Application of Time Series Analysis to Operating Room Management

Franklin Dexter, M.D., Ph.D.,* Alex Macario, M.D., M.B.A.,† Fang Qian, Ph.D.,‡ Rodeny D. Traub, Ph.D.§

Background: Allocation of the correct amount of operating room (OR) “block time” can provide surgeons with access to sufficient OR time to complete their elective cases while optimally matching staffing with the elective case workload (to maximize labor productivity). To evaluate how to predict accurately total hours of elective cases performed by a surgical group using data from surgical services information systems, the authors addressed the following questions: (1) How many previous 4-week periods of data should be used to minimize error in forecasting a surgical group’s total hours of elective cases? (2) Using the number of 4-week periods from question #1, can we detect trends or correlations between successive periods that could be used to improve forecasting accuracy? (3) How can results from questions #1 and #2 be used to calculate an upper prediction bound (upper limit) for the total hours of elective cases that will be completed in a future period? Prediction bounds can be used to budget staffing accurately.

Methods: Time series analysis was performed on total hours of elective cases over 39 consecutive 4-week periods from 17 surgical groups.

Results: The average of 12 consecutive periods’ total hours of elective cases had an appropriate error profile. The observations within each series of 12 consecutive 4-week periods followed a normal distribution, with each observation of total hours of elective cases not correlated with the subsequent observation.

Conclusions: The average of the most recent 12 4-week periods can be used to predict surgical groups’ future use of block time. (Key words: Operating room economics; staff scheduling; surgical services.)

The single greatest cost to a hospital of delivering surgical care occurs in the operating room (OR).¹ Salaries of OR staff (nursing and anesthesia) account for the majority of OR costs.² Within many hospitals and integrated health care delivery systems, an OR manager or governing body has the authority to minimize the cost of caring for surgical patients. To have an important impact on costs of patient care in the surgical suite, such OR managers try to maximize labor productivity by employing the fewest full-time nurses and anesthesiologists necessary to care for the patients.³ In this setting, the day on which to perform each elective case needs to be selected optimally to match the days on which full-time OR personnel are scheduled to work.⁵

One common method to provide surgeons with access to sufficient OR time to complete their elective cases and to match caseload with OR staffing is to allocate “block time” (i.e., OR time).³ Of ⁶ With block scheduling, OR time is assigned to a surgeon, surgical group, or surgical department for their exclusive use up to some designated time before the day of surgery (e.g., at 10 A.M. 2 working days before surgery). Labor productivity is maximized by filling the block time with as many hours of cases as possible.³ A key factor in filling block time is to forecast accurately how much block time to allocate to each surgeon or group of surgeons who share a block.⁵ Of ⁶ The surgical group needs sufficient block time to complete its elective cases. It is also important that the surgical group not be allocated too much extra time, which would increase the amount of time for which OR staff are scheduled but for which there are no elective cases. These goals require accurate forecasts of each surgical group’s future needs for OR time to complete its elective cases.

Models from management science that forecast a variable
Surgeons are scheduling cases into this period.

Forecasted Period

6 four-week periods used in forecast

Fig. 1. Historical data is used to forecast a future 4-week period, while skipping the 4-week period before the 4-week period being forecasted. In this example, the average of six 4-week periods is used to forecast total hours of elective cases for the eighth 4-week period while surgeons are scheduling patients into the seventh 4-week period.

(total hours of elective cases performed by a surgical group) based on past patterns of change of the variable over time are known as time series models. Total hours of elective cases can be forecasted by calculating the average total hours of elective cases over a predetermined past period (e.g., 11 months). In this study, we use scheduling data from our institution to address the following three questions. (1) How many previous 4-week periods of data should be used to minimize error in forecasting a surgical group's total hours of elective cases? For example, it is unknown whether using the previous 8 weeks or the previous 96 weeks of data minimizes forecasting error, the difference between actual and forecasted mean total hours of elective cases. Minimizing the error will result in the allocation of block time to surgical groups being more accurate. (2) Using the number of 4-week periods from question #1, can we detect trends or correlations between successive periods that could be used by more sophisticated statistical methods to improve the accuracy of the forecast? (3) How can results from questions #1 and #2 be used to calculate an upper prediction bound (upper limit) for the total hours of elective cases that will be completed in a future period? These limits would be useful, for example, to be 90% confident that sufficient OR staff have been budgeted to provide enough open OR time so that all of a surgical group’s patients can have surgery within 4 weeks of requesting to be scheduled for surgery.

