Predicting Case Volume from the Accumulating Elective Operating Room Schedule Facilitates Staffing Improvements

Vikram Tiwari, Ph.D., William R. Furman, M.D., Warren S. Sandberg, M.D., Ph.D.

ABSTRACT

Background: Precise estimates of final operating room demand can only be made 1 or 2 days before the day of surgery, when it is harder to adjust staffing to match demand. The authors hypothesized that the accumulating elective schedule contains useful information for predicting final case demand sufficiently in advance to readily adjust staffing.

Methods: The accumulated number of cases booked was recorded daily, from which a usable dataset comprising 146 consecutive surgical days (October 10, 2011 to May 7, 2012, after removing weekends and holidays), and each with 30 prior calendar days of booking history, was extracted. Case volume prediction was developed by extrapolation from estimates of the fraction of total cases booked each of the 30 preceding days, and averaging these with linear regression models, one for each of the 30 preceding days. Predictions were verified by comparison with actual volume.

Results: The elective surgery schedule accumulated approximately three cases per day, settling at a mean ± SD final daily volume of 117 ± 12 cases. The model predicted final case counts within 8.27 cases as far in advance as 14 days before the day of surgery. In the last 7 days before the day of surgery, the model predicted the case count within seven cases 80% of the time. The model was replicated at another smaller hospital, with similar results.

Conclusions: The developing elective schedule predicts final case volume weeks in advance. After implementation, overly high- or low-volume days are revealed in advance, allowing nursing, ancillary service, and anesthesia managers to proactively fine-tune staffing up or down to match demand. (ANESTHESIOLOGY 2014; 121:171-83)

VARIABILITY in daily surgical case volume suboptimizes the resources planned for the day of surgery because staff scheduling decisions are usually made weeks in advance. Typically, managers plan for the maximum demand, staffing all their allocated operating rooms (ORs). As the day of surgery approaches, managers may try to close ORs that have no booked cases, or consolidate ORs that have few cases, and thereby developing an opportunity to reduce staff labor costs. However, OR managers and anesthesia leaders typically can estimate final demand with precision only 1 or 2 days in advance of the day of surgery. By then, it is too late to plan OR closures and flex staff off or into other assignments.[1–3]

Most work on surgical volume prediction has focused on longer-term planning. To maximize longer-term efficiency in the OR suite, staffing is rationally planned by evaluating the tradeoff between underutilization and overutilization costs.[4,5] In the long run, if the distribution of OR labor costs and surgical demand (both in terms of case volume and procedure type) replicates the historical distributions used to develop the initial staffing parameters, OR efficiency will be maximized. Both the monthly aggregated surgical volume[6] and the individual surgical subspecialty's 4-week volume[7] can be fairly accurately predicted using statistical methods, and these can be used for planning monthly staffing. However, variability in surgical case volume in the short term (weekly and daily time scales) provides an additional opportunity to extract further improvements in matching OR costs to revenue opportunities.

What We Already Know about This Topic

- Precise estimation of operating room demand more than a few days in advance would allow flexible staffing decisions

What This Article Tells Us That Is New

- In a review of 146 consecutive surgical days at one academic medical center, case volume could be predicted with high accuracy 1 to 2 weeks in advance, allowing a closer match of staffing to demand

Although there are no published studies on this topic, stories of entire surgical services attending meetings or interview days, or highly busy surgeons in nonacademic centers going on vacations, with consequent unanticipated shortfalls in OR volume, are commonplace. Often, there is a rebound upswing

This article is featured in "This Month in Anesthesiology," page 3A. Corresponding article on page 6.

Submitted for publication September 2, 2013. Accepted for publication March 21, 2014. From the Department of Anesthesiology, Vanderbilt University School of Medicine, Nashville, Tennessee.

of demand on the days surrounding such absences, straining staff resources. Advance notification of the OR manager would allow planning for such events, but common experience hints that this is vulnerable to the limits of human diligence. Managers of the OR and ancillary services such as the supply chain, pathology, and radiology need a tool to predict total surgical case volume several days to a few weeks in advance and to give daily insight into OR suite volume for the 1 to 2 weeks leading up to surgery so that they can make intermediate-term adjustments that match capacity to demand as precisely as possible. To be useful, the tool should predict the total scheduled elective volume (with an allowance for the expected number of urgent and emergent add-ons that will appear just before and on the day of surgery), and it should make that prediction accurately, weeks in advance.

We speculated that the developing elective OR schedule, which accumulates over time as multiple surgical clinics add cases to fill the available time in their allotted blocks, contains information useful for predicting the ultimate aggregated elective and nonurgent add-on OR volume for each day. In other words, if a given OR day's volume differs from the canonical (historical daily) volume, that difference would have been reflected days to weeks in advance in that day's elective schedule as it developed. For ORs with some degree of labor flexibility, the ability to predict the aggregated daily surgical case volume 1 to 2 weeks in advance would be an important signal for adjusting staffing. We therefore undertook a study to determine, working from the elective schedule as it develops over time, whether: (1) surgical case volume can be predicted, and if so, (2) with what confidence, (3) how many days in advance, and (4) if the predictions could be used to plan OR closures and reduce OR staff labor expenses by cancelling shifts or shifting staff to other duties, locations, or assignments. Predicting the number of day-of-surgery add-on cases or the hours of overtime on the day of surgery was not in the scope of the project; rather, we sought to minimize their impact by giving OR managers more certainty about the elective workload in the weeks and days leading up to the day of surgery itself. We performed this work as an engineering project with observational assessment of impacts in the live, working environment of a 55 OR academic medical center (Vanderbilt University Hospital, Nashville, Tennessee—hereafter referred to as the "adult hospital"), with subsequent extension to the adjacent free-standing 18 OR pediatric hospital (Monroe Carell Jr. Children's Hospital at Vanderbilt, Nashville, Tennessee—hereafter referred to as the "children's hospital").

Materials and Methods

This study was reviewed by the Vanderbilt University Institutional Review Board (Nashville, Tennessee) and was determined to qualify as a nonhuman research study.

Setting and Prior State Description

This work was performed at a level 1 trauma center, with a mean ± SD daily adult surgical volume of 117 ± 12 cases (95% CI of mean = 2 cases) for the period from which data were used to develop the model. The average case count and the day-to-day variability in case numbers can be appreciated in figure 1. Median case volume was 117 cases, 25th percentile was 108 cases, and the 75th percentile was 126 cases. The mean daily case volume was also 117, and the distribution of daily volume was not statistically significantly different from a normal distribution using the Shapiro–Wilk test of normality ($P = 0.11$; the threshold for determination that a distribution is statistically different from a normal distribution using this test is generally considered to be $P < 0.05$). Total daily case volume was highly correlated with total case hours. Using 30 months of data from our institution, we found the Pearson's R correlation to be 0.92. Therefore, daily case volume is a good proxy of the daily workload. We also examined the day-to-day variability of daily case volume using Run Chart methodologies to determine whether this parameter demonstrated common cause or special cause variation.

On average, 6% of total cases are unscheduled day-of-surgery add-ons (added within the last 15 h before the start of the day of surgery, after the elective schedule closes). On the day of surgery, add-on cases are scheduled into holes in the elective schedule (which are rare because the OR scheduler works to eliminate holes during the week leading up to surgery) or into the ends of staffed OR blocks. Overtime utilization has been an accepted practice to complete cases for the past decade. Similarly, our OR practice is not to maintain an open room for add-on cases, and this has been consistent for the past decade. Accordingly, all nonurgent, urgent, and emergent cases added on the day of surgery have been handled the same way throughout the duration of the project: nonurgent cases are performed as time becomes available, urgent cases in the first reasonable room, and emergencies are booked into the first open room and the displaced cases worked into the OR schedule in other rooms on the day of surgery.

