BMJ Health & Care Informatics

Predicting surgical department occupancy and patient length of stay in a paediatric hospital setting using machine learning: a pilot study

Yuval Barak Corren [ID],[1] Joshua Merrill,[2] Ronald Wilkinson,[2] Courtney Cannon,[2] Jonathan Bickel,[2] Ben Y Reis[2,3]

Check for updates

[1]Computational Health Informatics Program, Boston Children's Hospital, Boston, Massachusetts, USA
[2]Enterprise Analytics, Boston Children's Hospital, Boston, Massachusetts, USA
[3]The Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA

Correspondence to
Dr Yuval Barak Corren;
yuval.barakcorren@childrens.harvard.edu

ABSTRACT

Objective Early and accurate prediction of hospital surgical-unit occupancy is critical for improving scheduling, staffing and resource planning. Previous studies on occupancy prediction have focused primarily on adult healthcare settings, we sought to develop occupancy prediction models specifically tailored to the needs and characteristics of paediatric surgical settings.

Materials and methods We conducted a single-centre retrospective cohort study at a surgical unit in a tertiary-care paediatric hospital in Boston, Massachusetts, USA. We developed a hierarchical modelling framework for predicting next-day census using multiple types of data—from bottom-up patient-specific orders and procedures to top-down temporal variables and departmental admission statistics.

Results The model predicted upcoming admissions and discharges with a median error of 17%–21% (2–3 patients per day), and next-day census with a median error of 7% (n=3). The primary factors driving these predictions included day of week and scheduled surgeries, as well as procedure duration, procedure type and days since admission. We found that paediatric surgical procedure duration was highly predictive of postoperative length of stay.

Discussion Our hierarchical modelling framework provides an overview of the factors driving capacity issues in the paediatric surgical unit, highlighting the importance of both top-down temporal features (eg, day of week) as well as bottom-up electronic health records (EHR) derived features (eg, orders for patient) for predicting next-day census. In the practice, this framework can be implemented stepwise, from top to bottom, making it easier to adopt.

Conclusion Modelling frameworks combining top-down and bottom-up features can provide accurate predictions of next-day census in a paediatric surgical setting.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ A shortage of surgical beds and frequent fluctuations in surgical-unit occupancy present major challenges for hospitals, and often result in last-minute surgical cancellations. Forecasting bed occupancy can mitigate this problem and has been shown to be feasible in adult settings.

WHAT THIS STUDY ADDS

⇒ This study demonstrates that accurate occupancy prediction is also feasible in paediatric settings and highlights the main features driving these predictions. We present a hierarchical modelling framework, combining top-down departmental variables with bottom-up patient-level data.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This framework can improve resource utilisation in paediatric surgical units as well as guide the development of other practical clinical prediction models.

INTRODUCTION

Substantial fluctuations in surgical-unit occupancy present a major and ongoing challenge for healthcare providers. Many tasks that depend on future bed availability, including staffing, scheduling and patient transfer management, are directly impacted by these uncertainties. Failure to accurately predict surgical department census and the resulting lack of bed availability are among the leading causes of planned surgery cancellations.[1 2] Last-minute cancellations can lead to great frustration for the patient, the patient's family and the clinical team.[3] Lastly, delays in surgery can also result in grave medical complications and increased morbidity and mortality.[4]

In recent years, a growing body of work has focused on clinical occupancy prediction in adult healthcare settings,[5–13] yet very few studies have focused on predicting capacity in paediatric hospitals or paediatric surgical departments.[14 15] Yet while inpatient beds for adult patients are projected to increase over time, the number of paediatric inpatient beds has been dropping,[16] increasing the pressure on inpatient resources devoted to paediatric services and highlighting the importance of focusing specifically on this subpopulation.

In this study, we develop an occupancy prediction model for a surgical unit in a tertiary-care paediatric hospital. We apply an integrated hierarchical modelling framework that incorporates both top-down population-level data and bottom-up individual-level data. In addition to this main modelling framework, we also develop a separate smaller model focused on postsurgical length of stay (LOS), in order to study the key risk factors associated with long postoperative stays.

