The Nepali Mathematical Sciences Report
Regd No.: 53/033-034, ISSN: 2392–411X
http://www.journal.cdmathtu.edu.np
Year: 2025, Volume: 42, No. 2
Pages: 149–161
https://doi.org/10.3126/nmsr.v42i2.88557
Research Article

Automatic Operating Room Scheduling to Reduce Surgical Delays

Ramtapasya Kumar Sah[1] and Tanka Nath Dhamala[*2]

[1]Central Department of Mathematics, Tribhuvan University, Kathmandu, Nepal
me.sahram@gmail.com
[*2]Central Department of Mathematics, Tribhuvan University, Kathmandu, Nepal
tanka.nath.dhamala@gmail.com

Received: 18 July, 2025 Accepted: 20 August, 2025 Published Online: 30 December, 2025

Abstract

Efficient operating room scheduling (ORS) is a cornerstone of modern healthcare delivery, directly impacting patient outcomes, resource utilization and hospital finances. In Kathmandu's government hospitals, patients face surgical delays ranging from six months to two years due to manual scheduling processes that are time-consuming and inflexible. Drawing inspiration from a case study at Aichi Medical University in Japan, this work reviews their ORS model which uses regression analysis and mixed-integer programming (MIP). Preliminary simulations demonstrate the system's ability to schedule 5 operations across 3 rooms with 5 surgeons within a 180-minute window, achieving significant efficiency gains in just one second. This review explores the methodology, results and transformative potential of this approach, with implications for global healthcare improvement.

Keywords: Operating room scheduling, Surgical delays, Mixed-integer programming, Regression analysis, Health care optimization, Automatic scheduling.
AMS(MOS) Subject Classification: Subject classification here.

1 Introduction

Large-scale government hospitals such as Teaching Hospital, Kanti Bal Hospital, Bir Hospital, Patan Hospital, Nepal Police Hospital, Civil Hospital are facing hardness to provide sufficiently satisfactory service to patients visiting them. People often wait between six

*Corresponding author

months to two years for a surgical appointment. Manual scheduling of surgeons' time which takes about two to three hours, is inefficient and sudden schedule changes pose further problems for both patients and hospital management[12].

Inspired by a case study at Aichi Medical University Hospital in Japan by Ito et al. [12], this research overviews their scheduling system over small sample size and generates efficient schedules in seconds for smaller size with optimal solution using multiple regression analysis and MIP. The primary aim of this study is to understand the modernization of healthcare infrastructure and systems.

2 Literature Review

Efficient operating room scheduling is critical, as operating rooms account for significant hospital revenue and expenditure [13, 17]. Cardoen et al. [4] emphasize that aging populations worsen waiting list issues, necessitating optimized scheduling. Dexter et al. [7] note that delayed start times increase overtime costs. Denton et al. [6] proposed a two-stage linear programming model to improve scheduling under uncertainty, while Blake et al. [3] used integer programming at Mount Sinai Hospital, saving approximately $20,000 annually.

Stochastic optimization has been effective in operating room scheduling. Lamiri et al. [15] combined Monte Carlo and MIP for cost-effective planning and Denton et al. [5] improved surgery sequencing. Metaheuristic approaches, such as simulated annealing [9] and Ant Colony Optimization [24], reduced wait times and costs. Baesler et al. [2] demonstrated simulated annealing's effectiveness in Chile while Razmi et al. [20] reduced OR costs in Iran using a stochastic model.

Recent studies incorporate advanced techniques. Soh et al. [23] introduced a scheduling metric improving room utilization. Kroer et al. [14] reduced overtime in Denmark using MIP and Monte Carlo. Ibrahim et al. [11] improved prediction accuracy by 17% by combining physician input with statistical models. Zhang et al. [25] used a two-level optimization model for cost reduction and Hamid et al. [10] minimized cancellations using MIP and simulation. Oliveira et al. [19] balanced urgency and fairness while Sauré et al. [21] applied Markov decision processes for robust scheduling. Ahmed et al. [1] integrated patient preferences and Maleki et al. [18] improved scheduling in Tehran using robust optimization. Gaon et al. [8] emphasized collaborative scheduling and Lopes et al. [16] used AI-based heuristics in Portugal to minimize costs and wait times.

Nepal-specific research is limited. Shrestha [22] documented long surgical delays in Kathmandu but offered no technological solutions. This study adapts global best practices to Nepal's context as well as others using regression analysis and MIP.