In the scheduling office, elective cases are fitted into allocated block time whose "end" is sufficiently elastic to allow modest overruns (approximately 1 h) at the discretion of the scheduling office manager. Seven calendar days before the day of surgery, all unused block time is released. At this point, the OR schedulers begin to fit elective add-on cases for surgeons who do not have block time on that particular OR day into the elective schedule. This scheduling technique has been found to be effective in filling unused block times in large ORs.[7,8]

Before the reported project, we developed monthly volume forecasts based on historical data combined with expert judgment based on business volume projected by each service line. These monthly volume projections were used to create the budgeted volume, which was then apportioned over the number of weekdays in a given month to derive the projected daily volume. Developing monthly budgeted volume in this manner has some theoretical validity[9] and satisfactorily estimates the monthly volume. However, these budgeted numbers fail to provide timely predictions of the daily and weekly variability in case volume.

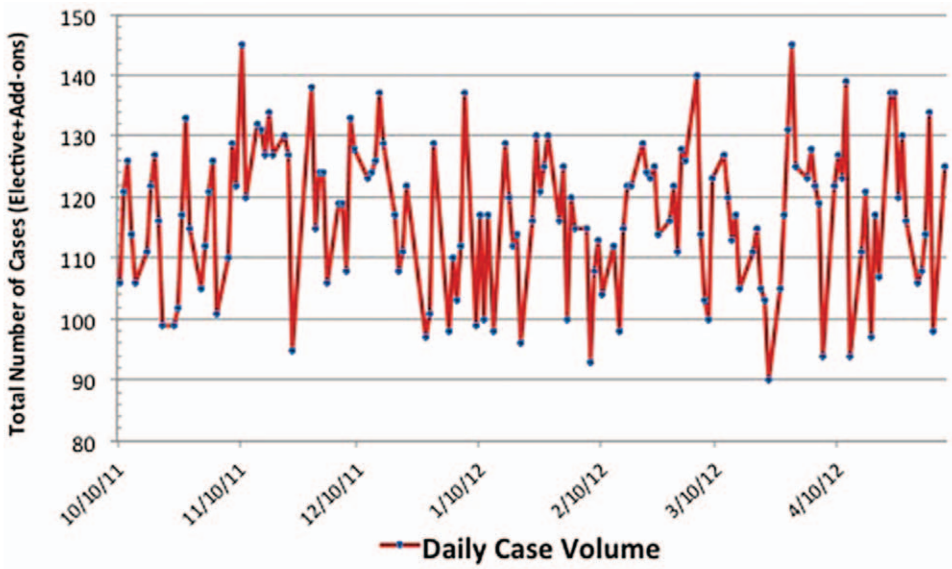


Fig. 1. Time series of daily surgical volume (adult hospital). Raw time series of the total cases (including day-of-surgery add-ons) performed each day in the 146 weekday run of data used to create the case count prediction model. Swings of 30 or more cases between nearby weekdays are common. Weekend days and holidays are not shown.

Data Collection

To assess expected future volume many days in advance of the day of surgery, we captured snapshots of the developing daily OR schedule, for each OR work day, as it existed at the end of each of the 30 days leading up to that day of surgery. To ensure consistency, OR schedule snapshots were taken at exactly 4:30 PM each day. This time was chosen because it is when the OR scheduling office closes for the day and hence captures the set elective schedule for any given day at the last time the schedulers work on it until the beginning of the next work day. Data were collected for 8 months (September 2011 to May 2012). We excluded OR days at the beginning of data collection for which the complete prior 30 days of case bookings into that surgery date were unavailable. After removing weekends and holidays, we ended with 146 consecutive surgical days; each with 30 prior (calendar) days of case-booking history (4,380 observations). For each scheduled case, the dataset consisted of surgery date, initial booking date, and the major organizational unit of our OR suite to which the case had been assigned.

Evaluating Alternative Volume Prediction Models

The objective of this project was to develop a technique to predict case volume multiple days to weeks in advance of the day of surgery. Two approaches guided the evaluation and development of volume prediction models. First, we used established time-series techniques, specifically, Autoregressive Integrated Moving Average, to see whether future days' volume could be predicted by just checking one or a few immediately past days' volume. We then explored a second approach that was motivated by the opinion held (though never scientifically tested) by expert anesthesiologists that the accumulating elective surgical case schedule may hold signals that if identified could perhaps help predict the final volume weeks before the actual day of surgery. Operationalizing this second general approach, we developed three models (two linear regression models and a linear extrapolation model) and selected two of these for inclusion in the final prediction tool.

Time Series of Daily Case Volume

We used time-series analyses to assess whether the final daily case volume itself could be used to predict the volume on future days. We used Expert Modeler of IBM SPSS v. 21 (IBM Corp., New York, NY) to explore which, if any, autocorrelation models could predict OR volume.

Linear Models of the Developing Schedule

We next developed and assessed three model approaches designed to take account of the temporal buildup of the OR schedule to predict the final case volume. The first approach ("Linear Trend Model") was a regression-based linear trend model. In the second modeling approach ("Percentage of Final Volume"), predicted surgical volume was a fixed proportion of the number of cases booked for each of the 30 preceding days. To account for the occasional periods of nonlinear trends in case bookings, especially in the final few days before the day of surgery, the distribution of the percentage of cases booked as of any day (up to 30 days) before the day of surgery was calculated and its mean used to predict the final case volume using a linear projection. The third modeling technique ("Days-Out Models") was also based on simple linear regression, but in this case, we developed a separate model for each of the 30 days preceding the day of surgery. Both the second and third approaches are recursive each prediction day. In other words, each model is recalculated on successive days to use the most current information about the number of cases booked by the date the prediction is generated for a given future surgery date. The final prediction was calculated by combining predictions from the two individual modeling techniques that proved valid. We developed the original modeling approach for our adult hospital and then tested for generalizability by reimplementing it at our free-standing children's hospital.

Statistical Analysis

After removing data for weekends and holidays, a sequential series of 146 surgery days was obtained. For each of these 146 days, the accumulated number of cases booked each day going back to 30 calendar days from the day of surgery was obtained. All individual models were validated with split-sample techniques. For prospective validation, volume predictions from the model were compared with actual corresponding daily volumes using 3 months of data (December 31, 2012 to March 29, 2013). Data are presented as mean ± SD, number (%), or 95% CI. Associations between time lags in the autoregressive time series are measured using autocorrelation, with their significance measured using the Ljung Box Q significant value. Stationary R^2 provides the fit of the time-series model. In regression analysis, the assumption of normality, of either the response or the explanatory variables, is not required; however, we checked for normality using the Shapiro–Wilk test and found all explanatory variables and the response variable to be mildly normal (null hypothesis is that data are from a normal distribution; the smallest P value was 0.12). Regression model fits were measured by the R^2 (which indicates the proportion of variance in the independent variable that is explained by the dependent variable), the standard error of the estimate, and the significance of the F value. Durbin–Watson test was used to confirm the lack of autocorrelation among the regression error terms, and Q-Q plots of residuals were examined to confirm normality of error terms. Presence of constant error variance was confirmed by examining scatter plots of the dependent and independent variables. Regression coefficients' significance was tested using t tests. Comparison of model precision between hospitals was performed using the Mann–Whitney U test. All analyses were performed using Microsoft Excel 2010 for Windows (Microsoft Corp., Redmond, WA), IBM SPSS for Windows v.21 (IBM Corp.), and JMP Pro 10.0.1 Release:2, 64-bit Edition (SAS Institute Inc., Cary, NC).

Results

Time Series of Daily Case Volume

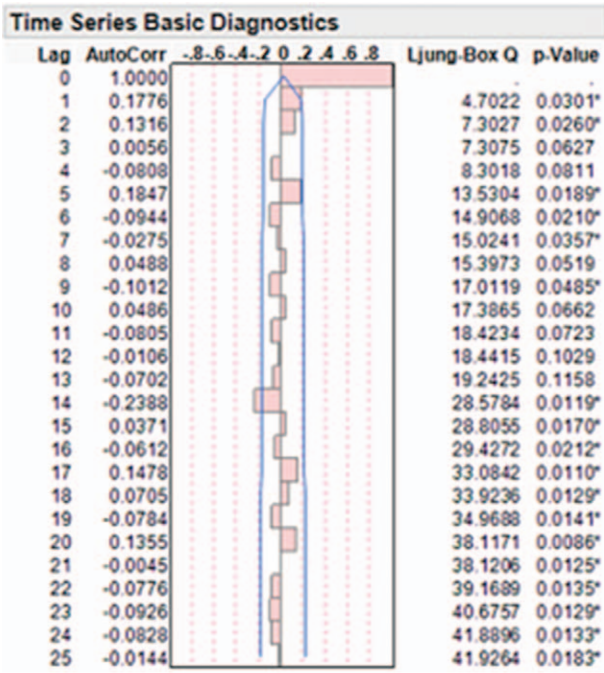
Figure 1 gives the time series of the daily case volume. Total weekday case volume is widely variable and does not demonstrate any obvious patterns visible from inspection of figure 1. We also confirmed this statistically through Run Chart methods (not shown), which indicated a process operating within control limits, with exceptions around holidays, exhibiting variability that can mostly be attributed to common causes.

