Data-Driven Hospital Scheduling: A Multi-criteria Evaluation of Algorithmic Approaches

Rafael Amorim, Afonso Lobo, João Lopes(✉), Tiago Guimarães, Júlio Duarte, and Manuel Filipe Santos

ALGORITMI/LASI Research Centre, University of Minho, Braga, Portugal
jlopes@dsi.uminho.pt

Abstract. It is critical in a health care institution to efficiently schedule medical assessments considering resources, clinical priorities and operational limits, and create a delicate balance to deliver efficient, timely and quality care. This work studies the medical exam scheduling problem using real data from Unidade Local de Saúde de Santo António (ULSSA), which has to allocate exams into operational time slots, with multiple competing constraints. An assessment of three optimization methods, namely Mixed-Integer Linear Programming (MILP), Particle Swarm Optimization (PSO) and Genetic Algorithms (GA) is presented. The proposed solutions are evaluated through a multi-criteria objective function that includes penalties for patient waiting time, deadline violations, priority non-compliance and idle time in scheduling blocks. This formulation allows for a realistic expression of operational-clinical trade-offs of the hospital. The results indicate that GA achieves the best overall performance among the evaluated methods, obtaining the lowest total penalty value and demonstrating superior ability to balance solution quality with scheduling flexibility under the defined multi-criteria objective function. In addition to optimizing the best-found solution, clustering analysis is performed on it to understand any hidden patterns in examination sequencing, specifically in the occurrence of consecutive examinations possessing the same attributes. Overall, the results highlight the usefulness of coupling optimization methods with data-driven pattern assessment to improve hospital scheduling systems. It will help in making decision-support systems in healthcare operations more adaptive, effective and clinically optimized.

Keywords: Medical Examination Scheduling · Healthcare Optimization

1 Introduction

Efficient scheduling of clinical services represents a critical challenge in modern healthcare systems. Hospital scheduling systems are essential for managing

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16761, pp. 619–630, 2027.
https://doi.org/10.1007/978-3-032-30527-5_40

patient access to medical professionals and diagnostic resources, playing a central role in ensuring timely care delivery and the efficient utilization of available capacity [1]. However, the inherent variability in service durations, uncertainty associated with patient behavior, and the limited availability of healthcare resources significantly increase the complexity of scheduling decisions [1,2]. As a result, healthcare scheduling has been extensively studied within the field of operations research [3,4].

Hospital scheduling problems typically involve multiple and often conflicting operational constraints, including patient prioritization, heterogeneous service times, resource availability, and shift-related limitations. These characteristics render the problem highly complex, frequently leading to formulations based on combinatorial optimization [5]. To address this complexity, a wide range of solution approaches has been proposed, including exact mathematical programming methods and metaheuristic techniques tailored to healthcare service planning [3,6]. Despite the significant body of research in this domain, several challenges remain. In particular, there is a lack of systematic comparative analyses of exact and metaheuristic approaches under realistic multi-criteria formulations that simultaneously capture operational efficiency and clinical priorities. Furthermore, limited attention has been given to the integration of optimization outputs with data-driven analytical techniques capable of extracting actionable patterns from scheduling solutions.

In this context, this study addresses the medical examination scheduling problem in a real hospital setting by developing and evaluating three optimization approaches: Mixed-Integer Linear Programming (MILP), Particle Swarm Optimization (PSO), and Genetic Algorithms (GA) [8,11,12]. The proposed solutions are assessed using a multi-criteria objective function that incorporates key operational constraints, including waiting times, deadline adherence, priority compliance, and resource utilization. Additionally, this work extends beyond optimization by incorporating a post-optimization clustering analysis on the best-performing solution, aiming to identify latent patterns in exam sequencing and scheduling structures. This combined approach contributes to the development of more informed and data-driven decision-support mechanisms for healthcare operations, bridging optimization techniques with knowledge discovery processes.