3 Methodology

Mixed Integer Linear Programming (MILP) is a powerful mathematical optimization framework where the objective function and constraints are linear and decision variables are a mix of integer (discrete) and continuous types. Integer variables are restricted to whole numbers or binary values (0 or 1) while continuous variables can take any real number within specified bounds. This duality enables MILP to address problems involving discrete decisions such as selecting facilities or scheduling tasks along with continuous decisions such as determining resource quantities. MILP problems involve:

- **Continuous Variables**: Represent measurable quantities such as production levels or flow rates.
- **Integer Variables**: Represent discrete choices such as the number of machines or facilities.
- **Binary Variables**: Represent logical decisions such as whether to undertake a project.

The objective typically focuses on minimizing costs, maximizing profits, or optimizing resource allocation. Constraints enforce operational, logical or resource-based limits expressed as linear equations or inequalities. MILP problems are computationally challenging (NP-hard) due to the discrete nature of integer variables. Key solution techniques include:

- **Branch-and-Bound**: Solves the linear programming (LP) relaxation (ignoring integer constraints) then branches on fractional variables to explore integer solutions iteratively refining bounds to find the optimal solution.
- **Cutting Planes**: Introduces additional constraints to eliminate fractional solutions while preserving valid integer solutions.
- **Branch-and-Cut**: Combines branch-and-bound with cutting planes for enhanced efficiency.
- **Heuristics**: Employs methods like diving or metaheuristics to quickly identify good though not always optimal solutions.

Commercial optimization solvers such as Gurobi, CPLEX, Xpress and open-source tools like CBC, SCIP, GLPK and HiGHS implement these techniques with advanced methods. These solvers efficiently handle problems with thousands of variables and constraints, enabling practical applications in large-scale scenarios. MILP is a cornerstone of optimization due to its ability to model and solve complex real-world problems. Its key advantages include:

- **Versatility**: Applicable across domains like logistics, energy, manufacturing, scheduling and finance.
- **Exact Solutions**: Guarantees globally optimal solutions, critical for high-stakes decision-making.
- **Discrete Decision Modeling**: Integer and binary variables allow precise modeling of logical conditions (e.g., fixed costs, mutual exclusivity) and discrete choices (e.g., facility location).
- **Scalability**: Advances in algorithms (e.g., branch-and-cut) and parallel computing enable efficient handling of large-scale problems.
- **Integration**: Seamlessly interfaces with simulation tools (e.g., Aspen Plus) and data-driven models broadening its applicability.

MILP faces the following challenges:

- **Computational Complexity**: Solution times can increase exponentially with problem size due to the combinatorial nature of integer variables.
- **Linearity Requirement**: Nonlinear relationships must be linearized potentially increasing model complexity and computational effort.

The proposed system by Ito et al. [12], employs a mathematical model to optimize operating room scheduling by minimizing overtime and ensuring efficient resource utilization.

Index Set

- O: Operations
- R: Operating rooms
- T: Time periods
- D: Departments
- S: Surgeons
- $G, A \subset O$: Operations requiring general/local anesthesia
- $H, C \subset R$: operating rooms for general/local anesthesia ($H \cap C = \emptyset$)
- $B \subset O$: Operations with fixed start times

Parameters

- M: Time to prepare operating room instruments
- L: operating room closing time
- F_{dr}: Time department d can use operating room r
- d_o, σ_o: Expected duration and standard deviation of operation o
- a: Penalty limit for unsuitable operating room assignments
- δ: Overestimation factor for operation durations
- b_o: Fixed start time for operation $o \in B$
- P_{dr}: Penalty for department d using operating room r
- S_d: Earliest start time for department d
- W_{od}: 1 if operation o belongs to department d, else 0
- V_{os}: 1 if surgeon s performs operation o, else 0

Variables

- x_{ort}: 1 if operation o starts in operating room r at time t, else 0
- E_o: Overtime for operation o
- Z_{dr}: Excess time department d uses operating room r
- U_r: Variability bound for consecutive operations in operating room r