Autoregressive models can forecast outputs of those systems that exhibit persistence or autocorrelation in a time series; that is, some subset of past values (in our case, past few days' daily case volume) can be used to predict future states (in our case, future daily volume). Given that our objective was to predict OR volume as much as two weeks in advance, time-series analyses did not seem promising as an initial approach. Nevertheless, we evaluated time-series–based autoregressive models for the daily volume; however, this method of predicting a day's case volume based solely on immediately preceding days was inadequate. The autocorrelation analysis given in figure 2 shows that lags 1, 2, 5, 6, 7, 9, and 14 were statistically significant (α = 0.05); however, the autocorrelations themselves are very small and too small to be operationally meaningful. We tested other autoregression models (using different lags and moving averages) with the actual day-of-surgery volume, but none was a good fit. The best-fitting model generated was Autoregressive Integrated Moving Average (1, 0, 14), which had a Ljung Box Q significant value of 0.224, a Stationary R^2 of 0.117 and a mean absolute error of 9.2. In the appendix, we provide a brief explanation of the terms used to define Autoregressive Integrated Moving Average models.

The lack-of-fit of autoregressive models was expected because in OR suites that have capacity allocation based on block schedules, the daily volume is unlikely to be strongly correlated with any of the immediately preceding days' case volume. To check the effect of weekly block schedules on daily volume, we also tested 5-day seasonal Autoregressive Integrated Moving Average models. The model fit improved, but the models still lacked the ability to identify higher (or lower) than average case volume days far in advance (appendix). This validated the basic premise for this study that dynamic techniques, possibly based on the real-time status of current volume booked, might be superior to the traditional forecasting models based on static historical information.

Linear Trend Model

Visual data examination shows that the accumulation of surgical cases into the elective surgery schedule seems to follow an almost linear trajectory until 1 or 2 days before the day of surgery. As an example of this trend, see figure 3, which shows the number of cases booked for each day of the week of July 22, 2013, starting from 14 days before the day of surgery. In the last 2 days before the day of surgery, the number of cases added is best described as random within a range (0% to +20% of the final case volume). To evaluate this linear trend, we tested a regression-based linear trend model. Using time (t) as the single explanatory variable, we developed an estimate for the final count (Y_t), $Y_t = a + bt + E_t$, where a is the intercept, b is the slope, and E_t is an error term. We took the mean of the number of cases booked for



Time Series Basic Diagnostics

Lag	AutoCorr	Ljung-Box Q	p-Value
0	1.0000	.	.
1	0.1776	4.7022	0.0301*
2	0.1316	7.3027	0.0260*
3	0.0056	7.3075	0.0627
4	-0.0808	8.3018	0.0811
5	0.1847	13.5304	0.0189*
6	-0.0944	14.9068	0.0210*
7	-0.0275	15.0241	0.0357*
8	0.0488	15.3973	0.0519
9	-0.1012	17.0119	0.0485*
10	0.0486	17.3865	0.0662
11	-0.0805	18.4234	0.0723
12	-0.0106	18.4415	0.1029
13	-0.0702	19.2425	0.1158
14	-0.2388	28.5784	0.0119*
15	0.0371	28.8055	0.0170*
16	-0.0612	29.4272	0.0212*
17	0.1478	33.0842	0.0110*
18	0.0705	33.9236	0.0129*
19	-0.0784	34.9688	0.0141*
20	0.1355	38.1171	0.0086*
21	-0.0045	38.1206	0.0125*
22	-0.0776	39.1689	0.0135*
23	-0.0926	40.6757	0.0129*
24	-0.0828	41.8896	0.0133*
25	-0.0144	41.9264	0.0183*

Fig. 2. Correlogram for daily surgical volume time series. Autocorrelation analysis of the 146 weekday run of case count data used to create the final case count autoregressive time-series models. Autocorrelation analysis was performed to determine whether each day's observed volume could be predicted from the case counts of prior days. *Asterisks* indicate prior (lag) days for which the current day volume shows a statistically significant autocorrelation.

each of the 30 days before a day of surgery (termed TMinus30 to TMinus1 for convenience) for the 146 days of data. As an example, for $t = 1, Y_{30} = \frac{\sum_{n=1}^{146} y_n^{30}}{146}$, where y_n^{30} is the nth data point (out of 146) corresponding to number of cases booked 30 days before the day of surgery (TMinus30). The value of Y_{30} = 31 cases, that is, 30 days before the day of surgery (t = 1), the average number of cases booked was 31. At TMinus29 (t = 2) Y_{29} = 33 cases.

The linear fit was good (adjusted R^2 = 0.97, standard error = 4.1, intercept coefficient = 20.81, regression coefficient = 2.63 with P = 0 at α = 0.05). This implied that the rate of new case booking was approximately three new cases per day. Although this simple model was useful, this was inadequate as it lacks the ability to exploit real-time information about the changes in the linear trajectory of case bookings as additional cases get booked each day.

This model's output provided useful insights about the almost linear nature in which cases appear on the schedule. However, it suffered from the problem of "averages of averages," which made its output susceptible to misinterpretation due to the wide swings in case volume, and therefore restricted its ability to be used in a live environment. This technique was therefore not included in making the final predictions.

Percentage of Final Volume Model

To account for the occasional periods of nonlinear trends in case bookings, especially in the final few days before the day of surgery, the distribution of the percentage of cases booked as of any day (up to 30 days) before the day of surgery was calculated, and the mean of this percentage was then used to predict the final case volume using a linear projection. For example, figure 4 shows that 66% of the time 7 days before the day of surgery (TMinus7), between 70 and 85% of the final case volume, has already been booked (mean = 73.3%, 95% CI of the mean = 1%). Therefore, at TMinus7, if 80 cases have been booked, the predicted final volume can be estimated as: $80 / 0.733 \cong 110$. Use of the most current information in this manner allows either reinforcing the prediction made the previous day(s) or updating our estimates for the day-of-surgery volume.

Days-Out Models

A different methodology, based on recursively predicting the final case count by using the most updated information available, was devised. Prediction models based on simple linear regression were developed. A separate prediction model referenced to each of the 30 days before the day of surgery was created, where the final case volume (dependent variable) was regressed on cases booked until each of the 30 preceding days (independent variable).

Table 1 gives the regression output for each of the 30 regression models. Figure 5 shows the scatter plot for number of cases booked 1, 5, 14, and 30 days before the day of surgery and the final volume. Each scatter plot has 146 data points, a regression line with its 95% CI zone (darker shaded area), and the 95% prediction zone (lighter shaded area).