This paper is organized as follows. Section 2 provides the background, covering the scheduling process, a review of related work, and the identification of operational gaps in the literature. Section 3 describes the three optimization approaches proposed in this study. Section 4 outlines the Design Science Research methodology followed. Section 5 formally defines the problem and the multi-criteria objective function. Section 6 presents and discusses the results, including the multi-criteria evaluation, the comparison with the hospital's current scheduling process, the clustering analysis, and the main contributions of the proposed approach. Finally, Sect. 7 draws the conclusions and outlines directions for future research.

2 Background

2.1 Scheduling Process

The problem involves assigning exam requests to available time slots. Each exam is characterized by attributes relevant to the decision-making process, such as the request creation date, the estimated duration, the priority level, the exam description, and the deadline for completion. The time structure consists of a single time slot per day, within a fixed daily operating window. The problem was thus modeled as a process of allocating exams to daily shifts with limited time capacity. Solutions are constructed from an ordered sequence of exams, which determines the order in which they are considered for allocation, always ensuring that the total duration of the exams does not exceed the capacity of the shift in which they are scheduled. To evaluate the quality of the generated solutions, operational criteria are considered, such as patient wait times, delays relative to the deadline, clinical priorities, efficient use of shifts, and operational continuity between exams with similar descriptions.

2.2 Related Work

The operations research literature has seen substantial studies on healthcare scheduling problems because of their impact on service efficiencies, resource utilization, and patient waiting time [1]. These problems exist in almost every hospital environment such as outpatient scheduling, surgical planning and diagnostic examination allocation. Each of them exhibits a different level of complexity thanks to the clinical and operational constraints involved [5,7].

As various type of optimization approaches have been devised for these challenges. Methods based on mathematical programming, particularly Mixed-Integer Linear Programming (MILP), facilitate the explicit formulation of operational constraints and optimization goal, which are well suited for structured problems [3]. Nevertheless, these strategies exhibit important computational restrictions that restrict their scalability and usability when the difficulty and dimensions of the problem rise [6].

Heuristic and metaheuristic methods have become an extensively researched alternative to address large and complex scheduling problems. These techniques allow for an efficient search within large solution spaces and an ability to achieve good quality solutions within acceptable computational times. In particular multi-constraint and multi-objective environments [3,4]. Nature-inspired algorithms and population-based techniques, like Particle Swarm Optimization and Genetic Algorithms, have made remarkable strides in healthcare scheduling [8].

Also, in the scheduling literature, the definition of job sets containing a certain similarity in characteristics is partly used for batching, sequence-dependent scheduling where the order of execution affects the systems' performance [6]. Hospital environments are especially affected by this since tests that are similar may be grouped together to make operations easier and to cut down on set-up or gap times.

2.3 Operational Gaps

Despite the significant number of existing studies, some relevant limitations persist in the literature on healthcare scheduling. Many studies focus on the application of a single optimization method, making it difficult to directly compare different approaches when applied to the same problem [3]. Similarly, a considerable portion of the studies relies on simplified scenarios or simulated data, which do not always adequately capture the complexity and operational constraints present in real hospital settings [5,6]. There is also limited exploration of the structural patterns present in the generated schedules. In particular, the relationship between the criteria defined in the objective function and the organization of exams over time remains largely unexplored. It is therefore important not only to compare different optimization approaches to the same problem using real data, but also to analyze, in a complementary manner, the patterns emerging in the generated solutions, allowing for a more detailed interpretation of each model's behavior in scheduling problems.

3 Proposed Algorithms

This study aimed to evaluate three different optimization approaches: Mixed-Integer Linear Programming (MILP), Genetic Algorithm (GA), and Particle Swarm Optimization (PSO). The MILP method relate to an exact method that is based on mathematical programming. While GA and PSO are population-based metaheuristics that are widely used to solve complex combinatorial optimization problems [8,11,12].