Mathematical Model

The mathematical formulation of the ORS problem proposed by [12] is

$$\text{minimize } \alpha \sum_{o \in O} E_o + \beta \sum_{r \in R} \sum_{d \in D} Z_{dr} + \gamma \sum_{r \in R} U_r \tag{3.1}$$

subjected to the constraints

$$\sum_{t \in T} (t + d_o + \delta\sigma_o) x_{ort} - L \leq E_o, o \in O, r \in R \tag{3.2}$$

$$\sum_{t \in T} \sum_{o \in O} d_o W_{od} x_{ort} \leq F_{dr} + Z_{dr}, d \in D, r \in R \tag{3.3}$$

$$\sum_{t'=t}^{t+d_o} \sigma_o x_{ort'} - \sum_{o' \in O} \sum_{t"=t+d_o}^{t+d_o+d_{o'}} \sigma_{o'} x_{o'rt"} \leq U_r, o \in O, r \in R, t \in T \tag{3.4}$$

$$\sum_{o \in O} \sum_{r \in R} \sum_{t'=t-d_o}^{t} V_{os} x_{ort'} \leq 1, t \in T, s \in S \tag{3.5}$$

$$\sum_{t \in T} \sum_{o \in O} P_{dr} W_{od} x_{ort} \leq a, d \in D, r \in R \tag{3.6}$$

$$\sum_{t \in T} \sum_{r \in H} x_{ort} = 1, o \in G \tag{3.7}$$

$$\sum_{t \in T} \sum_{r \in C} x_{ort} = 1, o \in A \tag{3.8}$$

$$b_o = \sum_{t \in T} \sum_{r \in R} t x_{ort}, o \in B \tag{3.9}$$

$$S_d \leq \sum_{r \in R} \sum_{t \in T} W_{od} t x_{ort}, o \in O, d \in D \tag{3.10}$$

$$\sum_{o \in O} \sum_{t'=t-(M+d_o+\delta\sigma_o)} x_{ort'} \leq 1, r \in R, t \in T \tag{3.11}$$

$$E_o \geq 0, o \in O \tag{3.12}$$

$$U_r \geq 0, r \in R \tag{3.13}$$

$$Z_{dr} \geq 0, d \in D, r \in R \tag{3.14}$$

$$x_{ort} = 0, 1, o \in O, r \in R, t \in T \tag{3.15}$$

The objective function 3.1 is formulated as a weighted sum of three terms, each addressing a distinct aspect of scheduling efficiency in a hospital's operating room management.The first term $\alpha \sum_{o \in O} E_o$ represents overtime for operation o, defined as the amount of time the operation's completion exceeds the operating room closing time L. It is calculated in Constraint 3.2 where $t + d_o + \delta\sigma_o$ is the completion time of operation o if it starts at time t in operating room r, with d_o as the expected duration and $\delta\sigma_o$ as an overestimation for uncertainty. E_o captures any excess beyond L, and $E_o \geq 0$. The term $\sum_{o \in O} E_o$ sums the overtime across all operations, weighted by α.It is important for resource utilization, cost control and patient-staff well-being.

The second term $\beta \sum_{r\in R} \sum_{d\in D} Z_{dr}$ is responsible for department operating room overuse with variable Z_{dr} that represents the extra time department d uses operating room r beyond the allocated time F_{dr}. It is defined in Constraint 3.3 in which left-hand side calculates the total time department d uses operating room r. If this exceeds F_{dr}, Z_{dr} captures the excess, and $Z_{dr} \geq 0$. The term $\sum_{r\in R} \sum_{d\in D} Z_{dr}$ sums the overuse, weighted by β.It is important for fair resource allocation, operational efficiency , inter-departmental coordination and penalty avoidance.

The third term represents the variability in consecutive operations with variable U_r is the upper bound on the difference in standard deviations of durations between consecutive operations in operating room r. It is defined in Constraint 3.4 where the first term represents the standard deviation σ_o of operation o's duration, and the second term represents $\sigma_{o'}$ of the next operation. U_r bounds the difference $|\sigma_o - \sigma_{o'}|$, and $U_r \geq 0$. The term $\sum_{r\in R} U_r$ sums the variability bounds, weighted by γ. It is important for schedule robustness, risk management and staff-resource planning.

Constraint 3.2 defines overtime E_o as the amount by which operation o exceeds the operating room closing time L and ensures that late-running surgeries are penalized in the objective function and encourages scheduling surgeries earlier in the day to avoid overtime.

Constraint 3.3 limits the total time department d uses operating room r to $F_{dr} + Z_{dr}$ and Z_{dr} captures **excess time** beyond allocated F_{dr}. It prevents one department from **monopolizing operating room resources** at the expense of others and encourages departments to **optimize their surgical schedules**.

Constraint 3.4 controls the **variability** between consecutive surgeries in the same operating room and U_r ensures that two back-to-back surgeries do not have **wildly different durations**. It maintain smoothness of operations in operating rooms and reduces **idle time** between surgeries and makes schedules **more predictable** for staff.