The final prediction model of OR volume is then calculated by combining predictions from the percentage of final cases booked method (Percentage of Final Volume Model) with the recursively generated regression predictions (Days-Out Models) and taking the average of the two. This method of combining forecasts increases forecast accuracy by cancelling forecast errors from different forecasting methods.[10] Weighing the forecasts equally, that is, averaging them, has been demonstrated to be as accurate as more complex weighing schemes.[10–12]

Table 2 shows the result of prospective model validation (December 31, 2012 to March 13, 2013) against actual total number of OR cases performed. We found that the mean absolute error was between 3.75 cases (for predictions made 1 day before the day of surgery) and 8.27 cases (for predictions made 14 days before the day of surgery). In the last 7 days before the day of surgery, our model was able to predict the actual final case count within ±7 cases (6% of mean number of cases) 80% of the time. The model underpredicted the final actual total case count by more than seven cases (in the final 7 days before the day of surgery) less than 10% of the time (see fig. 6 for representative model output). Recursively predicting the final volume each day (by using the prediction

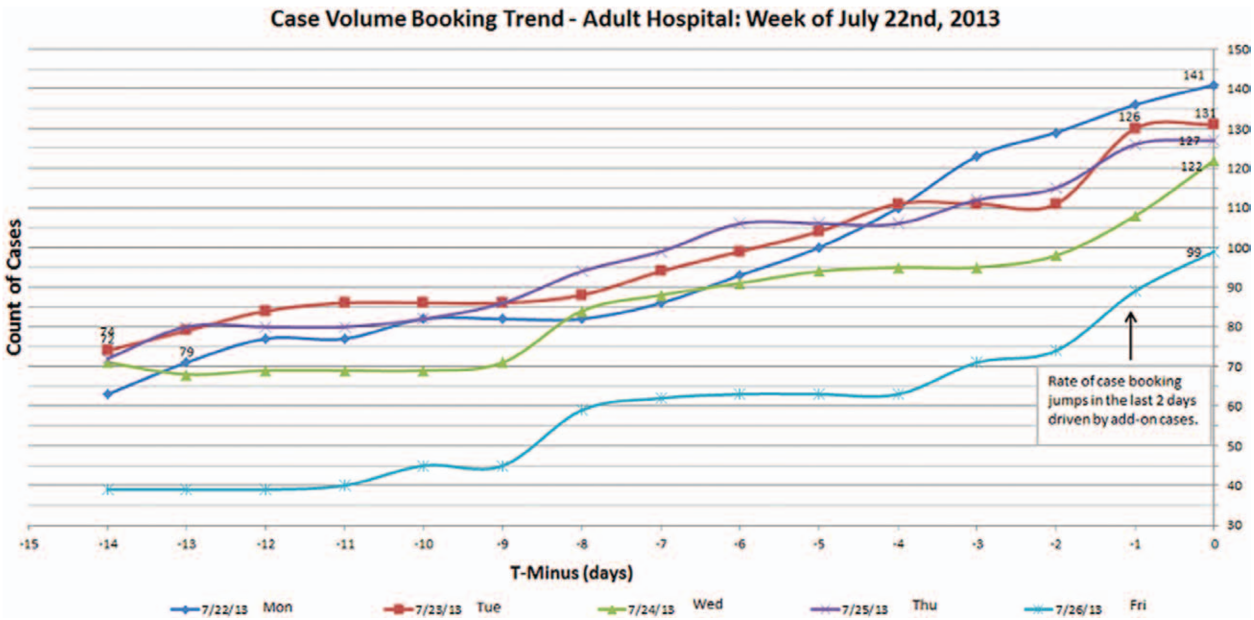


Fig. 3. Scheduled surgery cases booked for each day of the week of July 22, 2013 from 14 days-out. The count of elective cases booked for each day by the end of every calendar day for all 14 days (including weekend days) leading up to the actual day of surgery (day 0) are shown. The counts of cases at day 0 are the actual cases performed.

model appropriate for that day and taking into account the most recent case booking information) allowed OR managers to see any shifts in the predictions over the 2 weeks leading up to the day of surgery. This helped build face validity for the model's output with the end users (OR managers, pathology managers, and supply chain mangers). During the pilot phase (January to March 2013), model output was shared twice weekly with OR leaders, who enthusiastically began to use it for planning. Consistently low (or high)-volume days stand out because separate regression models for each of the 30 days leading up to the surgery day are created. This imparts predictive validity to the model's output. Despite this quantifiable validity, the easy perceptibility of the output imparted by its visual representation generated the much needed face validity required to get buy-in for implementation of the model.

The model's predictions are sufficiently accurate, especially 4 to 6 days before the day of surgery, to allow managers to adjust staffing by rescheduling, allowing total staff numbers to be adjusted downwards (or upwards). After a successful pilot phase, including consultation with OR operational managers about the tool's performance, the perioperative leadership made an executive decision to have the TMinus5 predictions reported out widely. We made this decision because the mean absolute error at Tminus5 during the pilot phase of the model's implementation was 4.92 (table 2), better than any other estimate of daily volume. The model's 5-day-out predictions (T-5) are now reported in the Daily Case Report (fig. 7) right next to the budged case volume. Predictions for 1 through 14 days in advance are also provided in a figure similar to figure 6. Table 3 shows that the mean absolute error of the model's 5-day-out prediction (T-5) was 7.17 cases, as opposed to the budgeted numbers' error of 9.47 cases (May 1 to July 25, 2013). This difference roughly translates to 1 OR-day worth of cases every day. For the first 25 days of July 2013, the prediction model predicted cases within ±7 cases 78% of the time (at Tminus5), whereas budget-based predictions were within seven cases 50% of the time (table 3).

Using the same methodology, we developed a similar case volume prediction model for our free-standing children's hospital. We refined the modeling technique by creating separate calculations for each day of the week (*e.g.*, Mondays calculated separately from Tuesdays, separately from Wednesdays, and so on), as opposed to the all-days-combined model created for the adult hospital. Preliminary analysis with 1 month of prospective data suggests that the model's output is at least as good, if not better than the adult hospital's model. For the period July 22 to August 21, 2013, the mean absolute error for TMinus5 is three cases, which is 4.5% of the average daily volume. Because the pediatric hospital's daily case volume is substantially smaller (approximately 50% of the adult hospital's volume), an appropriate metric for comparison among the models of the two hospitals is the median absolute percentage error. For the period July 1 to August 29, 2013, predictions from the adult hospital's model had a median absolute percentage error of 3.57%, compared with 2.78% for predictions from children's hospital's model. The distribution of absolute percentage error was found to be nonnormal. Therefore, to test whether there is statistically significant difference between the prediction error medians from the two models, the nonparametric Mann–Whitney U test was used. No significant difference was found: *P* value (asymptotic) = 0.61, so we cannot reject the null hypothesis

