

Can machine learning optimize the efficiency of the operating room in the era of COVID-19?

Natasha Rozario
Duncan Rozario, MD

Accepted Oct. 14, 2020

Correspondence to:

N. Rozario
Faculty of Mathematics
University of Waterloo
200 University Ave W
Waterloo ON N2L 3G1
mail@natasharozario.com

DOI: 10.1503/cjs.016520

SUMMARY

The cancellation of large numbers of surgical procedures because of the coronavirus disease 2019 (COVID-19) pandemic has drastically extended wait lists and negatively affected patient care and experience. As many facilities resume clinical work owing to the currently low burden of disease in our community, we are faced with operative booking protocols and procedures that are not mathematically designed to optimize efficiency. Using a subset of artificial intelligence called “machine learning,” we have shown how the use of operating time can be optimized with a custom Python (a high-level programming language) script and an open source machine-learning algorithm, the OR-Tools software suite from the Google AI division of Alphabet Inc. This allowed the creation of customized models to optimize the efficiency of operating room booking times, which resulted in a reduction in nursing overtime of 21% — a theoretical cost savings of \$469 000 over 3 years.

The coronavirus disease 2019 (COVID-19) pandemic resulted in the declaration of an emergency on Mar. 17, 2020, under the *Emergency Management and Civil Protection Act*, leading to the cessation of elective surgery in Ontario. It was estimated that between March 15 and June 13 there was a provincial backlog of 148 364 surgeries.¹ A gradually staged resumption of services started on May 16, 2020, exacerbating existing wait list issues and resulting in significant morbidity and mortality for patients. At our institution, we use a surgeon’s average procedure time for their last 10 cases as the booking time for future procedures. Mathematically, does that optimize the utilization of operative time? Cases that run late incur staff overtime costs, affect physician scheduling, decrease morale and delay the ability to complete urgent or emergent cases. Conversely, operating rooms (ORs) that finish early do not maximize the use of available time and are a waste of costly resources. In some centres, cases that will finish late are cancelled, resulting in significant patient dissatisfaction.

Oakville Trafalgar Memorial Hospital is a 469-bed facility in Oakville, Ontario, that performed 13 717 surgical procedures in 2017–18. In December 2015, we moved to a new 1.5-million-square-foot facility; it is 3 times larger than the previous hospital and has 10 ORs.

OPTIMIZING BOOKING TIMES USING A MACHINE LEARNING MODEL

The current common method of booking operative procedures using the average of the last 10 case times results in about 50% of cases running overtime, as case times follow a Gaussian distribution. Procedures with highly variable case times can cause a cascade of delays that result in more ORs running overtime.²

An optimization problem is defined using a decision variable for every type of procedure. Each variable is bound by the range of actual case times in the data set. An OR is considered to be running overtime if the sum of actual

procedure times plus changeovers exceeds the scheduled case time for a given day. At our institution, an OR is considered to be running undertime if the actual finish time is more than 15 minutes earlier than the scheduled finish time for a given day.

The objective of the optimization is to minimize both overtime and undertime cases in an OR. The relative cost of running undertime and the cost of running overtime at a local institution can be entered into the Python script (Appendix 1, available at canjsurg.ca/lookup/doi/10.1503/cjs.016520/tab-related-content). The Python script reads from an Excel spreadsheet containing a set of booking data from our PICIS OR booking system (<https://www.harriscomputer.com/en/>), formatted as shown in the example surgeon case audit report (Table 1). Our goal was for 80% of ORs to finish on time and to minimize the number of ORs that finish early. As such, the overtime cost is defined as $|x - 0.2n|$, where x is the number of rooms that run overtime in the model, and n is the number of days in the data set. The undertime cost is defined as (y) where y is the number of rooms that run undertime in the model. By multiplying these costs by different weights appropriate for a specific institution, the priority put on overtime/undertime rooms can be changed to yield different scheduling times.

Principles such as conflict-driven clause learning and intelligent backtracking are used to efficiently create a model that satisfies all the constraints and meets the objective. The model assumes that if an OR is available early, both the surgeon and the patient will also be available early.

Using 36 months of anonymized historical OR booking data from 2017 to 2019, comprising 10 553 cases (Appendix 1), a custom-created Python script and the OR-Tools optimization suite, we entered our desired optimization parameters to obtain ideal procedure times, and then assessed retroactively how our new model could have changed outcomes. We used OR patient-in-room and patient-out-of-room times for 15 surgeons from multiple divisions.

MEAN VERSUS MACHINE LEARNING MODEL RESULTS

The full set of anonymized data are shown in Table 2. Using the standard mean case time method, ORs were overtime 48% of the time and undertime 37% of the time. ORs finished within 15 minutes of the scheduled finish time 15% of the time.

If the scheduling times calculated by the machine learning model had been used, those same ORs would have been overtime 27% of the time and undertime 18% of the time. This would result in the completion of 97% of the previous volume of cases in the standard time, with the same number of OR minutes used. In turn, this would result in an overtime cost savings of \$469 000 over 3 years (approximate rate of 2.5 nursing staff per room at \$75 per hour, minimum 15-minute block). With the scheduling times suggested by the model, ORs would finish within 15 minutes of the scheduled time 55% of the time, yielding a much more consistent finish time.