Constraint 3.5 ensures a surgeon is **not assigned to two overlapping surgeries** and prevents **double-booking surgeons**, which would be physically impossible.

Constraint 3.6 limits the total penalty for assigning department d's surgeries to non-preferred operating rooms and encourages **preferred operating room assignments** (e.g., cardiac surgeries in operating rooms with heart-lung machines).

Constraint 3.7 and 3.8 ensure surgeries requiring **general anesthesia** ($o \in G$) go to operating rooms equipped for it ($r \in H$) and **local anesthesia** surgeries ($o \in A$) go to operating rooms in C.

Constraint 3.9 forces operations $o \in B$ to start at a **fixed time** b_o (e.g., scheduled by the surgeon) and respects **surgeon preferences** and **emergency cases** needing priority.

Constraint 3.10 ensures department d's surgeries do not start before S_d and allows **department-specific scheduling policies** (e.g., pediatric surgeries in the morning).

constraint 3.11 ensures **each surgery is assigned to at most one operating room at one time** and constraints 3.12, 3.13 and 3.14 shows non-negativity of the terms. Also,constraint 3.15 shows that it is binary variable.

Implementation

Using Python a random dataset of 500 samples was generated with normally distributed variables including predictors age, sex, surgeon experience, department, complexity and body mass index (BMI) where the dependent variable, operation duration is derived from a linear combination of these factors. Categorical variables like sex, department, and complexity are discretized and encoded using one-hot encoding while binary variables like sex are thresholded into 0s and 1s. The data is then cleaned and converted into numeric format where necessary for regression compatibility. The data is then processed to handle missing values and ensure all features are numeric. Using statsmodels' Ordinary Least Squares (OLS) regression, the model is trained to understand how each variable affects operation duration and is fitted to predict operation duration based on the specified predictors as in table 1. The model's statistical summary provides with R-squared > 0.5 insights into feature significance and overall model fit. Then it was applied to 10 new test samples, encoding categorical variables appropriately and generating predicted operation durations.

The model was implemented using Python's PuLP library. The algorithm processes input data (e.g., operation durations, surgeon availability) and generates an optimized schedule visualized as a Gantt chart. The system handles dynamic changes such as emergency cases, by reserving ORs and adjusting schedules in real time.

The model was tested using a simulated dataset of 5 operations across 3 operating rooms within a 180 minute window. Regression analysis predicted operation durations, while MIP optimized operating room allocation. The system's scalability was demonstrated by its ability to handle multiple constraints, such as anesthesia types and fixed start times. The standard deviation column in Table 1 represents the variability in predicted operation durations. Incorporating variability ensures the model accounts for uncertainties in real-world scenarios, improving the robustness of the scheduling system.

This system serves as a model for other low-resource healthcare systems globally. Future validation using real-world data will further strengthen its credibility. The regression model used is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \epsilon$$

Where:

- Y: Dependent variable (predicted operation duration, d_o)

- $X_1, X_2, X_3, X_4, X_5, X_6$: Independent variables (predictors: Age, Sex, Surgeon Experience, Department, Complexity, Body Mass Index respectively)
- β_0: Intercept
- $\beta_1, \beta_2, \ldots, \beta_k$: Regression coefficients
- ϵ: Error term

Table 1: Prediction of Operation Duration

SN	Age	Sex	Surgeon Experience	Cardiology	Neurology	Complexity	BMI	Predicted Duration	Std. Dev.
1	45	0	8	1	0	2	22	48.08	19.76
2	60	1	12	0	1	3	30	90.91	32.72
3	55	0	10	0	0	1	26	31.55	19.68
4	70	1	15	1	0	2	24	57.32	26.85
5	40	0	5	0	1	3	28	84.32	28.80
6	50	1	9	0	0	1	25	40.63	19.61
7	65	0	11	1	0	2	27	55.34	25.61
8	35	1	7	0	1	3	23	84.11	27.95
9	58	0	13	0	0	1	29	28.47	20.38
10	48	1	8	1	0	2	26	62.51	23.73

4 Results and Discussion

First, operation durations (40–90 minutes) were predicted using regression analysis. Then, in a simulation involving 5 operations, 3 operating rooms, and 5 surgeons, the above mathematical model scheduled surgeries within a 180-minute window, ensuring no overlaps and reserving rooms for emergencies in less than one second.