All methodologies share a common representation for solutions. A candidate solution is constructed as an ordered sequence of exam requests defining the order in which exams are considered for allocation. Next, a decoding procedure is implemented to convert the obtained sequence into a feasible schedule by assigning examinations to daily time slots sequentially within capacity constraints. In particular, each exam is assigned to the first available time slot that has enough available capacity, and where the total bookable duration does not exceed the shift limit. Different optimization algorithms can try out other permutations of exam sequences in the same search space. Thus using the same representation makes it consistent.

All methods optimize the same multi-criteria objective functions capturing key operational and clinical aspects, such as patient waiting times, deadline adherence, priority compliance and resource utilization. This objective function allows for a fair comparison between exact and meta-heuristic methods, which have entirely distinct characteristics.

The MILP model provides a mathematical description of the scheduling problem and its respective decision variable and constraints that characterizes their allocation. The HiGHS solver is used to solve the model with a time limit of 60 s and a tolerance for relative optimality gap of 10% for a balance between solution quality and computational tractability.

The GA has a population of 80 individuals for 12 generations, with a crossover probability of 0.9 and a mutation probability of 0.2. Tournament selection with size 3 is used, as well as elitism that maintains the best two individuals per generation. The Particle Swarm Optimization algorithm involved a population size of 100 particles, while performing in 40 iterations, utilizing inertia weight of 0.65, cognitive coefficient and social coefficient are set as 1.4. To ensure a proper balance between exploring the search space and identifying the optimal solution, these parameters were adjusted as the models were tested and implemented.

4 Methodologies

This study follows a Design Science Research (DSR) approach [9,10], which provides the overarching methodological framework guiding the research process. DSR is particularly well suited for this work as it focuses on the design and evaluation of artefacts (in this case, computational models) intended to solve real-world problems in organizational contexts such as healthcare operations.

The research process is structured around three DSR stages that map directly onto the contributions of this paper. In the problem identification stage, the scheduling problem is defined using real hospital data from ULSSA, and operational constraints are elicited with the support of medical professionals. This grounds the artefact design in authentic clinical and organizational requirements rather than simplified assumptions.

In the artefact design stage, three optimization approaches are developed under a unified problem formulation and shared objective function. This design choice ensures that differences in performance are attributable to the algorithms themselves, rather than to inconsistencies in problem representation.

Finally, in the evaluation stage, the artefacts are assessed through computational experiments, compared against each other and against the hospital's current scheduling practice. Each algorithm is executed 30 times to account for the stochastic nature of metaheuristic methods, with results reported as mean and standard deviation. This evaluation closes the DSR cycle by producing actionable knowledge about which algorithmic approaches best address the identified problem.

5 Problem Definition

Each test is defined by attributes relevant to decision-making, which were identified with the help of medical professionals who supported the strategic approach to this problem, identifying the expected duration, clinical priority level, request creation date, and completion deadline as the most important parameters. Following an initial analysis of the data, a time structure was defined consisting of a single daily shift, corresponding to the operating period between 8:00 a.m. and 11:00 p.m., during which a higher volume of exams was historically recorded. In addition to capacity constraints, it was determined that the scheduling process must take into account multiple operational factors, including patient wait times,

delays relative to deadlines, adherence to clinical priorities, and efficient use of available capacity. Furthermore, operational continuity is considered between consecutively scheduled exams with similar characteristics.

The problem can therefore be formulated as a combinatorial optimization problem, where different permutations of examination sequences lead to distinct scheduling solutions. The quality of each solution is evaluated through a multi-criteria objective function:

$$\begin{aligned} Z = {} & w_1 \textstyle\sum_{i \in A} \max(0, W_i) + \\ & w_2 \textstyle\sum_{i \in A} \max(0, D_i) + \\ & \textstyle\sum_{i \in A} \left(w_3 N_i^{(1)} + w_4 N_i^{(\geq 2)} \right) + \\ & w_5 \textstyle\sum_{p \in P} C_p - w_6 \sum_{p \in P} B_p + \\ & w_7 \textstyle\sum_{p \in P} I_p + w_8 \sum_{p \in P} I_p \, \delta \end{aligned}$$

where A denotes the set of scheduled examinations and P the set of available time slots (shifts).