Methods

Overview of the Time Series Analysis

We considered total hours of elective cases over 4-week periods (see Discussion for rationale). We propose using the average of total hours of elective cases over a specified previous number of 4-week periods to forecast the total hours of elective cases in a future 4-week period (fig. 1). If surgeons have assigned block time into which cases are scheduled, when using historical data to forecast a future 4-week period, the current 4-week period is skipped because surgeons are scheduling cases into that 4-week period (fig. 1). The forecast is updated every 4 weeks. In performing each update, the data from the most recent period are incorporated, and the oldest data are discarded.

Data Collection

The University of Iowa's surgical services information system tracks cases at a tertiary surgical suite and an Ambulatory Surgery Center. The 156-week interval between July 3, 1994, and June 28, 1997, contains 39 consecutive 4-week periods. During this interval, 27,453 elective cases were performed by nine groups of surgeons at the tertiary surgical suite and 13,470 elective cases were performed by eight groups of surgeons at the Ambulatory Surgery Center. The number of different surgeons in each group (surgical department) ranged from five in a group of hematologists performing bone marrow harvests at the Ambulatory Surgery Center to 73 in General Surgery at the tertiary surgical suite (including pediatric, vascular, gastrointestinal, and plastic surgery). The two surgical suites were scheduled independently. Some surgical groups operated at both suites. Seventeen combinations of group and suite were available for analysis, each with 39 consecutive observations of total hours of elective cases over 4-week periods.
For each elective surgical case, the data extracted from the OR information system were (1) the date on which the case started; (2) the time when the patient entered an OR; (3) the time when the patient left the OR; (4) the OR into which the patient entered; and (5) the surgical group. Case duration was defined to equal the number of hours from when the patient entered the OR until the patient left the OR. For forecasted total hours of elective cases to be relevant to block allocation, cases with durations longer than the duration of the University of Iowa’s regularly scheduled 8-h OR day were considered to have durations equal to 8 h.

Allocation of surgical block time needs to include the cases’ concomitant turnover times. We defined turnover time to equal the time from when a patient left an OR until another patient undergoing elective surgery entered the same OR, provided that the next patient entered during regularly scheduled hours of the same day. Turnover time was assigned to the surgical group of the preceding case. This definition of turnover time did not distinguish between cleaning and setup time versus avoidable or unavoidable delays. Therefore, interpreting the time interval from when one patient left an OR to the time when the next patient entered that OR as turnover exaggerated the time actually used to perform the needed cleanup and preparation between the two cases. Consequently, turnover times longer than 1 h were truncated to equal 1 h.

Data were extracted using the Access database language (Microsoft, Redmond, WA). The statistical analysis was performed in Excel (Microsoft).

**Question 1: How Many Previous 4-week Periods of Data Should Be Used To Minimize Error in Forecasting a Surgical Group’s Total Hours of Elective Cases?**

We evaluated a range of possible multiples of previous 4-week periods to determine an appropriate number of previous periods to use to forecast total hours of elective cases. We considered the forecasting error to equal the actual future value minus its corresponding forecasted value. We analyzed 17 different time series, one for each of the surgical groups caring for patients in the tertiary surgical suite (nine surgical groups) and Ambulatory Surgery Center (eight surgical groups). For each time series, we calculated the forecasting errors based on analysis of all 39 4-week periods in our database. We weighted each observation of total hours of elective cases over a 4-week period equally.

