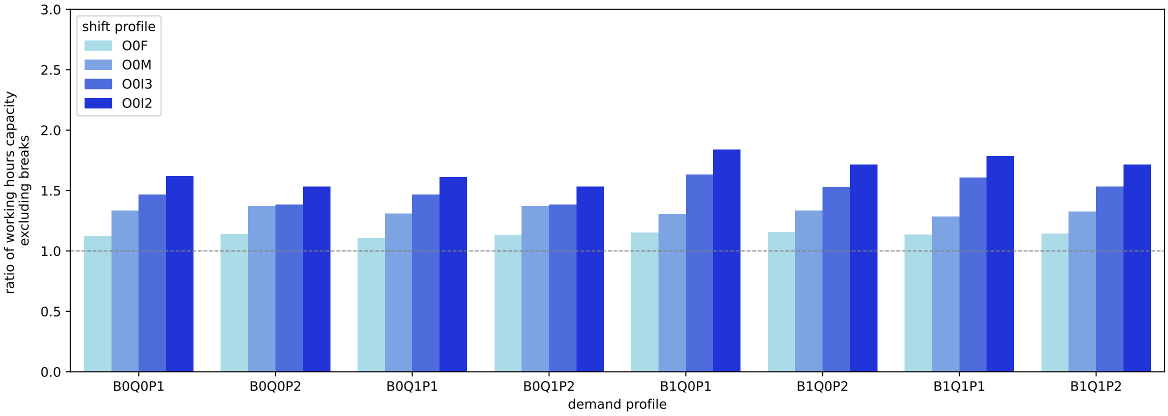
Including flexible break assignments in the solution process, significantly affects the workforce size for inflexible systems. Hence, the consideration of break assignments in staffing decisions could tackle the reduction of under-staffing in practical settings.

Using the weekly demand as a basis for staffing decisions, the evaluation provides key factors to estimate the required workforce sizes. The analysis of the key factors underlines that staff is scheduled and utilized more efficiently in flexible shift settings. This effect is especially strong if break planning is included into the staffing decision.

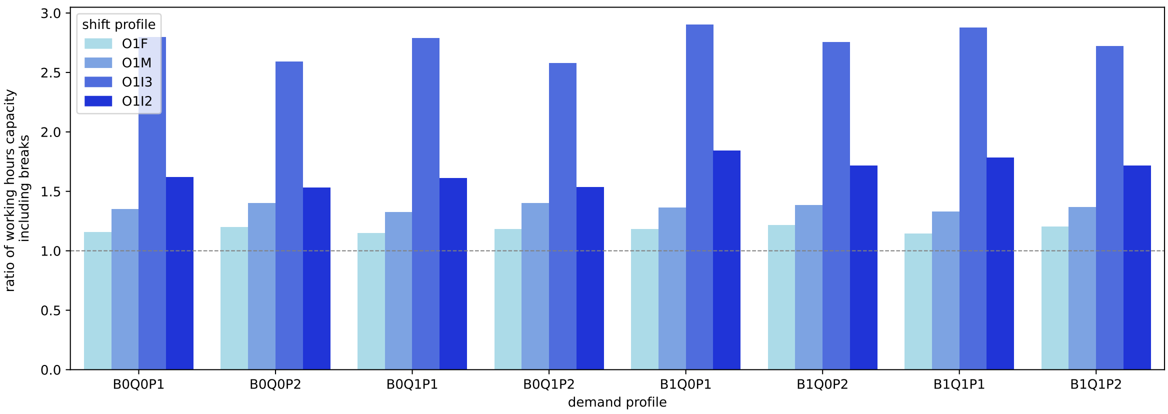
Consequently, hospitals face the trade-off of scheduling their employees to a clearly arranged but rigid pool of shifts or improving employee's utilization, and lower staffing cost.