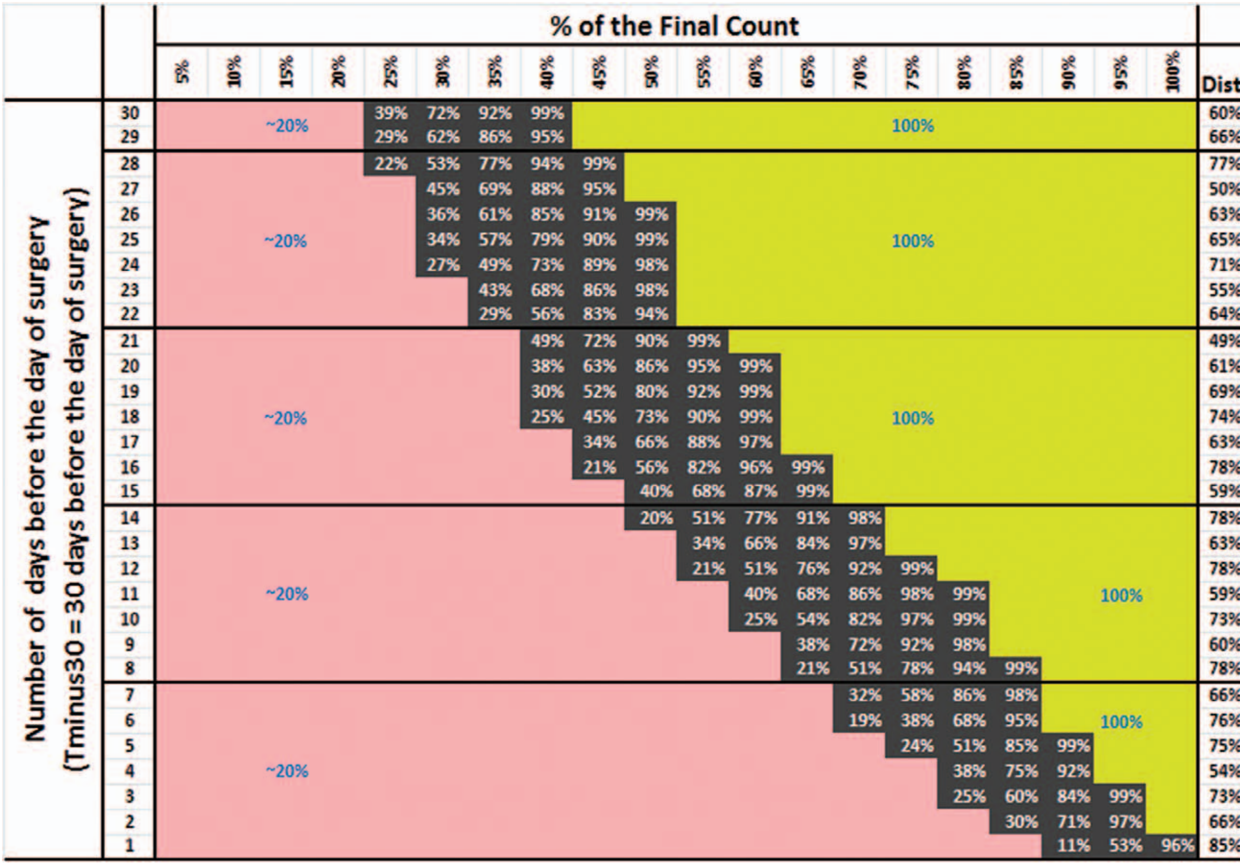


Fig. 4. Nomogram of percentage of cases booked as a function of days before the surgery date, in the baseline data used to create the "Percentage of Final Count" model. Thirty days before the day of surgery (TMinus30), 39% of the time 25% of the final case volume was found to have been already booked. Similarly, 72% of the time, 30% of the final volume was booked. The number 60% under "Dist." column is the difference between 99 and 39%, implying that 60% of the time, between 25 and 40% of the final case volume was booked at TMinus30. Mean = 26.4 ± 0.5% (95% CI of mean = 0.1%). If, for example, at TMinus30, 30 cases were already scheduled, then just using the "Percentage of Final Count" model, the following prediction would be made—"there is an absolute certainty that the final volume will not be less than 75 (=30/0.4) cases, and there is 60% probability that the final volume will be between 120 (=30/0.25) cases and 75 (=30/0.4) cases."

Table 1. Summary Output of 30 Linear Regressions (Days-Out Models)

	Model Summary			Unstandardized Coefficients				
Model Name	R^2	Std. Error of the Est.	Durbin–Watson	Constant/ Coefficient	B	Std. Error	*t*	Sig.
TMinus01	0.87	4.41	1.92	(Constant)	11.63	3.48	3.34	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.95	0.03	30.50	0.00*
TMinus02	0.76	5.85	1.80	(Constant)	30.06	4.07	7.39	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.86	0.04	21.57	0.00*
TMinus03	0.68	6.82	1.76	(Constant)	40.90	4.41	9.28	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.78	0.05	17.45	0.00*
TMinus04	0.66	7.07	1.65	(Constant)	44.63	4.42	10.09	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.76	0.05	16.55	0.00*
TMinus05	0.67	6.97	1.70	(Constant)	47.52	4.16	11.43	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.75	0.05	16.91	0.00*
TMinus06	0.66	6.97	1.77	(Constant)	48.92	4.08	11.98	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.76	0.05	16.88	0.00*
TMinus07	0.65	7.13	1.71	(Constant)	48.82	4.23	11.55	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.79	0.05	16.32	0.00*

(*Continued*)

Table 1. (*Continued*)

Model Name	Model Summary: R^2	Model Summary: Std. Error of the Est.	Model Summary: Durbin–Watson	Unstandardized Coefficients: Constant/ Coefficient	B	Std. Error	*t*	Sig.
TMinus08	0.64	7.27	1.70	(Constant)	50.86	4.23	12.02	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.81	0.05	15.82	0.00*
TMinus09	0.65	7.18	1.62	(Constant)	52.17	4.07	12.83	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.83	0.05	16.16	0.00*
TMinus10	0.57	7.91	1.62	(Constant)	60.57	4.17	14.54	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.75	0.06	13.75	0.00*
TMinus11	0.52	8.31	1.64	(Constant)	63.62	4.31	14.75	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.74	0.06	12.57	0.00*
TMinus12	0.53	8.24	1.59	(Constant)	64.25	4.19	15.33	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.75	0.06	12.79	0.00*
TMinus13	0.51	8.40	1.58	(Constant)	64.98	4.30	15.13	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.77	0.06	12.31	0.00*
TMinus14	0.48	8.65	1.63	(Constant)	66.64	4.41	15.11	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.78	0.07	11.61	0.00*
TMinus15	0.47	8.74	1.57	(Constant)	68.77	4.32	15.93	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.79	0.07	11.37	0.00*
TMinus16	0.43	9.06	1.63	(Constant)	72.96	4.28	17.05	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.76	0.07	10.49	0.00*
TMinus17	0.39	9.39	1.55	(Constant)	77.36	4.21	18.36	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.71	0.07	9.61	0.00*
TMinus18	0.36	9.65	1.52	(Constant)	80.73	4.14	19.48	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.68	0.08	8.96	0.00*
TMinus19	0.37	9.55	1.50	(Constant)	80.86	4.02	20.12	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.70	0.08	9.21	0.00*
TMinus20	0.37	9.54	1.55	(Constant)	81.34	3.96	20.56	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.72	0.08	9.24	0.00*
TMinus21	0.35	9.68	1.50	(Constant)	82.66	3.97	20.81	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.72	0.08	8.86	0.00*
TMinus22	0.35	9.70	1.51	(Constant)	83.83	3.86	21.72	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.73	0.08	8.83	0.00*
TMinus23	0.31	10.02	1.46	(Constant)	87.42	3.82	22.91	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.69	0.09	7.98	0.00*
TMinus24	0.26	10.36	1.46	(Constant)	91.47	3.72	24.58	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.62	0.09	7.09	0.00*
TMinus25	0.22	10.62	1.49	(Constant)	94.51	3.64	25.97	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.57	0.09	6.41	0.00*
TMinus26	0.23	10.59	1.47	(Constant)	94.03	3.67	25.60	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.60	0.09	6.48	0.00*
TMinus27	0.23	10.57	1.47	(Constant)	93.87	3.67	25.57	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.63	0.10	6.53	0.00*
TMinus28	0.22	10.66	1.50	(Constant)	94.11	3.77	24.93	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.66	0.10	6.28	0.00*
TMinus29	0.20	10.75	1.50	(Constant)	95.09	3.75	25.35	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.67	0.11	6.06	0.00*
TMinus30	0.19	10.84	1.52	(Constant)	96.39	3.69	26.13	0.00*
		F-Statistic sig. = 0.00*		CasesSch	0.67	0.12	5.80	0.00*

Linear regression output for Days-Out Models. *F*-Statistic is significant for all models at 95% confidence level; however, predictability of final volume is higher for models where the explanatory variable (CasesSch) refers to days that are close to the day of surgery (as shown by lower values of the standard error of the estimate). Durbin–Watson values are all approximately 2, implying no autocorrelation among the error terms. The constant term (intercept of the regression line) and the regression coefficient (B) are statistically significant at 95% confidence level (*P* values are negligible for all 30 regression models). Predictor Variable: TMinusXX, where XX refers to the number of days before the day of surgery. Dependent variable: final case volume.

* Statistically significant at 95% confidence level.

of equality of the two medians. Conclusive opinion about superiority of one model over the other can only be made after further rigorous empirical tests, which will be conducted after accumulation of more data.