To determine forecasting errors for different previous number of 4-week periods, we considered all possible combinations of forecastings using the data. For example, for determining forecasting errors when using two previous periods of data, we used the first 8 weeks (first and second periods) to forecast from the 13th to 16th weeks (fourth period) for the 17 combinations of surgical group and suite. Likewise, we used the 5th to 12th weeks (second and third periods) to forecast from the 17th to 20th weeks (fifth period). At the beginning of period $p$, a forecast that used data from $n$ previous periods $(p - 1, p - 2, \ldots, p - n)$ forecasted expected demand for period $p + 1$. Overall, we obtained 612 assessments of forecasting error when using two periods of historical data ([17 combinations of surgical group and suite] × [39 periods − 2 periods for forecasting − 1 period skipped]). The use of more periods in generating a forecast permitted fewer assessments of forecasting error for that number of periods.

For each number of historical time periods used, we calculated three measures of the forecasting errors:

1. Mean ± SE of the forecasting errors. The mean of the forecasting errors assessed whether the forecast consistently overestimated or consistently underestimated the total hours of elective cases.
2. Mean ± SE of the absolute values of the forecasting errors. This measure is important because if the block time allocated to a surgical group is the group's forecasted total hours of elective cases, then the period's absolute forecasting error will equal the sum of unused block time and hours of cases performed outside of block time during the 4-week period. The financial cost to the surgical suite of forecasting errors can then be proportional to the mean of the absolute values of the forecasting errors, depending on how the OR staff are compensated (e.g., part-time hourly, full-time hourly, or salaried). This economic assessment of forecasting errors assumes that when block time is allocated, the surgical group can only schedule an elective case outside of the block time if there is not enough open time in the block to perform the case.
3. Mean ± SE of the absolute values of the percentage forecasting errors. This third measure was included because the absolute values of the forecasting errors as previously described gave larger values for surgical groups that had the largest numbers of total hours of elective cases in a suite. Expressing these values as percentages provided equal weight to all surgical groups.
Figure 2 gives an example of using 12 4-week periods to forecast total hours of elective cases for one of the 17 surgical groups.

Our primary goal in question #1 was to determine the optimal number of previous 4-week periods to use in forecasting total hours of elective case time for a surgical group. Using fewer 4-week periods (less historical data) would generally increase the mean of the absolute values of the forecasting errors because the method would respond to outliers and random noise. Using more 4-week periods would decrease the responsiveness of the forecasting method to increases or decreases in long-term tendencies of the mean of total hours of elective cases over each 4-week period. To balance these two objectives, we defined “optimal” as the smallest number of periods that minimized the mean of the absolute values of the forecasting errors.

**Question 2: Using the Number of 4-week Periods from Question #1, Can We Detect Trends or Correlations between Successive Periods That Could Be Used by More Sophisticated Statistical Methods To Improve the Accuracy of the Forecast?**

Data that are normally distributed can be described by their mean and SD. We evaluated whether the total hours of elective cases over each of the optimal number of previous periods contained readily detectable information besides the mean and SD that could be used to forecast future values. If there were such information, it would most likely be characterized by a correlation between successive observations of total hours of elective cases over 4-week periods. For example, if there were a trend (slow increase or decrease) in the data over the duration of the number of 4-week periods from question #1, as could be detected by linear regression, then differences between an observation and its preceding observation would, or average, be positive or negative, respectively. As another example, if a surgical group were less busy than average during one period (e.g., around Christmas) then the group may tend to be more or less busy than average during the next period. We tested for these potential trends or correlations by using the mean square successive difference test for each set of the optimal number of previous 4-week periods of total hours of elective cases.

The mean square successive difference test’s statistic (von Neumann’s ratio) assessed correlation between successive observations in each set of 12 total hours of elective cases during each period. If this ratio of the average squared difference between successive observations of total hours of elective cases was large relative to the sample variance of the observations, then the observations would not be randomly distributed about the mean. For each mean square successive difference test performed, N = 12. Each time series of 39 4-week periods provided 28 sets of 12 4-week periods. Because we had 17 time series, each a combination of surgical group and suite, the mean square successive difference test was performed 476 times (17 X 28). In that this test assumed that the observations of total hours of elective cases over 4-week periods were normally distributed, it had a high statistical power to detect correlation between successive differences. The assumption of normality was tested in question #3.