6 Conclusion

We provide an IP model for staff scheduling based on flexible shifts and including flexible break assignment. With a focus on minimizing the workforce size, we aim to evaluate the effects of different settings, namely demand profiles and shift profiles. We decompose the model and give model reformulations to solve the problem efficiently using a column generation heuristic. Solving 1,600 instances, our solution approach provides near-optimal results taking run times of a few seconds on average.



(a) settings excluding break planning



(b) settings including break planning

Figure 10: Average ratio of surplus working capacity per setting (each bar corresponds to the average of 25 instances)

To generate universal insights on staffing decisions with flexible and volatile settings, we set up multiple settings for our computational experiments. Our study's instances are systematically varied, applying eight demand and eight shift profiles. The results show, that flexible shift profiles own the potential to significantly reduce the workforce size and should be considered in healthcare staffing decisions.

Further research in this context may analyze the effects of shift selection and shift design for such systems in detail. Additionally, future research should concentrate on integrating other kinds of flexibility and their effects on staffing levels. For instance, the employment of part-timers could generate benefits in the staffing process that enable better matching of supply and demand. Albeit a high level of flexibility in shift systems with break assignments leads to decreasing personnel costs, it is also necessary to take the wellbeing into account. Inserting too much flexibility in staff's schedule might have negative consequences on job motivation or individual workload. In current practice, hospitals have to find a middle course to match hospital's and personnel's interest.

References

[1] Álvarez, Esteban, Ferrer, Juan-Carlos, Muñoz, Juan Carlos and Henao, César Augusto: Efficient shift scheduling with multiple breaks for full-time employees: A retail industry case. In: *Computers & Industrial Engineering*, Volume 150:S. 106884, 2020. ISSN 03608352. doi:10.1016/j.cie.2020.106884.

[2] Bechtold, Stephen E. and Jacobs, Larry W.: Implicit modeling of flexible break assignments in optimal shift scheduling. In: *Management Science*, Volume 36(11):S. 1339–1351, 1990. ISSN 0025-1909. doi:10.1287/mnsc.36.11.1339.

[3] Benazzouz, Touria, Echchatbi, Abdelwahed and Bellabdaoui, Adil: A literature review on the nurses' planning problems. In: *International Journal of Mathematics and Computational Science*, Volume 1:S. 268–274, 2015.

[4] Bölt, Ute: Statistische krankenhausdaten - grunddaten der krankenhäuser 2022. In: Klauber, Jürgen, Wasem, Jürgen, Beivers, Andreas, Mostert, Carina and Scheller-Kreinsen, David (Ed.) *Krankenhaus-Report 2024*, Springer Berlin Heidelberg, Berlin, Heidelberg, S. 437–464. 2024. ISBN 978-3-662-68791-8.

[5] Böðvarsdóttir, Elín Björk, Bagger, Niels-Christian Fink, Høffner, Laura Elise and Stidsen, Thomas J. R.: A flexible mixed integer programming-based system for real-world nurse rostering. In: *Journal of Scheduling*, Volume 25(1):S. 59–88, 2022. ISSN 1094-6136. doi:10.1007/s10951-021-00705-7.

[6] Brunner, Jens O., Bard, Jonathan F. and Kolisch, Rainer: Flexible shift scheduling of physicians. In: *Health Care Management Science*, Volume 12(3):S. 285–305, 2009. ISSN 1572-9389. doi:10.1007/s10729-008-9095-2.

[7] Brunner, Jens O., Bard, Jonathan F. and Kolisch, Rainer: Midterm scheduling of physicians with flexible shifts using branch and price. In: *IIE Transactions*, Volume 43(2):S. 84–109, 2010. ISSN 0740-817X. doi:10.1080/0740817X.2010.504685.

[8] Burke, Edmund K., de Causmaecker, Patrick, Berghe, Greet V. and van Landeghem, Hendrik: The state of the art of nurse rostering. In: *Journal of Scheduling*, Volume 7(6):S. 441–499, 2004. ISSN 1094-6136. doi:10.1023/B:JOSH.0000046076.75950.0b.