METHODS
Setting

We analysed data from a surgical unit in a tertiary-care paediatric hospital in Boston, Massachusetts, USA. The data included all clinical and administrative orders, surgical procedures and procedure durations for patients admitted to the surgical unit from September 2015 to March 2020 (4.5 years). The data also included the daily census of the top three admitting departments, comprising 90% of all admissions to the surgical unit:

emergency department (ED), postanaesthesia care unit (PACU) and surgical intensive care unit (ICU). Data from the first 3.5 years of the study period (1 September 2015–28 February 2019) were used for model training, while data from the last year (1 March 2019–29 February 2020) were used for model validation.

Census prediction

The daily census was calculated based on the number of patients in the surgical unit each day at 8:00 hours. To predict next-day census, the following formula was used:

Predicted census = Current census + Expected admissions − Expected discharges

An overview of the modelling framework is presented in figure 1. This figure may serve as a useful guide to the description of the multipart model in the following paragraphs. We applied a hierarchical modelling approach that fused both population-level temporal trends (top-down) together with individual-patient-specific factors (bottom-up).
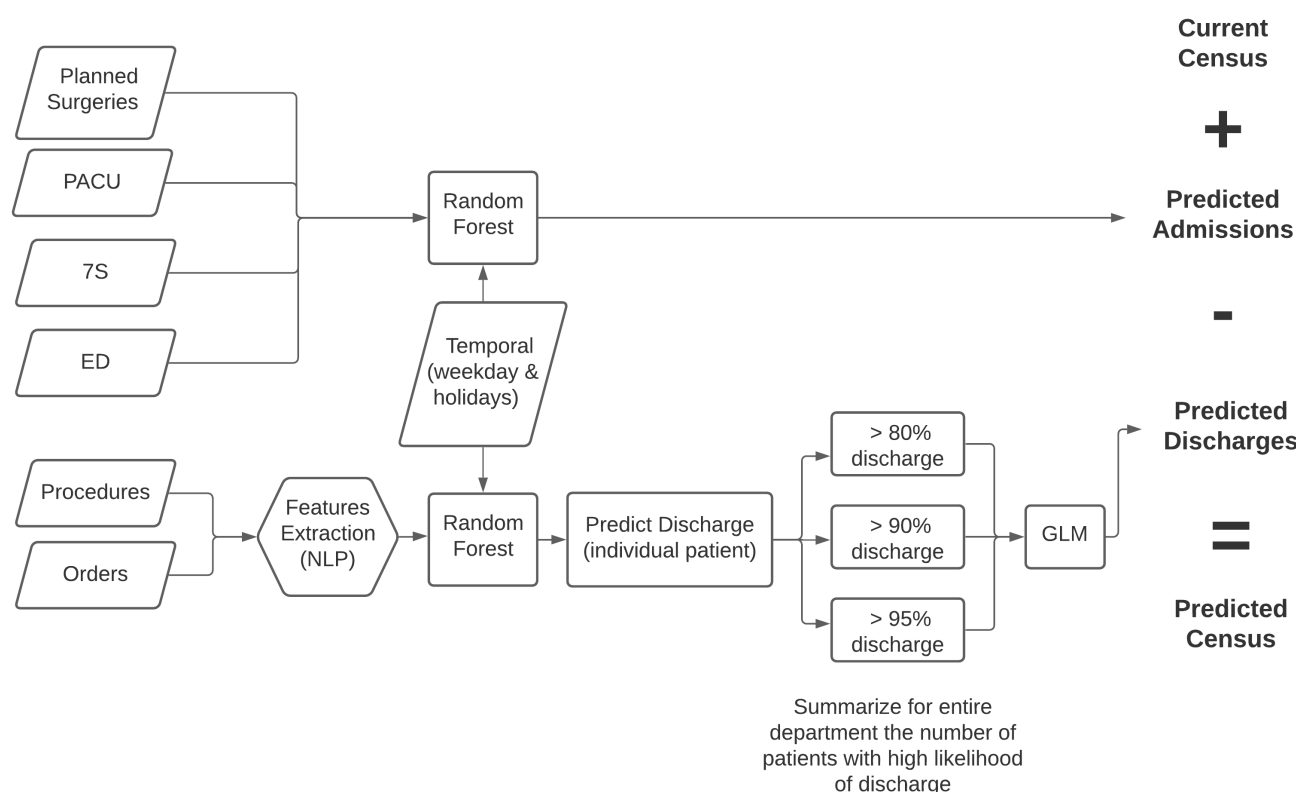