---

Fig. 2. Example of using 12 4-week periods to forecast total hours of elective cases for one of the surgical groups. (Panel 1): The raw data for total hours of elective cases per week for one surgical group from a 156-week interval are displayed. (Panel 2): From the data in panel 1, total hours of elective cases during 39 (156 ÷ 4) consecutive 4-week periods are calculated. (Panel 3): Forecasts are made by taking the mean of total hours of elective cases over 12 successive 4-week periods. (Panel 4): The forecasting errors or differences between actual future values and their corresponding forecasted values are given. (Panel 5): The percentage forecasting errors or absolute values of the forecasting errors expressed as percentages of the actual values are given.
Question #3: How Can Results from Questions #1 and #2 Be Used To Calculate an Upper Prediction Bound (Upper Limit) for the Total Hours of Elective Cases that Will Be Completed in a Future Period?

Once the optimal number of previous periods had been established, it was important to assess whether the differences between actual and forecasted total hours of elective cases (forecasting errors) were normally distributed over the optimal number of 4-week periods. If the forecasting errors were normally distributed, then the mean is the appropriate descriptor for the central tendency (typical number) of the total hours of elective cases over the optimal number of previous periods. If the forecasting errors are skewed, the median might be a better estimator of central tendency.

We evaluated whether each set of the optimal number of previous periods of total hours of elective cases had forecasting errors that were normally distributed. We used Lilliefors test to determine whether the forecasting errors were normally distributed for each set of the optimal number of previous periods. Using a P value < 0.05 to assess statistical significance, we would reject the hypothesis that forecasting errors over the optimal number of previous 4-week periods tend to be normally distributed if the test specified that more than 5% of the series had forecasting errors that were not normally distributed.

If the errors are normally distributed, then the Student t-distribution can be used to calculate upper prediction bounds (upper limits) for forecasts of total hours of elective cases. There is a specified probability that the actual total hours of cases during a forecasted period will be less than or equal to its upper prediction bound. Prediction bounds refer to the next forecasted period, whereas confidence bounds refer to the mean among many future forecasts. The 100(1 - α)% upper prediction bound equals

\[
\bar{X} + s \cdot t_{1-α} (N - 1) \cdot \sqrt{1 + 1/N}
\]

where \( \bar{X} \) is the mean of the observations of the total hours of elective cases over 4-week periods, \( s \) is the SD of the observations, \( N \) is the number of historical periods, and \( t_{1-α} (N - 1) \) is the 100(1 - α)th percentile of the t distribution with \( N - 1 \) degrees of freedom. For \( N = 12 \), the 60%, 80%, 90%, and 95% upper prediction bounds equal \( (\bar{X} + s \times 0.27) \), \( (\bar{X} + s \times 0.91) \), \( (\bar{X} + s \times 1.42) \), and \( (\bar{X} + s \times 1.87) \), respectively.

Results

Question 1

Differences were small in the three measures of the forecasting errors between choices of 9 previous 4-week periods and 20 previous 4-week periods (table 1). Based on the criterion given in Methods, the optimal number of 4-week periods of data to minimize error in forecasting total hours of elective cases was 12.
Question 2

The total hours of elective cases for each 4-week period during each series of 12 consecutive 4-week periods did not show statistically significant trends or correlations between successive periods ($P < 0.02$ for 1.0% of the series [5 of 476]).

Question 3

The total hours of elective cases for each 4-week period during each series of 12 consecutive 4-week periods were normally distributed ($P < 0.05$ for 3.4% of the series [16 of 476] and $P < 0.01$ for 0.8% of the series [4 of 476]). The upper prediction bounds (upper limits) for the total hours of elective cases that will be completed in a future period can be calculated using equations that assume that 12 successive observations of total hours of elective cases during 4-week periods are normally distributed.

