Evaluating Predict-then-Optimize methods for surgical case planning

Pieter Smet[1] and David Ajit Kirpekar-Sauer[2]

[1] KU Leuven, Department of Computer Science, Gebroeders de Smetstraat 1, 9000 Gent, Belgium

[2] Technical University of Denmark, Anker Engelundsvej, 2800 Kongens Lyngby, Denmark

1 Introduction

We consider a Surgical Case Planning (SCP) problem in which patients are selected from a waiting list for elective surgery and assigned surgery dates. We account for the limited capacity of two key resources: operating rooms (ORs) and anesthesiologists who must attend each procedure [1]. This problem is solved after both the master surgery schedule (MSS) of the ORs and the rosters for the anesthesiologist have been determined, but before the patients' actual clinical conditions and potential staff absences are known. As a result, the SCP decisions are subject to significant uncertainty [7]. We model this uncertainty through two problem parameters that are unknown at the time of solving: (i) surgery durations and (ii) the absence or presence of scheduled anesthesiologists.

To address this uncertainty, we adopt a Predict-then-Optimize (PTO) approach. In the first stage, two prediction models are used to predict surgery durations and anesthesiologist absences. These predictions are then used in the second stage, where a deterministic optimization model is solved to generate a solution for the SCP problem. It is well known that for some problems, improving prediction accuracy does not necessarily lead to better solutions of the optimization model [4]. The relationship between prediction accuracy and solution quality becomes even more complicated when two prediction models are involved, as is the case in the SCP problem. In this work, we empirically evaluate this relationship, thereby enabling an analysis of how solution quality changes as a function of prediction accuracy.

2 Problem description

For a given planning period, the considered SCP problem involves selecting a subset of patients from a waiting list and assigning each of them a feasible surgery date, OR and anesthesiologist, while respecting the capacity constraints of ORs and anesthesiologists. The anesthesiologist rosters and MSS specify the available hours for both of these resources on each day of the planning period. All patients on the waiting list are eligible for surgery in the planning period. Each patient has a due date before which they must be operated on, as well as an estimated surgery duration. Additionally, clinical requirements may restrict compatibility between patients and certain anesthesiologists or ORs. For example, certain patients require anesthesiologists with specific skills, while certain procedures depend on equipment that is not available in all ORs.

Two problem parameters are unknown at the time of solving:

1. The surgery duration t_p of patient p. Although historical averages are often used in practice, these may be inaccurate. We assume instead that this duration can be predicted using patient-specific features $\mathbf{z}_p$, such as their age, clinical condition and medical history [5].
2. The presence of anesthesiologist a scheduled to work on a day d, represented as a binary parameter b_{ad}. Unexpected absenteeism may occur due to illness or external factors. We assume that absences can be predicted from anesthesiologist-specific features $\mathbf{z}_a$, such as workload intensity, recent absenteeism frequency and seasonal illness trends [3].

The objective is to minimize a weighted sum of two objectives: (i) the number of external workers required to cover for absent anesthesiologists and (ii) the total waiting time of all patients. Unscheduled patients are carried over to the next planning period, incurring additional waiting time.

3 Methodology

To solve the SCP problem, we apply the PTO approach illustrated in Figure 1. Two prediction tasks are carried out: a regression task to predict surgery duration $\hat{t}_p$, and a binary classification task to predict anesthesiologist absences $\hat{b}_{ad}$. The optimization model is then solved using integer programming to obtain an optimal solution $x^*(\hat{t}_p, \hat{b}_{ad})$ from the set of feasible solutions $\mathcal{X}$.

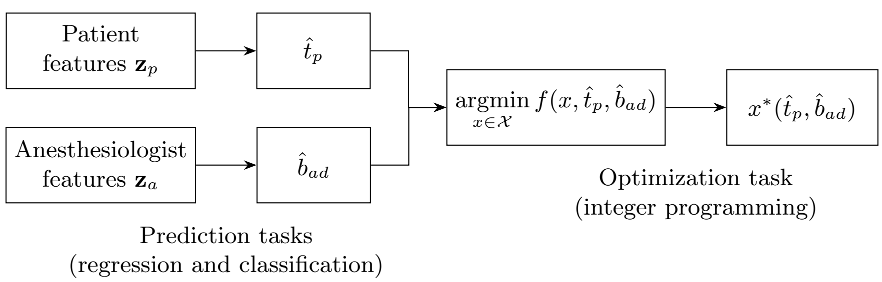

Fig. 1. Schematic overview of the PTO approach.

We conduct a series of experiments to analyze how the accuracy of the regression and classification models impacts solution cost. To this end, we use the algorithms introduced in [2,6] to simulate the predictions that models would produce at a specified level of accuracy. By systematically adjusting this accuracy, we obtain an empirical mapping between prediction performance metrics and solution cost. We vary the prediction accuracies of both models simultaneously to assess interaction effects between the two sources of uncertainty.

Solution cost is computed by simulating the realization of the actual surgery durations and anesthesiologist absences given a solution $x^*(\hat{t}_p, \hat{b}_{ad})$ obtained with predictions. Once the uncertain parameters are revealed on each day, we apply one of several rescheduling policies to restore feasibility, if necessary. These policies may involve postponing surgeries, reassigning them to a different OR or anesthesiologist, or a combination of these actions. Each change incurs a cost, which is minimized during the rescheduling process.

Preliminary computational results based on data from a Danish hospital indicate that KPIs such as the number of postponed, canceled or resource-changed patients, are affected in a similar way by the accuracy of surgery duration predictions. Notably, the results demonstrate that improvements in prediction accuracy do not necessarily lead to better KPI values. Furthermore, the accuracy of anesthesiologist absence predictions appears to have only a limited influence on these results. Overall, these findings suggest that the relationship between prediction accuracy and operational performance is not straightforward, highlighting the need for further investigation into the interaction between prediction models and scheduling decisions.

Acknowledgments This research was partially supported by KU Leuven C24E/23/012 'Human-centred decision support based on new theory for personnel rostering'.

References

1. Cardoen, B., Demeulemeester, E., Beliën, J.: Operating room planning and scheduling: A literature review. European journal of operational research **201**(3), 921–932 (2010)
2. Doneda, M., Smet, P., Carello, G., Lanzarone, E., Vanden Berghe, G.: Robust personnel rostering: how accurate should absenteeism predictions be? Journal of Scheduling pp. 1–16 (2025)
3. Lawrance, N., Petrides, G., Guerry, M.A.: Predicting employee absenteeism for cost effective interventions. Decision Support Systems **147**, 113539 (2021)
4. Mandi, J., Kotary, J., Berden, S., Mulamba, M., Bucarey, V., Guns, T., Fioretto, F.: Decision-focused learning: Foundations, state of the art, benchmark and future opportunities. Journal of Artificial Intelligence Research **80**, 1623–1701 (2024)
5. Park, R.T., Stucky, C.H., Moser, C.H.: Machine learning surgery duration predictions compared to traditional methods: A systematic review. Perioperative Care and Operating Room Management **41**, 100581 (2025)
6. Smet, P., Doneda, M., Lanzarone, E., Carello, G.: Prediction accuracy versus rescheduling flexibility in elective surgery management. arXiv preprint arXiv:2507.15566 (2025)
7. Wang, L., Demeulemeester, E., Vansteenkiste, N., Rademakers, F.E.: Operating room planning and scheduling for outpatients and inpatients: A review and future research. Operations Research for Health Care **31**, 100323 (2021)