- W_i: waiting time of examination i;
- D_i: delay of examination i relative to its deadline;
- $N_i^{(1)}$: number of pending examinations with a higher priority level by one level;
- $N_i^{(\geq 2)}$: number of pending examinations with a higher priority level by two or more levels;
- C_p: number of type changes between consecutive examinations in shift p;
- B_p: number of consecutive pairs of examinations of the same type in shift p;
- I_p: idle time in shift p;
- δ: binary indicator equal to 1 if there exist unscheduled examinations, and 0 otherwise.

The objective is to minimize Z, reducing waiting times, delays, priority violations, and idle time, while promoting the grouping of examinations with similar characteristics within each shift.

6 Results

6.1 Multi-criteria Evaluation

The optimization approaches were compared using the formulating multi-criteria objective function which takes into consideration important operational aspects of the scheduling process, including patient waiting time, delayed deadline, prioritized patient compliance, and shift capacity utilization.

The findings demonstrate that the methods exhibit a significant difference in performance, which leads to different quality solutions. The objective function values indicate the capability of the models to minimize operational inefficiencies without being in contradiction to clinical priorities.

Figure 1 presents the total penalty values obtained by the three evaluated models: Genetic Algorithms (GA), Mixed-Integer Linear Programming (MILP,

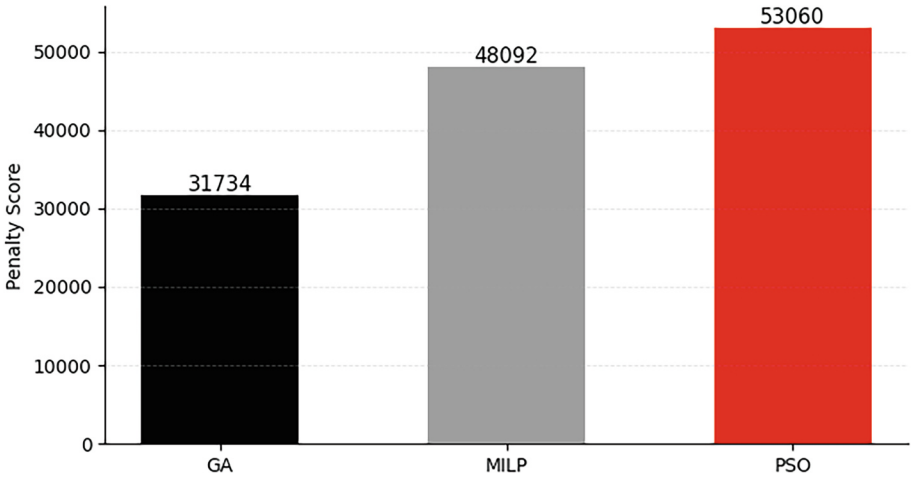

Fig. 1. Comparison of total penalty values achieved by the different optimization models under the defined multi-criteria objective function.

and Particle Swarm Optimization (PSO). As the objective function aggregates multiple operational criteria, lower penalty values indicate more effective scheduling solutions according to the defined objectives.

The results show that GA achieves the best performance, with the lowest total penalty of all the approaches we analyzed, due to its ability to explore all possible solutions and adjust the order of tests to fit the specific constraints of the problem. This flexibility is fundamental and is what allowed it to achieve optimal performance across all the different criteria we considered. Compared to the other methods, the advantage of GA in handling the problem's complexities and finding the best solution is evident. The MILP and PSO models exhibited higher penalty values, meaning they were not as effective at satisfying all conditions of the objective function simultaneously, leading to the conclusion that, for this specific scheduling problem and given the size of the dataset and the identified constraints, GA is more effective at striking a balance between solution quality and scheduling flexibility.

Based on these results, the GA model was selected for further comparison with the scheduling process currently used in the hospital.