Discussion

Budgeting Labor Costs

Block OR scheduling requires accurate forecasting of each surgical group’s future needs for OR time. A common approach to forecasting OR needs is to use the average of a surgical group’s recent total hours of elective cases. OR managers may do this informally, for example, by reviewing annual usage of OR time by the surgical group. By applying tools from management science, we showed that differences in forecasting errors among different number of previous 4-week periods are small between choices of 9 (36 weeks) and 20 (80 weeks) previous periods. The average of total hours of elective cases from 12 consecutive 4-week periods had an optimal error profile, wherein observations older than 48 weeks (12 periods × 4-week periods) did not influence the forecast. Nevertheless, surgical suites with information systems that record total hours of elective cases over six to eight previous periods may not experience important declines in forecasting accuracy (table 1).

Forecasting surgical groups’ needs for OR time is necessary to ensure there are adequate staff available either to complete the work in a manner that minimizes labor costs or during regularly scheduled OR hours (question #3). OR managers can handle the uncertainty in future OR workload by calculating the upper prediction bound (upper limit) for total hours of cases. We found that total hours of elective cases over 4-week periods can be normally distributed and unchanging over time. Consequently, an OR manager can use equations based on normal distributions to calculate the upper prediction bounds (upper limits) for total hours of elective cases. For example, if, during the past 12 4-week periods, a surgical group performed $X \pm s = 165 \pm 22$ total hours of elective cases, then there is a 90% chance the group will perform $< 165 + (22 \times 1.42) = 196$ h of elective cases during the next period. These data would then be used to budget OR staff. Similar statistical principles can be used to apply our results to determine how much OR block time to allocate to surgical groups so as to minimize future staff labor costs.

Importance of OR Workload Forecasting to Anesthesia Staffing

Under capitated contracts, anesthesia groups are paid a fixed monthly premium to provide all anesthesia services for a specified number of insured patients. Revenue is fixed. In addition, there are provider organizations that have salaried anesthesiologists or anesthetists. Under both capitated revenue or salaried settings, an anesthesia group minimizes staffing costs by having surgical groups allocated sufficient block time to complete their elective cases, but not too much extra time. Allocation of additional block time would increase “down” time, when salaried anesthesiologists and anesthetists are not caring for patients. A necessary step to achieve this goal is to use surgical services information systems to forecast surgical groups’ total hours of elective cases. To manage uncertainty in future workload, the anesthesia group can choose a percentage upper prediction bound (e.g., 50% vs. 90%) for the total hours of elective cases that will be completed in a future period. The percentage quantifies the likelihood of having sufficient OR time to complete the surgeons’ elective cases during the regularly scheduled OR hours of the next 4-week period. The higher the percentage, the greater the convenience provided to the surgeons, but the greater the cost of anesthesia staffing.

Rationale for Using a 4-week Period as the Length of Time for Forecasting Total Hours of Elective Cases

A major decision that the OR manager must make is the length of time over which to forecast total hours of elective cases. We used 4 weeks as the length of time for forecasting total hours of elective cases. We illustrate our rationale for this decision by considering a hypothetical surgical group.

A hypothetical surgical group performs an average of 12 h of elective cases each week. The surgical suite
allocates block time in multiples of 8 h each week. If the surgical group were assigned two 8-h blocks each week, the group would have too much unused OR time. In contrast, if the group were assigned one 8-h block each week, the group would not have enough OR time to complete their cases within their block time. The total hours of block time can be adjusted more precisely using blocks every 2 weeks rather than blocks every week. In addition, the surgical group would have difficulty completing their cases within their block time if the group were allocated one block per week and their one block per week coincided with a national holiday. For these reasons, both to reduce the impact of rounding forecasted total hours of cases up or down to the nearest multiple of the hours in a work day and to reduce the impact of holidays on block time allocations, many surgical suites allocate block time in increments of blocks every 2 weeks. In our example, the surgical group would be assigned three 8-h blocks every 2 weeks.

Using 2-week periods of block allocation, the shortest period over which total hours of elective cases should be forecasted is 4 weeks. This statement is true unless the OR manager requires that the day of the week that the surgeons have their block time be identical every week (e.g., every Monday). This requirement is not realistic for many surgical suites, because surgical groups may have time away for meetings or vacations on one of their block-time days. To complete their elective cases within their block time, the surgical group would need to have a different day of the week assigned to complete their cases.