We began to share the model output, *via* the Daily Case Report, on August 13, 2013. Between August 13, 2013 and August 30, 2013, predictions of the final volume made 5 days before the day of surgery (TMinus5) were within ±7

Table 2. Comparing Prediction Model's Output against Actual Final Volume on the Day of Surgery

	Absolute Value of [Prediction Minus Actual] (Count of Cases)—December 31, 2012 to March 29, 2013													
	T-1	T-2	T-3	T-4	T-5	T-6	T-7	T-8	T-9	T-10	T-11	T-12	T-13	T-14
Mean absolute error (count of cases)	3.75	4.00	4.58	4.71	4.92	5.40	5.23	5.73	7.21	7.21	6.75	8.02	8.50	8.27
Mean absolute percentage error	3.1%	3.3%	3.8%	3.9%	4.1%	4.5%	4.4%	4.8%	6.0%	6.0%	5.6%	6.6%	7.0%	6.8%
Probability prediction exceeds actual by >7 cases	3.8%	3.8%	5.8%	9.6%	7.7%	15.4%	13.5%	19.2%	28.8%	21.1%	15.4%	19.2%	23.1%	17.3%
Probability prediction exceeds actual by 0 to 7 cases	53.8%	48.1%	51.9%	46.2%	50.0%	36.5%	44.2%	34.6%	32.7%	36.5%	40.4%	26.9%	21.2%	21.2%
Probability prediction is less than actual by 0 to 7 cases	36.54	40.4%	34.6%	36.5%	28.8%	36.5%	32.7%	34.6%	23.1%	26.9%	21.2%	28.8%	26.9%	32.7%
Probability prediction is less than actual by >7 cases	5.7%	7.7%	7.7%	7.7%	13.5%	11.5%	9.6%	11.5%	15.4%	15.4%	23.1%	25.0%	28.8%	28.8%
Probability prediction is between ±7 cases of the actual	90.4%	88.5%	86.5%	82.7%	78.8%	73.1%	76.9%	69.2%	55.8%	63.5%	61.5%	55.8%	48.1%	53.8%

Model performance against actual case counts during validation period. Mean absolute error of the model's output is calculated by taking the mean of the absolute value of the forecast errors, where errors are the difference between the predicted volume (for each of the 14 days preceding the day-of-surgery) and the final actual volume. Mean absolute percentage error is the mean of the absolute value of the individual errors, which are expressed as percentage of the actual final case volume. Prediction probabilities become increasingly accurate closer to the day of surgery.

cases 92.3% of the time, as compared with 61.54% for the budgeted volume predictions, which are calculated based on longer-term historical data. During this same period, the mean absolute error of the TMinus5 predictions for the Children's Hospital was 4 ± 3.26 cases (Min: 0 case, Max: 12 cases), as compared with 8 ± 6.14 cases for the budgeted predictions (Min: 1 case, Max: 22 cases).

Discussion

We used an engineering approach to develop a prediction model of final OR case volume based on accumulation of the elective schedule. This model demonstrates that it is possible to predict day of surgery case volume 1 to 2 weeks in advance with sufficient confidence to make staffing decisions that affect organizational finances. This facilitates tuning planned capacity to more closely match actual demand.

Our OR suite managers routinely use the output of this model to identify both higher than average and lower than average-volume days, and make appropriate staffing decisions. For example, quantitative estimates for July 5, 2013 (a Friday) were known 30 days in advance, allowing confident staff planning. Historically, July 5 in general, and especially if it falls on a Friday, has been a low-volume day in our OR suite. The anecdotal expectation of our managers was that volume on Friday July 5 would be extremely low, much lower than the budgeted volume, and similar to what is typical after the Thanksgiving Day holiday, which always occurs on a Thursday. If we had simply looked at several prior years' OR case volumes for July 5, without using our modeling methodology, we would have had no way to predict how low (or high) the volume would be. Without a basis for prediction (provided by the model), we would only been able to guess and would have made a significant underestimate of demand. However, our model predicted only a mild volume shortfall for this particular July 5. Having developed enough confidence in the model's predictions, the OR managers planned staffing appropriate to a typical Friday, consistent with the model. Actual case volume for July 5 (fig. 7) proved to be within 2% of the median for the 5 Fridays in the month of July 2013, and our staff matched our needs.

The model has been especially useful in identifying unexpected low-volume days, as happened in the second week of May 2013, when all the surgeons in one specialty attended an out-of-town conference but did not provide advance notification to OR managers. A low-volume prediction (12 days before the day of surgery) triggered an immediate investigation confirming the cause. An alert was sent to all surgery clinics announcing extra OR capacity available on those days. All pending add-on cases were assigned to ORs and then excess rooms were closed (three on Monday, May 6, and four on Tuesday, May 7). The original study hypothesis was that accurate volume predictions weeks in advance would

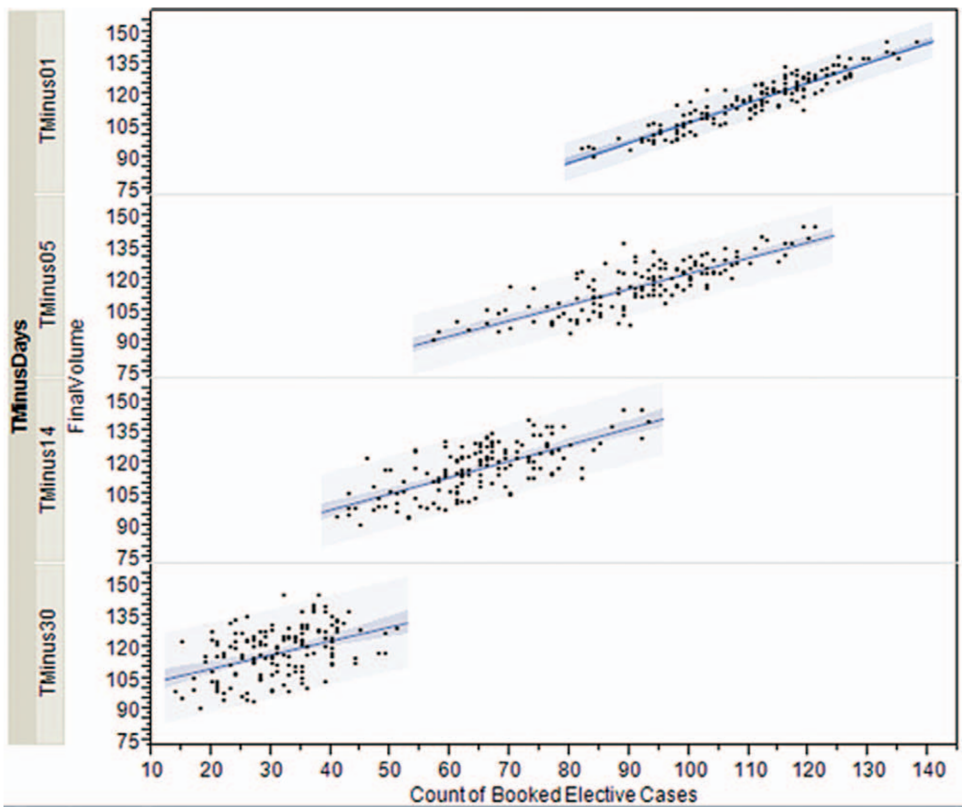


Fig. 5. Scatter plot of booked elective cases (at T-30, T-14, T-5, and T-1) and final case volume. Linear regressions of current booked elective case count *versus* actual final case count for 4 illustrative days before the actual day of surgery. Even as far back as TMinus30 days, there is a clear relation between the number of booked elective cases and the final case volume on the day of surgery in the operating rooms. 95% CI and prediction interval bands are shown as *dark* and *light areas*, respectively.

enable the cost savings by flexing staff off duty in response. However, the potential to replace revenue by temporarily reallocating capacity among services and by accommodating wait-listed cases with alacrity (and the potential to avoid lost revenue that might occur if surgeons defer cases because they do not expect to be able to get them on the schedule) is also very appealing to the medical center's leadership.