6.2 Hospital Solution Comparison

A practical comparison between the GA-based schedule and the hospital's existing scheduling practice was carried out to examine the real-world implications of the proposed approach. Both were evaluated on the same dataset, covering the three-month period from October to December 2021.

Restricting the analysis to examinations performed within the model's operational window—08:00 to 23:00—was a necessary condition for the comparison

to hold methodological ground. Without this constraint, differences in temporal scope would confound any direct assessment of the two approaches' relative performance.

Table 1. Performance comparison between the optimization model (GA) and the hospital scheduling process

Metric	GA Model	Hospital Scheduling
Penalty score	31734	787635.5
Consecutive identical exam descriptions (%)	13.07	12.60
Scheduled examinations	3382	7488
Unscheduled examinations	6648	1781
Scheduled priority 0 (%)	0	33.48
Scheduled priority 1 (%)	0	32.20
Scheduled priority 2 (%)	100	34.32
Scheduled after deadline (%)	0	0
Scheduled before or on deadline (%)	100	100

Table 1 summarises the global comparison between the two approaches. The GA model achieves a substantially lower total penalty—a gap that reflects, above all, a closer alignment with the criteria encoded in the multi-criteria objective function.

One result that warrants attention is the higher absolute number of examinations scheduled by the hospital. This is, in fact, expected: the hospital operates through multiple concurrent teams and resources, whereas the proposed model assumes sequential allocation within a single daily shift, which inherently caps overall throughput. The difference in scheduling volume is therefore a structural consequence of these distinct operating assumptions, not an indicator of inferior optimisation performance.

It is identified that the two approaches differ in how they handle priorities. The hospital distributes exams across the different priority levels in a relatively balanced manner, while the GA model schedules exclusively high-priority cases. This behavior results directly from the penalty structure of the objective function, in which delaying or failing to assign urgent exams incurs disproportionately high costs, which guides the model toward a systematic prioritization of the most critical requests. Regarding the sequence of exams, the GA model produces a slightly higher proportion of consecutive exams with identical descriptions. An efficiency gain is observed, modest in absolute terms but consistent with the model's optimization goal of grouping similar exams within a shift, reducing costs associated with switching between exam types.

Figure2 shows the weekly trend of this metric throughout the entire analysis period. Expressing the values as percentages distinguishes the comparison from variations in the absolute number of exams, providing a clearer picture of how

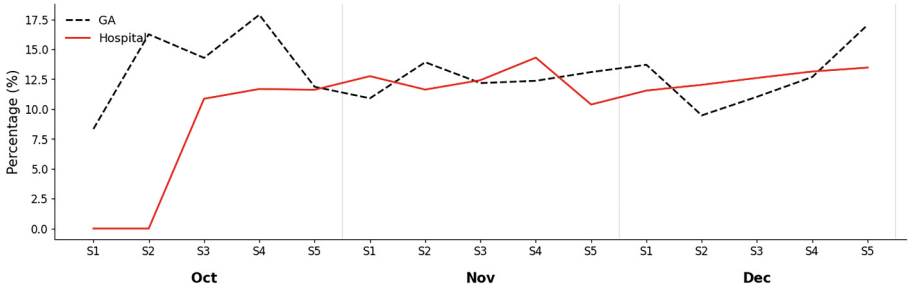

Fig. 2. Comparison of the weekly percentage of consecutively scheduled exams by description.

the metric has changed over time. Across most weeks, the GA model achieves values comparable to those recorded in the hospital scheduling process. Taken together, these results suggest that the proposed approach not only improves quantitative outcomes but also produces schedules with greater structural regularity, exploiting patterns in examination sequencing in a way the hospital's current process does not systematically attempt.

6.3 Clustering Analysis

Using K-Means clustering technique with k = 8 (determined using elbow method), k-means clustering was applied on unscheduled and scheduled examination data separately for the period OctoberâĂŞDecember 2021 using four normalized features: examination duration, priority, request creation time, and time window up to deadline. Table 2 summarizes the distinguishing profile of each group.