[9] Camiat, Fanny, Restrepo, Marìa I., Chauny, Jean-Marc, Lahrichi, Nadia and Rousseau, Louis-Martin: Productivity-driven physician scheduling in emergency departments. In: *Health systems (Basingstoke, England)*, Volume 10(2):S. 104–117, 2019. ISSN 2047-6965. doi:10.1080/20476965.2019.1666036.

[10] de Causmaecker, Patrick and Vanden Berghe, Greet: A categorisation of nurse rostering problems. In: *Journal of Scheduling*, Volume 14(1):S. 3–16, 2011. ISSN 1094-6136. doi:10.1007/s10951-010-0211-z.

[11] Erhard, Melanie J. P.: Flexible staffing of physicians with column generation. In: *Flexible Services and Manufacturing Journal*, Volume 33(1):S. 212–252, 2021. doi:10.1007/s10696-019-09353-8.

[12] Erhard, Melanie J. P., Schoenfelder, Jan, Fügener, Andreas and Brunner, Jens O.: State of the art in physician scheduling. In: *European Journal of Operational Research*, Volume 265(1):S. 1–18, 2018. ISSN 03772217. doi:10.1016/j.ejor.2017.06.037.

[13] Fuchs, Michaela and Weyh, Antje: Die arbeitsmarktsituation in krankenhäusern. In: Klauber, Jürgen, Wasem, Jürgen, Beivers, Andreas and Mostert, Carina (Ed.) *Krankenhaus-Report 2023*, Springer Berlin Heidelberg, Berlin, Heidelberg, S. 33–47. 2023. ISBN 978-3-662-66880-1.

[14] Fügener, Andreas and Brunner, Jens O.: Planning for overtime: The value of shift extensions in physician scheduling. In: *INFORMS Journal on Computing*, Volume 31(4):S. 732–744, 2019. ISSN 1091-9856. doi:10.1287/ijoc.2018.0865.

[15] Fügener, Andreas, Brunner, Jens O. and Podtschaske, Armin: Duty and workstation rostering considering preferences and fairness: a case study at a department of anaesthesiology. In: *International Journal of Production Research*, Volume 53(24):S. 7465–7487, 2015. ISSN 0020-7543. doi:10.1080/00207543.2015.1082667.

[16] Gross, Christopher Nikolaus, Brunner, Jens O. and Blobner, Manfred: Hospital physicians can't get no long-term satisfaction – an indicator for fairness in preference fulfillment on duty schedules. In: *Health Care Management Science*, Volume 22(4):S. 691–708, 2019. ISSN 1572-9389. doi:10.1007/s10729-018-9452-8.

[17] Guo, Jia and Bard, Jonathan F.: A column generation-based algorithm for midterm nurse scheduling with specialized constraints, preference considerations, and overtime. In: *Computers & Operations Research*, Volume 138:S. 105597, 2022. ISSN 03050548. doi:10.1016/j.cor.2021.105597.

[18] Isken, Mark W. and Aydas, Osman T.: A tactical multi-week implicit tour scheduling model with applications in healthcare. In: *Health Care Management Science*, Volume 25(4):S. 551–573, 2022. ISSN 1572-9389. doi:10.1007/s10729-022-09601-8.

[19] Kiermaier, Ferdinand, Frey, Markus and Bard, Jonathan F.: The flexible break assignment problem for large tour scheduling problems with an application to

airport ground handlers. In: *Journal of Scheduling*, Volume 23(2):S. 177–209, 2020. ISSN 1094-6136. doi:10.1007/s10951-019-00635-5.

[20] Kraul, Sebastian, Erhard, Melanie J. P., Brunner, Jens O. and Erhard, Melanie: Optimizing physician schedules with resilient break assignments. In: *Omega*, Volume 129:S. 103–154, 2024. ISSN 03050483. doi:10.1016/j.omega.2024.103154.