Figure 1 Overview of hierarchical modelling framework. Next-day surgical unit census is modelled as a function of the current census, the inflow of patients (ie, expected admissions) and the outflow of patients (ie, expected discharges). Patient inflow is predicted by a random forest model that uses the census data from the top three admitting departments (PACU, SICU, ED), as well as the schedule of planned surgeries, combined with temporal data on current weekday and nearby holidays. Discharge prediction for each individual patient is based on a random forest model that incorporates features extracted by natural language processing (NLP) tools from the orders and procedures datasets. These predictions also incorporate temporal effects related to the day of week and nearby holidays. All individual-patient discharge predictions are aggregated within a generalised linear model (GLM) to predict the overall number of expected discharges. ED, emergency department; PACU, post-anaesthesia care unit; SICU, surgical intensive care unit.

The top-down features included information on variations in occupancy associated with day of week, month of year and holidays, as well as the overall census in the top admitting departments. The bottom-up features included individual patient orders and procedures data.

Both prediction of expected admissions and prediction of expected discharges were based on a random forest model using the randomForest package in R.[17 18] The input features for the admission prediction model included the census in the top three admitting departments, the schedule of planned surgeries for the next day, as well as additional temporal information (month in year, day of week and nearby holidays). The input features for the discharge prediction model included the current day of admission, data on procedures and orders, and the temporal information described above. Further description of the features extracted from the procedures and orders data is provided in online supplemental appendix A.

Predictions for individual-patient discharges, based on both bottom-up and top-down features, were incorporated into a generalised linear model that took into account summary statistics on the total number of patients with a likelihood of discharge of over 80%, 90% and 95% within 24-hours to predict the total number of expected discharges (figure 1).

Model validation and performance evaluation

The model, developed using the training set, was validated on the testing set. Model performance was measured using median absolute error (MAE) and median absolute percentage error (MAPE). To evaluate model performance for binary predictions (such as the prediction of discharge within 24 hours for a given patient), we measured sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). To measure the contribution of each of the three main features of the model (temporal features, number of predicted admissions and number of predicted discharges) to the overall prediction, a linear regression, with each variable separately and with all variables combined, was used. Then, the adjusted R^2 was measured for each variable and for all variables combined.

Secondary analysis: procedure duration and LOS

Alongside the main census prediction modelling framework described above, which aimed to predict census within the next 24 hours, we also developed a separate LOS prediction model in order to study the factors affecting a patient's LOS in the surgery unit. Specifically, we aimed to study the relationship between different surgical procedures and the LOS in the surgical unit following the procedure. For this analysis, we focused only on patients who underwent a surgical procedure prior to their admission.

First, we calculated the median LOS and IQR for each individual procedure, for all procedures of a given department and for each key-term extracted from the procedure's description. In addition, we calculated the
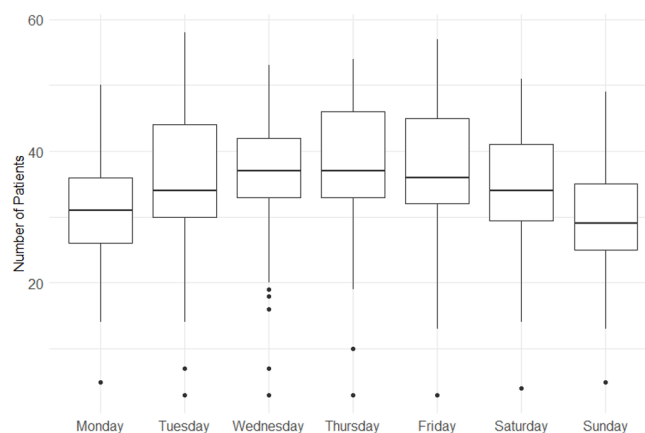

Figure 2 Occupancy of surgical department by day of week. This box-plot shows the median number of patients in the surgical unit at 8:00 hours each day, by day of week.

Spearman's r correlation coefficient between procedure duration and LOS. We used the Spearman correlation since the distribution of the variables was not found to be normal according to the Shapiro-Wilk normality test. Finally, to assess the ability of these different factors to predict a patient's total LOS, a random forest model was developed. The model was based only on procedure related data (online supplemental appendix A), using the same training/validation sets as before. The accuracy of the prediction was measured in terms of MAPE and MAE.