Table 2. Summary of K-Means clustering results for scheduled and unscheduled examinations ($k = 8$).

Group	Modality	Duration (min)	Priority	Temporal Window
Scheduled examinations				
S1–S4	CT	15–20	High (P2)	Favorable
S5–S8	MRI	40–45	High (P2)	Favorable
Unscheduled examinations				
U1–U4	Mixed	40–45	Low (P0–P1)	Tight or expired
U5–U8	Mixed	40–45	Low (P0–P1)	Less favorable

Scheduled clusters exhibit the same compositions of modality (basically CT and MRI), and shorter duration of examinations. The objective function

supports this behaviour by giving higher incentive to bunch similar examination types and also lower operational transitions. Unscheduled clusters have longer duration, lesser priority and worse time windows which makes them non-competitive for just those reasons as per the criteria. These results illustrate that the objective function governs scheduling behavior, and the output of clustering provides a clear and data-driven explanation of allocation.

6.4 Scheduling Approach and Main Contribution

The scheduling model which is proposed is based on sequential representation of the requests of examination and every solution corresponds to ordering of processing. A constructive procedure assigns examinations to free shifts subject to temporal capacity constraints to transform this sequence into a feasible schedule.

The quality of solution is assessed against a multi-criteria objective function inclusive of some of the main operational characteristics in a hospital environment, such as patient waiting time, deadline compliance, priority compliance, and capacity utilization. Through this framework, we can evaluate scheduling performance through multiple lenses. With real hospital data and experimental conditions, the proposed method is a distinctive approach to compare various optimization paradigms. In particular, both exact and metaheuristic methods are evaluated in the same decision framework. Additionally, this study goes beyond the conventional quantitative evaluation as clustering techniques were harnessed to assess structural patterns in the schedules generated. This further analysis then allows for a connection between the objective function design and observed scheduling behaviour.

This work offers three major contributions. First, it systematically evaluates different optimization approaches according to the case of a real-world hospital scheduling problem. Second, it formulates a multi-criteria objective function consistent with operational and clinical constraints. Lastly, it performs clustering-based analysis for insights into scheduling and model behaviour. As a whole, these contributions allow for performance evaluation and the explainability of healthcare scheduling systems.

7 Conclusion

Three algorithmic approaches were developed under a unified problem formulation and assessed through a shared multi-criteria objective function, enabling a systematic comparison between exact and metaheuristic methods. The results show that the GA achieves the best overall performance under the defined objective function, recording the lowest total penalty value and demonstrating a stronger capacity to simultaneously balance multiple operational and clinical criteria. This outcome suggests that population-based metaheuristic approaches are well suited to complex scheduling problems shaped by competing constraints, settings where exact methods tend to face computational limitations and metaheuristics such as PSO may struggle to consistently exploit the structure of the

solution space. The comparison with the hospital's current scheduling practice further supports the value of the proposed approach. Although the GA model operates under more restrictive assumptions (most notably, sequential allocation within a single daily shift), it achieves substantially lower penalty scores and yields schedules with greater structural regularity, particularly in the grouping of consecutive exams sharing similar characteristics. Taken together, these findings point to the practical potential of optimization-based tools to strengthen both operational efficiency and clinical priority compliance in real hospital settings.

The complementary clustering analysis offered further insight into how the objective function shapes scheduling behaviour. Examinations that were successfully scheduled consistently exhibited high priority levels, shorter durations, and favorable time windows, whereas unscheduled ones clustered around lower priorities, longer durations, and tighter or expired deadlines. This perspective confirms that the objective function effectively steers allocation decisions and provides a transparent account of the model's behaviour, adding to the overall interpretability of the scheduling system. Despite these contributions, the model rests on assumptions that constrain its direct applicability to real hospital operations. Most notably, the single-shift sequential allocation structure does not capture the concurrent use of multiple clinical teams and resources that is typical of real hospital environments. Addressing this limitation is the most pressing direction for future work.