Gantt Chart

Figure 1 presents the Gantt chart output, illustrating the optimized schedule. Each bar represents an operation, with colors distinguishing different surgeons and departments. The chart enhances coordination among hospital staff and provides a clear visualization of operating room utilization. The role of standard deviation in the regression model ensures variability is accounted for, enhancing the robustness of predictions.

The operational efficiency of the algorithm using CBC MILP Solver ,Version: 2.10.3 for single department is shown in table 2.

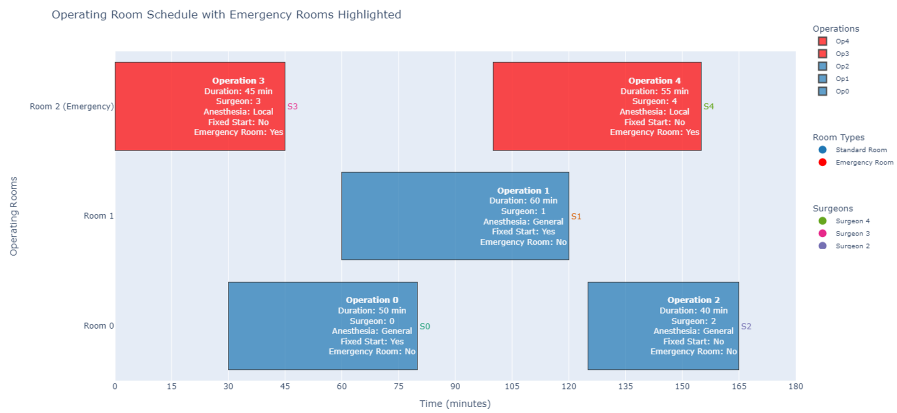

Figure 1: Automated OR Schedule

Table 2 Operational Efficiency Data

Number of Operations	Number of Rooms	Number of Surgeons	Average Time (Seconds)	Result
2	2	2	0.23	Optimal
3	3	3	2.13	Optimal
4	4	4	5.77	Optimal
5	5	5	4.52	Optimal
6	6	6	146.17	Optimal
7	7	7	600	No Solution

Challenges and Considerations

Implementation requires investment in technology and training. Reliable data on surgery durations and resources is essential, requiring robust data systems. Future enhancements could integrate machine learning for improved predictions.

5 Conclusion

The system delivers efficient, adaptable schedules in less than 3 minutes for number less or equal to 6 of operations, surgeons and operating rooms with single department.

Future work should focus on scaling the model, integrating real-time data, and evaluating long-term impacts on patient outcomes and hospital finances. By serving as a blueprint for low-resource healthcare systems, this study contributes to global healthcare improvement.

References

[1] Ahmed, A., Ali, H., Modeling patient preference in an operating room scheduling problem, *Operations Research for Health Care*, Vol. 25, 100257, 2020. https://doi.org/10.1016/j.orhc.2020.100257

[2] Baesler, F., Gatica, J., Correa, R., Simulation optimization for operating room scheduling, *International Journal of Simulation Modeling*, Vol. 14(2), pp. 215–226, 2015. https://doi.org/10.2507/IJSIMM14(2)3.287

[3] Blake, J. T., Donald, J., Mount Sinai Hospital uses integer programming to allocate operating room time, *Interfaces*, Vol. 32(2), pp. 63–73, 2002.

[4] Cardoen, B. E., Demeulemeester, B. J., Operating room planning and scheduling: A literature review, *European Journal of Operational Research*, Vol. 201(3), pp. 921–932, 2010.

[5] Denton, B., Viapiano, J., Vogl, A., Optimization of surgery sequencing and scheduling decisions under uncertainty, *Health Care Management Science*, Vol. 10(4), pp. 399–416, 2006. https://doi.org/10.1007/s10729-006-9005-4

[6] Denton, B. J., Viapiano, A. V., Optimization of surgery sequencing and scheduling decisions under uncertainty, *Health Care Management Science*, Vol. 10(1), pp. 13–24, 2007. https://doi.org/10.1007/s10729-006-9005-4

[7] Dexter, F., Macario, A., Applications of information systems to operating room scheduling, *Anesthesiology*, Vol. 85, pp. 1232–1234, 1996.