RESULTS

Variations in current census

There were 19 642 encounters in the surgical unit during the study period. Of these, 15 260 were used for training (78%) and 4382 used for validation (22%). The median number of patients in the surgical unit each day at 8:00 hours was 33 (IQR 28–40). The daily census varied from a minimum of 12 patients to a maximum of 58 patients per day, and the daily change in number of patients varied between a decrease of 20 patients to an increase of 17 patients.

Occupancy did not vary significantly from month to month (p=1), though weekday-to-weekend variations were more pronounced (p<0.001). On Sundays, the median number of patients was 27 while on Thursdays the median was 36 (figure 2). Variations in census were also noted around major US holidays. In all holidays but Columbus day, there was a decrease in the number of patients on both the day before and the day of the holiday. For most holidays, there was an overall decrease in the number of patients during the week surrounding the holiday (ie, no immediate compensatory increase in patients was observed after the holiday for the decrease in patients before the holiday). The holidays with the greatest decrease in the number of patients were Christmas Day (−8.4 patients), Thanksgiving Day (−8.2 patients) and Memorial Day (−6 patients). A detailed

summary of the effect of each holiday on occupancy is presented in online supplemental appendix B.

Predicting admissions

The median rate of admissions to the surgical unit was 12 patients per day (IQR 9–15). Ninety per cent of the patients came from three departments: PACU with 53% of admissions, ED with 31% of admissions and the surgical ICU with 6% of admissions. The admission prediction model was able to predict the number of new admissions to the surgical department with a MAPE of 16.7% and a MAE of two patients per day. The most significant factors identified by the model were the schedule of planned surgeries, the number of patients in the PACU and the day of the week.

Predicting discharges

The median rate of discharges from the surgical unit was 12 patients per day (IQR 9–15). The discharge prediction model was able to predict next-day discharge for individuals with an AUC of 0.84. At 90% specificity, the model was able to identify 57% of discharges with a PPV of 81% and NPV of 73% (online supplemental appendix C). The top factors associated with discharge timing included the number of days since admission (also representing the time since procedure), duration of the surgical procedure prior to admission and whether the procedure was planned or not.

As expected, 'discharge summary' orders were highly predictive of discharge with an OR of 13.4 for next-day discharge. Other less-intuitive orders associated with next-day discharge included those with descriptions containing the terms 'XR elbow' (OR of 5.6 (4.0–7.7)) and orders including the terms 'soft diet' (OR of 5.9 (4.9–7.0)). In contrast, patients receiving total parenteral nutrition, or any intravenous medications were unlikely to be discharged within 24 hours (OR <0.01). A summary of the top 40 orders to predict discharge (or lack of) can be found in online supplemental appendix C. Incorporating these individual-patient discharge predictions into a prediction of the overall discharges from the surgical unit, the generalised linear model was able to provide predictions with an error rate of 21.4% (MAPE) or a median of three patients per day (MAE).

Predicting next-day census

Integrating all of the above information into the census prediction formula (next-day census = current census + expected admissions – expected discharges), we were able to predict the next-day census in the validation set with a median error rate of 7% or three patients per day (IQR 1–5 patients per day) (figure 3). A prediction of an increase in the census was correct 85% of the time, while a prediction of a decrease in the census was only correct 60% of the time. The model was able to explain about 60% of the daily variability in census (R^2=0.58), with most of the variability explained by the day of week and nearby holidays (R^2=0.35), followed by the number of predicted
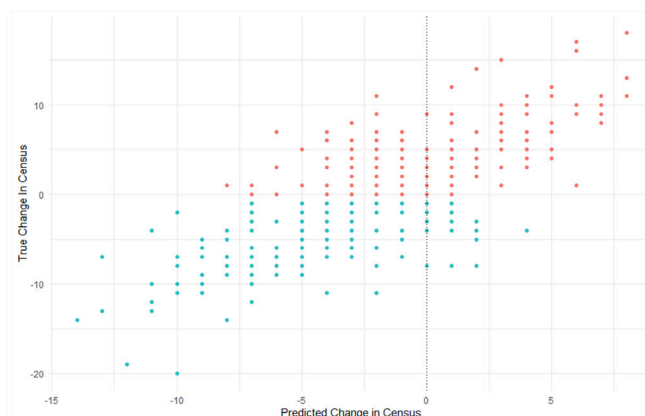

Figure 3 Predicted versus actual daily change in census during the validation time period. Observations to the left of the vertical dotted line (x=0) are days in which the model predicted a decrease in census. Observations to the right of the line are days in which the model predicted an increase in census. Days in which the actual census decreased are shown in blue (below y=0), while days in which the actual census increased are shown in red (above y=0).

admissions (R^2=0.26) and the number of predicted discharges (R^2=0.08).