[21] Lim, Gino J., Mobasher, Arezou, Bard, Jonathan F. and Najjarbashi, Amirhossein: Nurse scheduling with lunch break assignments in operating suites. In: *Operations Research for Health Care*, Volume 10:S. 35–48, 2016. ISSN 2211-6923. doi:10.1016/j.orhc.2016.07.001.

[22] Lübbecke, Marco E. and Desrosiers, Jacques: Selected topics in column generation. In: *Operations Research*, Volume 53(6):S. 1007–1023, 2005. ISSN 0030-364X. doi:10.1287/opre.1050.0234.

[23] Mattia, Sara, Rossi, Fabrizio and Smriglio, Stefano: Robust shift scheduling with discretionary rest breaks. In: , 2023. doi:10.2139/ssrn.4666608.

[24] Meyer, Markus, Meinicke, Moritz and Schenkel, Antje: Krankheitsbedingte fehlzeiten in der deutschen wirtschaft im jahr 2022. In: Badura, Bernhard, Ducki, Antje, Baumgardt, Johanna, Meyer, Markus and Schröder, Helmut (Ed.) *Fehlzeiten-Report 2023*, Springer Berlin Heidelberg, Berlin, Heidelberg, S. 435–520. 2023. ISBN 978-3-662-67513-7.

[25] Ngoo, Chong Man, Goh, Say Leng, Sze, San Nah, Sabar, Nasser R., Abdullah, Salwani and Kendall, Graham: A survey of the nurse rostering solution methodologies: The state-of-the-art and emerging trends. In: *IEEE Access*, Volume 10:S. 56504–56524, 2022. doi:10.1109/ACCESS.2022.3177280.

[26] Pandey, Peeyush, Gajjar, Hasmukh and Shah, Bhavin J.: Determining optimal workforce size and schedule at the retail store considering overstaffing and understaffing costs. In: *Computers & Industrial Engineering*, Volume 161:S. 107656, 2021. ISSN 03608352. doi:10.1016/j.cie.2021.107656.

[27] Statistisches Bundesamt (Ed.): Grunddaten der krankenhäuser 2022. 2023. URL `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankenhaeuser/_inhalt.html`.

[28] Statistisches Bundesamt (Ed.): Kostennachweis der krankenhäuser 2022. 2023. URL `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Krankenhaeuser/_inhalt.html`.

[29] Stolletz, Raik and Brunner, Jens O.: Fair optimization of fortnightly physician schedules with flexible shifts. In: *European Journal of Operational Research*, Volume 219(3):S. 622–629, 2012. ISSN 03772217. doi:10.1016/j.ejor.2011.10.038.

[30] Tiedemann, Jurek, Kunath, Gero and Werner, Dirk: Jahresrückblick 2023 - rückgang der fachkräftelücke, aber keine entspannung. 2024.

[31] Witkoski, Amy and Dickson, Victoria Vaughan: Hospital staff nurses' work hours, meal periods, and rest breaks. In: *AAOHN Journal*, Volume 58(11):S. 489–497, 2010. ISSN 0891-0162. doi:10.1177/216507991005801106.

[32] Xu, Shuling and Hall, Nicholas G.: Fatigue, personnel scheduling and operations: Review and research opportunities. In: *European Journal of Operational Research*, Volume 295(3):S. 807–822, 2021. ISSN 03772217. doi:10.1016/j.ejor.2021.03.036.

Appendix C.

Covidal

Bartenschlager CC, Ebel SS, Kling S, Vehreschild J, Zabel LT, Spinner CD, Schuler A, Heller AR, Borgmann S, Hoffmann R, Rieg S, Messmann H, Hower M, Brunner JO, Hanses F, Römmele C (2022): **COVIDAL: A machine learning classifier for digital COVID-19 diagnosis in German hospitals.**

Status: published in ACM Transactions on Management Information Systems; Category B.

The final version of the contribution and its supplementary material is available online at https://doi.org/10.1145/3567431.