In the short term, several extensions could increase the model's realism and robustness. The framework should be adapted to accommodate multiple concurrent shifts and parallel resource allocation, better reflecting the operational structure of hospital departments. A sensitivity analysis of the objective function weights would also be valuable, helping to clarify how different prioritization schemes influence scheduling outcomes and to support more balanced decisions across clinical priority levels. Beyond this, testing the proposed models on larger datasets and across different hospital services or exam types would help establish how broadly the findings generalize. Looking further ahead, more substantial developments become possible. The incorporation of dynamic and real-time scheduling capabilities would allow the model to handle the kinds of disruptions (cancellations, urgent requests, unexpected resource unavailability) that are routine in clinical practice but absent from the current static formulation. Introducing uncertainty in exam durations and patient arrival patterns, through stochastic or robust optimization, would bring the model closer to the variability that characterizes real healthcare operations. These extensions, taken together, would move the proposed framework toward a more adaptive and operationally grounded decision-support tool for hospital scheduling.

Acknowledgments. This work has been supported by FCT âĂŞ Fundação para a Ciência e Tecnologia within the R&D Unit Project Scopes UID/00319/2025 - Centro ALGORITMI (ALGORITMI/UM) and 2024.07481.IACDC.

References

1. Gupta, D., Denton, B.: Appointment scheduling in health care: challenges and opportunities. IIE Trans. **40**(9), 800–819 (2008). https://doi.org/10.1080/07408170802165880
2. Hall, R.: Handbook of Healthcare System Scheduling. Springer, Cham (2012). https://doi.org/10.1007/978-1-4614-1734-7
3. Abdalkareem, Z.A., Amir, A., Al-Betar, M.A., Ekhan, P., Hammouri, A.I.: Healthcare scheduling in optimization context: a review. Heal. Technol. **11**(3), 445–469 (2021). https://doi.org/10.1007/s12553-021-00547-5
4. Ala, A.L., Al-Fuqaha, A.A., Abu-Alhaj, M.M.: Outpatient appointment scheduling: a systematic review. Healthcare **10**(9), 1732 (2022). https://doi.org/10.3390/healthcare10091732
5. Hulshof, P.J.H., Kortbeek, N., Boucherie, R.J., Hans, E.W., Bakker, P.J.M.: Taxonomic classification of planning decisions in health care: a structured review of the state of the art in OR/MS. Health Syst. **1**(2), 129–175 (2012). https://doi.org/10.1057/hs.2012.18
6. Cardoen, B., Demeulemeester, E., Beliën, J.: Operating room planning and scheduling: a literature review. Eur. J. Oper. Res. **201**(3), 921–932 (2010). https://doi.org/10.1016/j.ejor.2009.04.011
7. Ahmadi-Javid, A., Jalali, Z., Klassen, K.: Outpatient appointment systems in healthcare: a review of optimization studies. Eur. J. Oper. Res. **258**(1), 3–34 (2017). https://doi.org/10.1016/j.ejor.2016.06.064
8. Squires, M., Tao, X., Elangovan, S., Gururajan, R., Zhou, X., Acharya, U.: A novel genetic algorithm-based system for the scheduling of medical treatments. Expert Syst. Appl. **195**, 116464 (2022). https://doi.org/10.1016/j.eswa.2021.116464
9. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. **28**(1), 75–105 (2004)
10. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. J. Manag. Inf. Syst. **24**(3), 45–77 (2007)
11. Huangfu, Q., Hall, J.A.J.: Parallelizing the dual revised simplex method. Math. Program. Comput. **10**(1), 119–142 (2018). https://doi.org/10.1007/s12532-017-0130-5
12. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of ICNN'95 - International Conference on Neural Networks, vol. 4, pp. 1942–1948. (1995). https://doi.org/10.1109/ICNN.1995.488968