Secondary analysis: surgery duration and LOS

As described in the methods section, alongside the main analysis of census prediction, we also developed a separate model in order to examine the factors affecting the overall LOS among patients who underwent a surgical procedure prior to their admission. Of all included surgical-unit patients, 68% (13 439/19 642) underwent a surgical procedure prior to their hospitalisation. Of these, 47% were for orthopaedics, 21% were for general surgery, 17% orofacial surgery, 9% plastic surgery, 2% gastrointestinal (GI) surgery and 2% genitourinary procedures (GU). Sixteen patients underwent more than one procedure. The 6229 admissions (32%) that were not preceded by a surgical procedure were either followed by a surgical procedure later during the hospitalisation (6.5%) or were without any documentation of a surgical procedure during the specific hospital encounter (25.5%).

Patients stayed in the surgical unit for a median of 1.7 days (IQR 0.9–3.3). LOS varied by the type and duration of the preadmission procedure. For example, the median LOS for patients undergoing 'medical' procedures (eg, placement of a central line) was 5.8 days, compared with only 1 day for those undergoing ophthalmology procedures (p<0.001). Patients undergoing spine halo application stayed for a median of 21 days (IQR 10–34), while patients undergoing tympanomastoidectomy stayed for a median 17 hours (IQR 12–22 hours, p<0.001). Procedures with descriptions that included terms such as 'exploratory' (eg, 'exploratory laparotomy') were associated with longer stays (median 6.1 days, IQR 2–10) while those that included terms such as 'percutaneous' (eg, 'elbow closed reduction with percutaneous pinning') were associated with shorter stays (22 hours
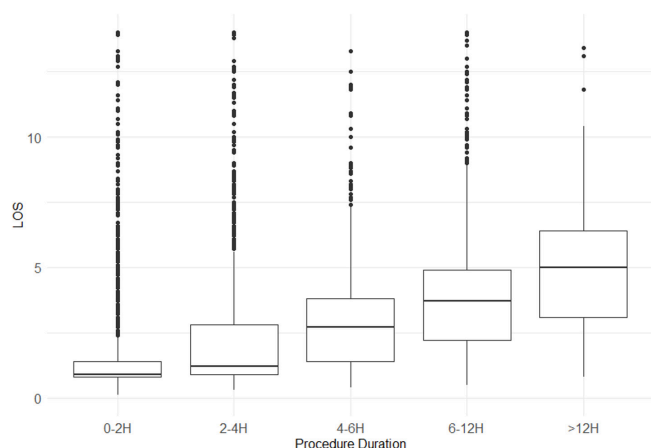
Figure 4 Length of stay (LOS) in the surgical unit by duration of surgical procedure. Boxplot showing the median LOS in days by the procedure duration in hours.

on median, IQR 19–38, p<0.001). Further details can be found in online supplemental appendix D.

Procedure duration was more predictive of total LOS than procedure type or procedure description text: Patients who underwent a procedure that took over 12 hours stayed on average 6 days longer than patients who underwent a procedure that took less than 2 hours (LOS of 8.4 days (95% CI 6.4 to 10.5) vs 2.3 days (95% CI 2.1 to 2.4) (figure 4)). Similarly, within a specific procedure type (eg, appendectomy), patients undergoing shorter procedures were likely to stay for a shorter time than those who underwent longer procedures (21 hours stay when the appendectomy took less than 1 hour vs 2 days when the appendectomy took longer than 1 hour, p<0.001). A detailed review of the correlation between procedure duration and the LOS can be found in online supplemental appendix D.