The utility of the case volume predictions extends beyond OR staff planning. Case count predictions are also used by our ancillary and downstream units whose workload is driven by OR volume. Managers at the case cart preparation center, with access to temporary staffing, use this information as they plan daily operations. In addition, advanced warning of higher-volume days prompts the case cart preparation center to proactively check for instrument conflicts among scheduled cases 5 to 7 days in advance, rather than looking for these conflicts on the day before surgery. This has helped reduce case cancellations and delays related to equipment availability conflicts. Our surgical pathology laboratories use the case volume predictions to adjust daily staff assignments in advance of surgery. Finally, the Department of Anesthesiology is operationalizing a new staffing plan based on the model predictions. Prior practice had been to staff for maximum demand by scheduling attending anesthesiologist coverage of every anesthetizing location to which the department has committed coverage. The model now allows long-term planning for the maximum demand minus two to four locations (out of a total of 78 anesthetizing locations per day) depending on model case volume projections, with an "invoke-able" buffer (additional staff on call in reserve). This plan saves two attending anesthesiologists per day (2.8 Full Time Equivalents per year). In practice, we have been able to cover the year's organic growth in case volume (approximately 5%) without increasing the size of the anesthesia faculty. Avoiding even a single faculty hire at current rates saves at least \$400,000.[13] Such savings are providing external validation to this effort comparable with other meaningful savings from information technology interventions in OR operations.[14]

The prediction model reduces the uncertainty about near-term procedural workload by providing daily visibility of expected OR volume. It has encouraged the lead scheduling OR nurse manager to more aggressively and confidently use the scheduling policies (which are described in Materials and Methods section) to accommodate cases requested by surgeons who do not have block time on the requested day of surgery by assigning their cases to rooms several days in advance, to steer add-on elective cases away from high-volume days where unused,

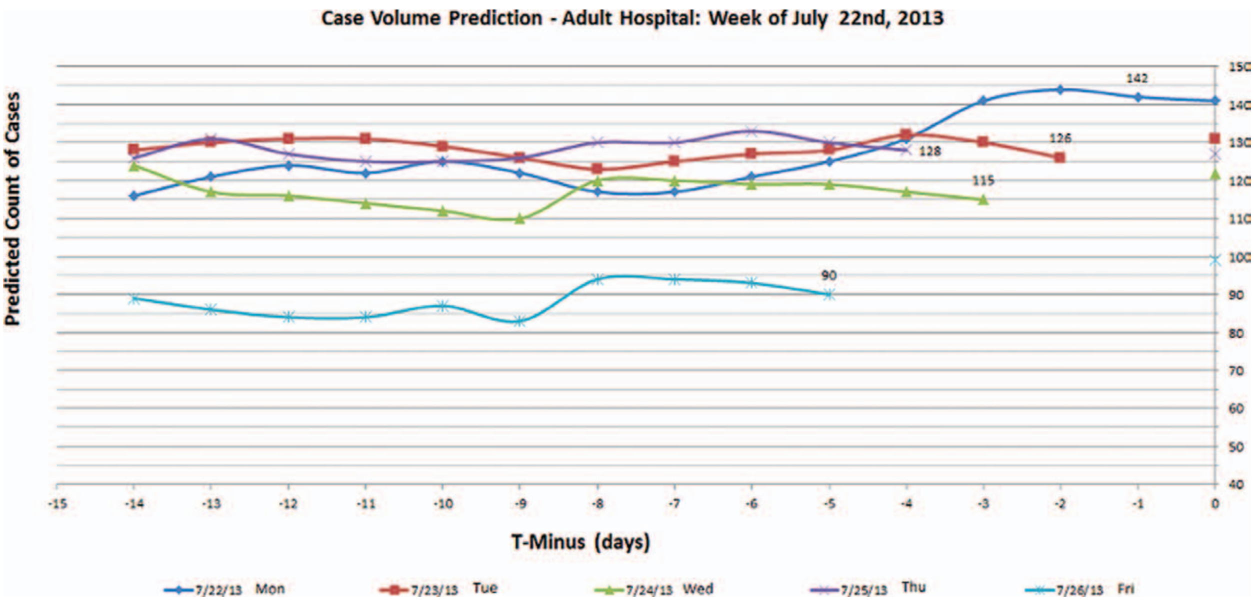


Fig. 6. Predictions for 5 weekdays in July, for the week ending on July 26, 2013, beginning 14 days from July 26. Managers are shown individual predictions of case volume for each weekday, beginning at T-minus 14 days from the Day of Surgery. Each estimate is recursively calculated each day, updated with new information from the previous day's bookings (fig. 3) for each Day of Surgery. The prediction view is shown as it would appear to managers on the morning of July 22, 2013, with the most recent prediction for Monday (7/22) being the T-Minus 1 prediction and T-Minus 5 being the most recent prediction for Friday (7/26). The actual final case counts for these days (which are the TMinus0 numbers in fig. 3) were Monday: 141, Tuesday: 131, Wednesday: 122, Thursday: 127, and Friday: 99. Managers tracking the predicted case count every day from 14 days out would have accurately concluded that Monday, Tuesday, and Thursday would be above average-volume days, with Wednesday right at the average, and Friday will be a low-volume day.

open block time would be hard to find, and to inform nurse managers of upcoming low-volume days so that staff can be flexed off several days in advance. Between December 2012 and July 2013, as a result of better labor-management policies, labor costs per case, as measured by our centralized finance department, have decreased by approximately 15%. Whether some or all of these reductions can be attributed to the availability of future volume predictions will be quantitatively analyzed once sufficient prospective data are available.

One potential limitation of this work is that it may not be generalizable to other hospitals. However, we have validated the approach in two distinct environments. The applicability of the methodology at two hospitals, each of which follows its own distinct surgical demand and elective case-booking techniques, shows generalizability and potential of the technique to be replicable at other institutions.

The almost linear nature in which the aggregated volumes of elective cases get booked every day, especially up to the last few days before the day of surgery, is interesting. Why aggregated elective case bookings follow an almost linear trajectory remains an unanswered question and is a focus of our further research. Higher than average-volume days continue to have cases booked at a higher rate (and not the average rate of three cases per day); and this trend is visible 2 to 3 weeks in advance. Similarly, for low-volume days, the rate of elective case accumulation is lower than for average-volume OR days. Our conjecture is that on days (and weeks) when the majority of the surgeons are in town (and not absent due to vacation, conference travel, or medical leave), the volume of surgical cases will be higher than average, and *vice versa*, but that this pattern can be detected weeks in advance.

Although more complex response functions might have provided a better fit than the simple linear functions we used, in the absence of any theoretical basis to force higher-order polynomial functions on the underlying data, we chose simplicity over abstractness and complexity. Several empirical published studies on thousands of time series have shown simpler forecasting techniques to be more effective than complex methods.[15,16] These reports document a number of empirical studies on literally thousands of time-series forecasts where complex methods fared no better, and often worse, than simple methods.

As a final point, this research answers the call[17,18] to do more translational research[19] in the delivery of health care. Conceptual insights from academic research are informative, but often are static and therefore nontransferable to widespread practice. There is a need to translate research findings into meaningful operational use. Translational healthcare delivery research, as distinct from "applied research" (where the problem addressed is a real-world issue with a narrow focus, and therefore less generalizable), is an iterative way of doing research that necessitates close interaction between

Total Cases	Month to Date				Mon	Tue	Wed	Hol	Fri	Sat	Sun	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Mon
Service Rollup	Actual	Budget	Var	Var %	7/01	7/02	7/03	7/04	7/05	7/06	7/07	7/08	7/09	7/10	7/11	7/12	7/13	7/14	7/15
Total Actual	2,884				118	136	125	13	108	14	13	134	126	131	133	125	20	12	135
Total Budget	2,815				123	123	123	15	97	15	15	123	123	123	123	123	15	15	123
Total Variance	69			2%	(5)	13	2	(2)	11	(1)	(2)	11	3	8	10	2	5	(3)	12
5-Day Out Prediction, Total					113	132	115		93			128	121	131	135	119			135

Fig. 7. Example of the daily case volume report provided to medical center leadership and to managers across the enterprise. The 5-days-out predictions (T-5) are inserted in the Daily Case Report, so that operating room managers and anesthesia leaders can see predicted volume for the next 5 days. The full 14 days-out predictions are also made available at the end of this daily report *via* charts similar to those in figure 6. In the above cropped prospective view of July 2013 report, the predictions are highlighted in *red*, the budgeted volume is highlighted in *gray*, and the final volume is highlighted in *yellow*. For the first 15 days of July (excluding weekends), the model's predictions (at TMinus5) were within ±7 cases of the final volume 80% of the time, whereas the budgeted volume was within ±7 cases only 40% of the time.