We used 4 weeks as the length of time over which to forecast total hours of elective cases for two other reasons. First, surgeons' vacations and national meetings generally last 1–2 weeks. If many surgeons in a group are at a national meeting for a week (e.g., American College of Surgeons), the forecast for the period overlapping with the national meeting will be more accurate if the forecasting period is 4 weeks long instead of 2 weeks long. Second, the Fair Labor Standards Act specifies that time and a half is paid when hours worked exceed 40 h per week averaged over the pay period. At many hospitals, the pay period is 4 weeks. For the forecasted total hours of elective cases, and in particular the upper prediction bound (upper limit) for the total hours of elective cases, to be useful for budgeting staffing costs accurately, the forecasting period should be a multiple of the pay period.

Study Limitations

We focused on the method (time series analysis) for forecasting surgical groups' total hours of elective cases. Importantly, the accuracy of the OR workload forecasts (i.e., the forecasting errors) depends on the method used to determine on which day patients are scheduled for surgery.3,13 How patients are scheduled into OR block time varies among surgical suites.13 Patients' requests to be scheduled for surgery arrive randomly. Usually, a patient needing surgery is given the first available surgical date for which the surgeon has open block time.13 Patients then wait to have surgery from the day they were given a surgical date until their day of surgery. Increasing the mean length of this waiting period (e.g., from 1 to 3 weeks) decreases the week-to-week variability in the surgical group's workload. Surgical suites with longer average lengths of time that patients wait to have elective surgery are more likely to have accurate forecasts of workload. This is because surgical suites with longer average patient waiting periods will have more flexibility in finding the best date to schedule a case to match the expected case duration with the open block time that day.13 Block time will be more completely filled, resulting in total hours of elective cases each 4-week period being more nearly constant.

The accuracy of the workload forecasts may also depend on surgical groups' “time away” policies. For example, if a surgical suite's administrative structure encourages surgeons within a given group to coordinate time away from the surgical suite (e.g., only one member takes vacation at a time), then forecasting accuracy may be improved. Because our analyses were limited to data from the University of Iowa's tertiary surgical suite and Ambulatory Surgery Center, other sites with less variability in caseload may be able to use fewer than 12 4-week periods when forecasting total hours of elective cases. However, our recommendations on how to calculate upper prediction bounds (upper limits) for total hours of elective cases only apply to the use of 12 4-week periods.

Our study considered how to forecast surgical departments' or groups' total hours of elective cases, not an individual surgeon's future caseload.13 Although we did not evaluate scenarios relevant to forecasting total hours of elective cases for individual surgeons, we expect that our results would not apply. Forecasting elective caseload for individual surgeons may require that additional qualitative information be used. For example, if a new surgeon joins a group, the surgeon's future hours of elective cases would not equal zero. If a surgeon always
is on vacation for 4 weeks around Christmas and New Year's Day, then his or her total hours of elective cases during each period will not be normally distributed. If a surgeon progressively phases out his or her practice, there may be a trend over time. If surgeons in a group alternate being "on-service" (e.g., a burn or trauma team) or choose who takes consultations in a manner to achieve an equal workload, then each surgeon's total hours of elective cases during a period may not be normally distributed. OR managers may need a system to manually "override" statistical forecasts for individual surgeon's total hours of elective cases. Our analysis does not consider these issues.

Conclusions

Our study confirms the validity of a common practice of using the average of the most recent year's total hours of elective cases to forecast future usage of OR time for surgical groups. Importantly, we showed that 12 successive 4-week periods of total hours of elective cases may be used to predict the OR staffing required to complete surgical groups' elective cases during regularly scheduled OR hours or to allocate5,6 the appropriate hours of block time to minimize OR labor costs.

References

1. Macario A, Vitez TS, Dunn B, McDonald T: Where are the costs in perioperative care? Analysis of hospital costs and charges for inpatient surgical care. Anesthesiology 1995; 83:1138-44

2. Dexter F, Macario A: Applications of information systems to operating room scheduling. Anesthesiology 1996; 85:1232-4