In online supplemental appendix E, we provide summary statistics for this separate model that predicts overall LOS based on procedure data. This model, separate from the main modelling framework, was able to predict LOS with a MAPE of 36% and MAE of 0.8 days, with the procedure duration serving by far as the most important factor for the prediction.

DISCUSSION

In this study, we developed and validated a model to predict next-day census in the surgical unit of a large tertiary paediatric hospital. We used a hierarchical modelling framework in order to isolate and evaluate the importance of individual prediction components (temporal, current census, discharges, admissions) and to evaluate which predictive factors best predict each of the components. This model was able to predict admissions with a median error of 16.7% (two patients per day), predict discharges with a median error of 21% (three patients per day) and predict next-day census with a median error of 7% (three patients per day). The factors found to be of most significance in predicting

next-day census were the day of the week, followed by the number of predicted admissions, and finally the number of predicted discharges.

In addition to the prediction of next-day census, a particular focus was given for the prediction and analysis of the LOS after surgery. Based only on procedure-related data, a separate random forest model was able to predict LOS with a median error of 36% or 0.8 days per patient. By far the most significant factor for this prediction of LOS was the procedure duration. Longer procedures were associated with longer hospital stays, both when comparing all procedures only by their duration and when comparing the same procedure for different durations. These results are consistent with prior studies that have shown an association between procedure duration and total LOS,[10 19 20] yet these other studies were all carried out in an adult population where different factors drive procedure duration.[21] The association between procedure duration and LOS is not surprising, as major procedures are expected to take longer than minor procedures and to require longer in-hospital recovery time. Studies also suggest that prolonged procedures are associated with an increased risk for perioperative surgical-site infections, thus contributing even more to the overall LOS.[10]

The hierarchical modelling framework used in this study enables better understanding of the factors driving capacity issues and prediction accuracy. This approach also lends itself well to stepwise model implementation, from basic top-down temporal features such as day-of-week and holidays to bottom-up features extracted from detailed order sets using advanced natural language processing tools. In this study, we show the contribution of these different 'building blocks' and highlight the importance of easy-to-obtain temporal information to the overall prediction. In addition, we provide a systematic investigation of paediatric surgical procedures with respect to postoperative hospital LOS. We show that the LOS can be predicted with good accuracy based on procedures data alone, and that procedure duration is highly predictive of the total LOS for most, but not for all paediatric procedures. Since paediatric surgical procedures differ greatly from adult procedures, and since previous studies have focused on the adult population, our results can provide an important contribution to the field of paediatric capacity planning.

This study has several limitations. First, it is a single-centre retrospective study of a tertiary paediatric hospital, and thus, the findings may not be applicable to other settings. Second, while the present model was developed to predict next-day census, other settings may require different timescales for prediction—from the prediction of hourly changes in census, to long-term prediction of the census in the next week, month or year. These different timescales will require retraining of the model according to the desired outcomes. Third, only procedures completed prior to the surgical-unit admission were included in the models, possibly omitting valuable information from the prediction. Nevertheless, we would

expect the models to only improve if this information would be added in future implementations. Fourth, as in any prediction model, the census prediction model is not 100% accurate. When used in clinical practice, clinicians and administrators can incorporate its predictions as one of several inputs in their decision-making processes. Lastly, our data included only data recorded prior to the COVID-19 pandemic, and thus may not be representative of data collected during the pandemic.

As the cost of building inpatient bed spaces continues to rise and financial pressure on paediatric hospitals increases, efficient utilisation of existing inpatient spaces becomes increasingly vital for healthcare sustainability. Predictive tools make the smoothing of elective procedures feasible and enable proactive planning to align staffing and other expensive resources. Similarly, in facilities with excess capacity, predictive tools help minimise waste of resources during periods of low occupancy. Facility planning and long-range staffing strategies will be more accurate with the help of high-performance system-wide prediction models. As healthcare resources are not stretched too much, quality and patient safety can be improved greatly.

We have shown that surgery occupancy prediction in a paediatric setting is plausible. We have further shown that surgical unit occupancy is dependent on both top-down temporal factors as well as bottom-up individual-patient information, including data on surgical procedures planned and performed, occupancy in other related departments, and clinical and administrative orders given to admitted patients. A hierarchical modelling framework that combines both types of factors has the potential to be better suited for predicting future surgical-unit occupancy, supporting decision-makers in their quest for improved scheduling, staffing and resource planning, reducing overcrowding and cancellations of surgeries in paediatric healthcare settings.