Table 3. Comparing Accuracy of Budgeted Volume to Model's Predicted Volume (at TMinus5 days): May–July 2013

	Actual *vs.* Budget	Actual *vs.* TMinus5
Mean absolute error	9.47	7.17
Mean absolute error (rounded)	9	7
SD	6.77	5.83
Median	8	6
Probability intervals for accuracy of model's predictions: comparing budgeted to TMinus5 predictions for July 1 to July 25th, 2013 (only weekdays)		
Probability prediction exceeds actual by >7 cases	44.4%	22.2%
Probability prediction is within ±7 cases of the actual	50.0%	77.8%
Probability prediction is lesser than actual by >7 cases	5.6%	0.0%

Pairwise comparison of model's predictions at TMinus5 days *vs.* actual case volume compared with budgeted volume *vs.* actual case volume for the same interval. The comparison includes each working weekday in the period May 1 to July 25, 2013. Budgeted daily volume is based on past 3 yr of historical volume, adjusted for expert estimation of planned growth due to addition of surgeons and programs, and spread over days of the week based on block allocation and historical case counts. The model predicted final count within seven cases 77.77% of the time, as compared with 50% for the budget-based predictions during the period July 1 to 25, 2013.

practitioners and researchers. Translational research progresses iteratively and builds on the successful implementation of the research findings of previous steps before determining the direction for the subsequent steps. The driving force for the current "engineering solution" research was to address the issue (faced by most medium- to large-sized ORs) in such a manner that the methodology and solutions could be readily adopted and implemented on the fly to make an immediate impact on practice. Since the case prediction report demonstrated in figures 6 and 7 was created, it has "gone viral" within our institution, and the predictions are now provided (by request) to approximately 50 managers across the enterprise, from pathology to bed management. Although there is considerable anecdotal evidence on the usefulness of these predictions for daily operational planning and staffing, after further continuous usage of the predictions by various subunits at our institution, it would also be possible to quantitatively measure the impact on staffing and other operational costs.

Acknowledgments

Support was provided solely from institutional and/or departmental sources.

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Tiwari: Department of Anesthesiology, Vanderbilt University, 1211 21st Avenue South, MAB 403G, Nashville, Tennessee 37212. vikram.tiwari@vanderbilt.edu. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

References

1. Guerriero F, Guido R: Operational research in the management of the operating theatre: A survey. Health Care Manag Sci 2011; 14:89–114
2. Nelson R, Kennedy MS: The other side of mandatory overtime: 'Flexing down' means nurses are losing money and patience. Am J Nurs 2008; 108:23–4
3. Dexter F, Shi P, Epstein RH: Descriptive study of case scheduling and cancellations within 1 week of the day of surgery. Anesth Analg 2012; 115:1188–95
4. Strum DP, Vargas LG, May JH, Bashein G: Surgical suite utilization and capacity planning: A minimal cost analysis model. J Med Syst 1997; 21:309–22
5. Strum DP, Vargas LG, May JH: Surgical subspecialty block utilization and capacity planning: A minimal cost analysis model. Anesthesiology 1999; 90:1176–85
6. Dexter F, Traub RD: The lack of systematic month-to-month variation over one-year periods in ambulatory surgery caseload—Application to anesthesia staffing. Anesth Analg 2000; 91:1426–30
7. Dexter F, Macario A, Traub RD, Hopwood M, Lubarsky DA: An operating room scheduling strategy to maximize the use of operating room block time: Computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time. Anesth Analg 1999; 89:7–20
8. Mazzei WJ: Maximizing operating room utilization: A landmark study. Anesth Analg 1999; 89:1–2
9. Dexter F, Macario A, Qian F, Traub RD: Forecasting surgical groups' total hours of elective cases for allocation of block time: Application of time series analysis to operating room management. Anesthesiology 1999; 91:1501–8
10. Clemen RT: Combining forecasts: A review and annotated bibliography. Int J Forecasting 1989; 5:559–83

11. Mahmoud E: Accuracy in forecasting: A survey. J Forecasting 1984; 3:139–59
12. Albright SC, Winston WL, Zappe C: Data Analysis and Decision Making with Microsoft Excel. Edited by Curt Hinrichs. Pacific Grove, Duxbury Press, 2002, pp 707
13. Kheterpal S, Tremper K, Shanks A, Morris M: Seventh and eighth year follow-up on workforce and finances of the United States anesthesiology training programs: 2007 and 2008. Anesth Analg 2009; 109:897–9
14. Spring SF, Sandberg WS, Anupama S, Walsh JL, Driscoll WD, Raines DE: Automated documentation error detection and notification improves anesthesia billing performance. Anesthesiology 2007; 106:157–63
15. Armstrong S: Research on forecasting: A quarter century review, 1960–1984. Interfaces 1986; 16:89–103.
16. Schnarrs S, Bavuso J: Extrapolation models on very short-term forecasts. J Bus Res 1986; 14:27–36.
17. Reid PP, Compton WD, Grossman JH, Fanjiang G: Building a Better Delivery System: A New Engineering/Health Care Partnership. Committee on Engineering and the Health Care System, Institute of Medicine and National Academy of Engineering, Washington, D.C., The National Academies Press, 2005, pp 85–8
18. Valdez RS, Ramly E, Brennan PF: Industrial and Systems Engineering and Health Care: Critical Areas of Research—Final Report. (Prepared by Professional and Scientific Associates under Contract No. 290-09-00027U.) AHRQ Publication No. 10–0079. Rockville, Agency for Healthcare Research and Quality, 2010, pp 38
19. Woolf SH: The meaning of translational research and why it matters. JAMA 2008; 299:211–3

Appendix. Five-day Seasonal ARIMA Model Results

Technique Used in SPSS	Model Type with 5-day Seasonality Added	Stationary R^2	MAPE	MAE	Normalized BIC
ARIMA model fitting	ARIMA (1, 0, 0) $(1, 0, 0)_5$	0.05	8.148	9.29	5.189
Expert Modeler	Simple seasonal	0.55	8.257	9.45	4.97

SPSS and Expert Modeler part of the statistical software owned by IBM Corp., New York, NY. ARIMA (p,d,q) $(P,D,Q)_s$: the terms within (p,d,q) refer to the nonseasonal component of the series, whereas (P,D,Q) refers to the seasonal components. "p" is the number of time-period lags or the number of autoregressive terms, "d" refers to any differencing required to make the series stationary, "q" refers to the number of lagged error terms or the number of moving average lags. Seasonality effects are modeled by including (P,D,Q). "P" is the number of seasonal autoregressive terms, "D" is the number of seasonal differences, "Q" is the number of seasonal moving average terms. The subscript "s" refers to the time span of repeating seasonal order. The table above compares results of fitting (to the baseline data) the seasonal ARIMA (1, 0, 0) $(1, 0, 0)_5$ model and the SPSS's inbuilt Expert Modeler that automatically identifies and estimates the best-fitting ARIMA model and thus eliminating the need to identify an appropriate model through trial and error. Given the two models, the one with lower BIC value will be preferred. Difference in BIC value of <2 between the models being compared indicates that the models are very comparable.
ARIMA (p,d,q) $(P,D,Q)_s$ = Autoregressive Integrated Moving Average; BIC = Bayesian information criterion; MAE = mean absolute error; MAPE = mean absolute percentage error.

To check the effect of weekly block schedules on daily volume, we also tested 5-day seasonal ARIMA models. However, the limitation in using immediately previous days' final volume (even if by day of the week, which is the case when we include the 5-day seasonality effect) is that it fails to catch excessively high- or low-volume days far in advance. In fact, the Maximum Absolute Error in the above ARIMA models (including the one reported in the Results section and subsection Time Series of Daily Case Volume) is 29 cases. This error is far higher than the errors produced by our models developed based on the "accumulating surgical schedule." The focus of the research in our study is therefore on the development and implementation of the methodology for predicting final case count based on the accumulating surgical schedule and not based on using previous days' case count or previous day-of-the-week case count to predict the future case count.