[8] Gaon, N., Gabai Schlosberg, Y., Zivan, R., Scheduling operations in a large hospital by multiple agents, *Engineering Applications of Artificial Intelligence*, Vol. 126, 107074, 2023. https://doi.org/10.1016/j.engappai.2023.107074

[9] Granja, C., Almada-Lobo, B., Janela, F., Seabra, J., Mendes, A., An optimization based on simulation approach to the patient admission scheduling problem using a linear programming algorithm, *Journal of Biomedical Informatics*, Vol. 52, pp. 427–437, 2014. https://doi.org/10.1016/j.jbi.2014.08.007

[10] Hamid, M., Hamid, M., Musavi, M., Azadeh, A., Scheduling elective patients based on sequence-dependent setup times in an open-heart surgical department using an optimization and simulation approach, *Simulation: Transactions of the Society for Modeling and Simulation International*, Vol. 95(12), pp. 1141–1164, 2019. https://doi.org/10.1177/0037549718811591

[11] Ibrahim, R., Kim, S.-H., Predicting surgery duration: Physician input, statistical models, and combined models, *University College London Marshall School of Business, University of Southern California*, 2018.

[12] Ito, M., Suzuki, A., Fujiwara, Y., Operating rooms scheduling system—A case study in Aichi Medical University Hospital, *International Symposium on Scheduling*, pp. 114–120, 2015.

[13] Jackson, R., The business of surgery, *Health Management Technology*, Vol. 23(7), pp. 20–22, 2002.

[14] Kroer, L. R., Foverskov, K., Vilhelmsen, C., Hansen, A. S., Larsen, J., Planning and scheduling operating rooms for elective and emergency surgeries with uncertain duration, *Operations Research for Health Care*, Vol. 19, pp. 107–119, 2018. https://doi.org/10.1016/j.orhc.2018.03.006

[15] Lamiri, M., Xie, X., Dolgui, A., Grimaud, F., A stochastic model for operating room planning with elective and emergency demand for surgery, *European Journal of Operational Research*, Vol. 185(3), pp. 1026–1037, 2008. https://doi.org/10.1016/j.ejor.2006.02.057

[16] Lopes, J., Guimarães, T., Duarte, J., Santos, M., Enhancing surgery scheduling in health care settings with metaheuristic optimization models: Algorithm validation study, *JMIR Medical Informatics*, Vol. 13, e57231, 2025. https://doi.org/10.2196/57231

[17] Macario, A., Vitez, T. S., Dunn, B., McDonald, T., Where are the costs in perioperative care?: Analysis of hospital costs and charges for inpatient surgical care, *Anesthesiology*, Vol. 83(6), pp. 1138–1144, 1995.

[18] Maleki, A., Hosseininesaz, H., Jasemi, M., A comparative analysis of the efficient operating room scheduling models using robust optimization and upper partial moment, *Healthcare Analytics*, Vol. 3, 100144, 2023. https://doi.org/10.1016/j.health.2023.100144

[19] Oliveira, M., Bélanger, V., Marques, I., Ruiz, A., Assessing the impact of patient prioritization on operating room schedules, *Operations Research for Health Care*, Vol. 24, 100232, 2020. https://doi.org/10.1016/j.orhc.2019.100232

[20] Razmi, J., Yousefi, M. S., Barati, M., A stochastic model for operating room unique equipment planning under uncertainty, *IFAC-PapersOnLine*, Vol. 48(3), pp. 1796–1801, 2015. https://doi.org/10.1016/j.ifacol.2015.06.347

[21] Sauré, A., Begen, M. A., Patrick, J., Dynamic multi-priority, multi-class patient scheduling with stochastic service times, *European Journal of Operational Research*, Vol. 280(1), pp. 254–265, 2020. https://doi.org/10.1016/j.ejor.2019.06.040

[22] Shrestha, R., Healthcare challenges in Nepal: A case study of Kathmandu's public hospitals, *Nepal Medical Journal*, Vol. 12(2), pp. 45–52, 2023.

[23] Soh, K. W., Walker, C., O'Sullivan, M., Wallace, J., Innovative operating room scheduling metric for creating surgical lists with desirable room utilisation rates, *Operations Management Research*, Vol. 17, pp. 544–567, 2024.

[24] Xiang, W., Yin, J., Lim, G., An ant colony optimization approach for solving an operating room surgery scheduling problem, *Computers and Industrial Engineering*, Vol. 85, pp. 335–345, 2015. https://doi.org/10.1016/j.cie.2015.04.010

[25] Zhang, J., Dridi, M., El Moudni, A., A two-level optimization model for elective surgery scheduling with downstream capacity constraints, *European Journal of Operational Research*, Vol. 276(2), pp. 602–613, 2019. https://doi.org/10.1016/j.ejor.2019.01.036