ORCID iD
Yuval Barak Corren http://orcid.org/0000-0002-9603-4728

REFERENCES

1 Wong DJN, Harris SK, Moonesinghe SR, *et al*. Cancelled operations: a 7-day cohort study of planned adult inpatient surgery in 245 UK National health service hospitals. *Br J Anaesth* 2018;121:730–8.
2 Gillies MA, Wijeysundera DN, Harrison EM. Counting the cost of cancelled surgery: a system wide approach is needed. *Br J Anaesth* 2018;121:691–4.
3 Herrod PJJ, Adiamah A, Boyd-Carson H, *et al*. Winter cancellations of elective surgical procedures in the UK: a questionnaire survey of patients on the economic and psychological impact. *BMJ Open* 2019;9:e028753.
4 Hanna TP, King WD, Thibodeau S, *et al*. Mortality due to cancer treatment delay: systematic review and meta-analysis. *BMJ* 2020;371:m4087.
5 Kortbeek N, Braaksma A, Burger CAJ, *et al*. Flexible nurse staffing based on hourly bed census predictions. *Int J Prod Econ* 2015;161:167–80.
6 Weissman C, Scemama J, Weiss YG. The ratio of PACU length-of-stay to surgical duration: practical observations. *Acta Anaesthesiol Scand* 2019;63:1143–51.
7 Baek H, Cho M, Kim S, *et al*. Analysis of length of hospital stay using electronic health records: a statistical and data mining approach. *PLoS One* 2018;13:e0195901.
8 Kasotakis G, Schmidt U, Perry D, *et al*. The surgical intensive care unit optimal mobility score predicts mortality and length of stay. *Crit Care Med* 2012;40:1122–8.
9 Schneider EB, Hyder O, Wolfgang CL, *et al*. Provider versus patient factors impacting Hospital length of stay after pancreaticoduodenectomy. *Surgery* 2013;154:152–61.
10 Tan T-W, Kalish JA, Hamburg NM, *et al*. Shorter duration of femoral-popliteal bypass is associated with decreased surgical site infection and shorter hospital length of stay. *J Am Coll Surg* 2012;215:512–8.
11 Clague JE, Craddock E, Andrew G, *et al*. Predictors of outcome following hip fracture. admission time predicts length of stay and in-hospital mortality. *Injury* 2002;33:1–6.
12 Djurasovic M, Kiskaddon E, Bratcher KR, *et al*. Determinants of hospital length of stay in patients undergoing lumbar spinal fusion. *The Spine Journal* 2014;14:S103–4.
13 Zheng F, Cammisa FP, Sandhu HS, *et al*. Factors predicting hospital stay, operative time, blood loss, and transfusion in patients undergoing revision posterior lumbar spine decompression, fusion, and segmental instrumentation. *Spine* 2002;27:818–24.
14 Dexter F, Epstein RH, Shi P. Forecasting the probability that each surgical case will either be ambulatory or the patient will remain in the hospital overnight having a length of stay of two or more days. *Cureus* 2020;12:e10847.
15 Shapiro NL, Bhattacharyya N. Complications and determinants of length of stay for inpatient pediatric otolaryngologic procedures. *Arch Otolaryngol Head Neck Surg* 2003;129:169–72.
16 Khare M, Rauch DA. Trends in national pediatric bed census. *Pediatrics* 2017;140:20.
17 Liaw A, Wiener M. Classification and regression by randomForest. *R News* 2002;2:18–22.
18 R Core Team. R: a language and environment for statistical computing. Vienna, Austria R Foundation for Statistical Computing; 2020. https://www.R-project.org/
19 Garbarino LJ, Gold PA, Sodhi N, *et al*. The effect of operative time on in-hospital length of stay in revision total knee arthroplasty. *Ann Transl Med* 2019;7:66.
20 Collins TC, Daley J, Henderson WH, *et al*. Risk factors for prolonged length of stay after major elective surgery. *Ann Surg* 1999;230:251.
21 Bravo F, Levi R, Ferrari LR, *et al*. The nature and sources of variability in pediatric surgical case duration. *Paediatr Anaesth* 2015;25:999–1006.