We have provided in Appendix 1 all of the anonymized, raw data which show the scheduling time calculated by both methods for the 10 most common procedures of 15 surgeons from multiple divisions.

DISCUSSION

The authors of the pivotal book *Prediction machines*³ state that, “prediction is at the heart of making decisions under uncertainty,” and that the drop in the cost of prediction is the key way that will democratize the power of artificial intelligence. Historical OR booking data are a well annotated, big data set that leads well to machine learning optimization to predict future outcomes and can easily be used in an iterative fashion to continuously improve.

Cost-effective predictions of operative booking times are crucial to optimizing the efficiency of the OR. In future models this could be done in real time to predict case completion times as circumstances (e.g., intraoperative complications) change during a case.⁴

The machine learning model developed at our institution is applicable to other institutions, which can use their own historical data to predict future models. Local constraints (e.g., the priority placed on undertime and overtime finishes, overtime rules) can be customized into the script to produce relevant booking times. Cases that inherently have greater variation (e.g., complex colonic resections) introduce more error into predictions.

Surgeon estimates of case times can be inaccurate, and historic patterns of using mean times do not optimize outcomes owing to case variation. Furthermore, taking an average of only 10 cases can introduce error owing to the small sample size. In addition, booking times change frequently and are particularly vulnerable to skewing by outlier cases. Conversely, the machine

Table 1. Surgeon case audit report*

Date	CR #	Procedure	Description	Schedule	In	An start	Pr start	Pr end	An end	Out	Actual duration	Booking duration
1/1/2017	111111	ABC	Example procedure description here	12:15	12:15	12:30	12:45	13:00	13:15	13:30	75	60

An = anesthesia; CR = case reference; Pr = procedure.
 *Date covers the period of Jan. 1, 2017, to Dec. 31, 2019.

Table 2. Summary 2017–2019*

Variable	Original	ML Model	Difference
Overtime frequency, %	47.54	26.55	-20.99
Undertime frequency, %	36.53	17.50	-19.03
Overtime minutes used	50639	19145	-31494
Overtime cost, \$	770437.50	301312.50	-469125.00
OR minutes used, %	72.56	72.65	0.09

OR = operating room.
 *Based on 10553 cases in 2975 surgeon OR days, with a total rate of machine learning cases achieved of 96.53%.

learning model is developed based on years of data, so outlier case times have minimal effect on the scheduling time for that procedure.

Whereas the original mean case method takes into account only an average of recent case times, the machine learning model analyses a number of variables, including average case times, variability of case times, frequency of procedure types and distribution of procedure types, to calculate an effective scheduling time that keeps the OR, not just the individual case, running on time. Because these factors are highly dependent on the surgeon, it is vital that this model is run on a surgeon-specific basis.

As more data, such as surgeon, anesthesiologist, type of surgery, previous surgery, surgeon estimate of complexity, type of anesthetic, American Society of Anesthesiologists class and other patient factors, are incorporated into the machine learning model, we can likely expect even more accurate estimates of case times. Machine learning optimization can create case times free of bias, as it facilitates objective, evidence-based decision-making. This method requires some familiarity with programming to run the script, and its adoption will depend on acceptance by surgeons and other involved parties.

CONCLUSION

The high-level Python programming language combined with the open source OR-Tools software suite from Google AI has the potential to easily and accurately predict operative booking times. We have shown a theoretical improvement of 21% in overtime rates and an estimated cost saving of \$469 000 over 3 years to illustrate how surgeon-specific data can be used to optimize bookings. Given the complexity of the underlying systems involved, it may be a challenge to get all involved stakeholders engaged with adopting this model. It will be important to emphasize that this model is designed to be predictive, not punitive, to optimize the efficiency of ORs. Given the system-level challenges we face now and in the months ahead as a result of the COVID-19 pandemic, a focus on the most efficient use of our limited and precious resources will be vital to provide the level of care our community requires.

Affiliations: From the Department of Mathematics, University of Waterloo, Waterloo, Ont. (N. Rozario); and the Oakville Trafalgar Memorial Hospital, Oakville, Ont. (D. Rozario).

Competing interests: None declared.

Contributors: Both authors contributed substantially to the conception, writing and revision of this article and approved the final version for publication.

References

1. Wang J, Vahid S, Eberg M et al. Clearing the surgical backlog caused by COVID-19 in Ontario: a time series modelling study. *CMAJ* 2020;192:E1347-E1356.
2. Zhu S, Fan W, Yang S et al. Operating room planning and surgical case scheduling: a review of literature. *Journal of Combinatorial Optimization* 2019;37:757-805.
3. Agrawal A, Gans J, Goldfarb A. *Prediction Machines — The Simple Economics of Artificial Intelligence*. Harvard Business Review Press; 2019.
4. Bartek MA, Saxena RC, Solomon S et al. Improving operating room efficiency: machine learning approach to predict case-time duration. *J Am Coll Surg* 2019;229:346-354.