Improving Hospital Operations by Scheduling

Seung-Chul Kim
Ira Horowitz

Follow this and additional works at: https://aisel.aisnet.org/iceb2001

This material is brought to you by the International Conference on Electronic Business (ICEB) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICEB 2001 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Scheduling elective surgery needs to take several factors into consideration such as operating surgeon’s schedule, patient's schedule and the bed availability in ICU. Often, the operating surgeons schedule elective surgeries without close communication with the ICU administrator, and the consequence is the frequent cancellation of scheduled surgeries due to bed shortages in the ICU. This paper proposes a method to reduce the cancellation of elective surgeries by an improved communication between surgeons and ICU physicians through which the number of elective surgeries scheduled each day is made stable during the predetermined planning horizon.

INTRODUCTION

Langer [1] has written about the “illusion of control” wherein individuals believe that they can influence the outcomes of purely chance events. Hospital administrators are not infrequently prone to similar delusions in their belief that elective surgeries can be scheduled based solely on the availability of the operating surgeon and his or her supporting staff, the availability of an operating theatre, and how well the timing suits the patient.

Assuredly, there are many elective surgeries that can be confidently scheduled taking only these three factors into consideration. Breast reconstruction following a radical mastectomy, for example, can be scheduled months in advance, the only uncertainty being the number of days that the patient will have to remain in the hospital following the surgery, and there is rarely a shortage of beds for this purpose. Other elective surgeries, however, such as a hip replacement for an elderly patient, often require that the patient spend some time in the hospital’s intensive care unit (ICU) following the surgery, which introduces four distinct and uncontrollable elements of uncertainty.

First, the amount of time that an elective-surgery patient may have to spend in the unit prior to release is stochastic. Moreover, as is shown below, the stochastic process is quite complex and not nearly as straightforward as one is tempted to presume it to be because of the “elective” qualifier. Second, the typical ICU serves patients from a number of different sources, including, for example, the emergency room, so that any new demand for an ICU bed is by its very nature stochastic. Third, the amount of time that any such patient will have to spend in the ICU is both patient and therapy dependent and is also stochastic. Finally, ICU capacity as measured through the number of beds in the unit is typically a scarce hospital resource, because ICU care is an unusually expensive therapy, one that requires higher-than-normal nurse-to-patient and doctor-to-patient ratios, as well as specialized equipment that may exist only in the unit. Thus, in combination with the first three elements, the availability of an ICU bed becomes a matter of chance.

As a consequence, an ICU administrator is occasionally, and at random, forced to reject applicants, or to discharge ICU patients prematurely without providing the requisite amount of intensive care. The ICU patients, however, are only one of several constituencies to which the administrator is accountable. The families and physicians of the patients, and the operating surgeons and the ICU physicians are among the other constituencies and their preferences in regard to admissions and discharge priorities often conflict. One particularly prominent conflict is that between the operating surgeons and the ICU physicians. The basis for the conflict is that the surgeons must schedule elective surgeries and the operating theatre well in advance and assume that there will be an empty bed in the ICU for the individual upon whom they will be operating. In contrast, the admissions priorities of the ICU physicians are based upon the needs of all applicants. Like Hobson, the administrator is often forced to resolve the conflict through an unpleasant choice, and the path of least resistance is commonly to deny admission to an elective-surgery patient and, in effect, force the surgeon to cancel and reschedule the surgery. Such last-minute cancellations can have several negative consequences. In particular, they can wreak havoc with the schedules of the surgeons and the supporting staff, and require changes in the schedule of the operating theatre, since a cancelled surgery will eventually have to be rescheduled. Cancellations can also impose considerable stress on the patients and their families.

We focus on the possibility of using elective-surgery quotas in conjunction with a scheduling window to improve the scheduling of elective surgeries and hence enhance the operations of the ICU. We explore the efficacy of this scheduling medium through a simulation model that employs the patient demand and length-of-stay data from an actual ICU. We consider the establishment of elective-surgery quotas within the modus operandi of that ICU as well as in conjunction with a Flexible Bed Allocation scheme (FBA) [2] that reserves one or more beds for the exclusive use of elective-surgery patients. It is shown that the combination of a daily-quota schedule and FBA for elective surgeries can greatly reduce the number of cancelled surgeries with relatively few and minimal negative consequences for the hospital’s other patients.
We consider and build our model on a foundation of six month’s worth of daily data from a typical multi-disciplinary ICU that receives patients from four basic sources: Ward; Accidents and Emergency (A & E); Operating Theatre (OT)-emergency; and Operating Theatre (OT)-elective. The ICU physicians review each referral. Qualified referrals queue for beds that are allocated on a first-come, first-served basis. When there are empty beds and an empty queue, all qualifying patients are admitted immediately. If no bed is empty, the physicians check to determine whether any patient is sufficiently recovered as to permit an expedited discharge to a general ward. When a patient’s illness is so severe that admission to the unit cannot be delayed, an expedited discharge might become a virtual necessity. When no bed is empty and an expedited discharge is not feasible, the new referral must join or begin a queue. When, however, it is an OT-elective patient who is the candidate for admission, the operation is cancelled and rescheduled, which is the source of a very significant and common problem. As shown in Table 1, although it is quite rare for a referral from any source other than elective surgery to be denied admission to the ICU due to a full house, about a quarter of the elective surgeries were cancelled because of the lack of an available bed.

We use a computer-simulation model to capture these stochastic arrival and service processes. The model, which uses the SLAM II simulation language with user-written FORTRAN programs, is discussed in detail elsewhere [3].

Simulation models have a well-established history in helping to improve managerial decision making in ICUs [4, 5, 6]. As in any simulation, we had to first establish the statistical properties of the patient demand (arrivals) and length-of-stay (service times) processes. It will come as no surprise that the classic queuing-model assumptions of Poisson arrivals and exponential service times were the first alternatives to gain our attention [3].

In the cases of referrals from the Ward, A & E, and OT-emergency, a Chi-square test was unable to reject the hypotheses of Poisson-distributed arrivals and exponential service times. Thus, we assume these theoretically convenient arrival and service patterns throughout our analysis. In the case of OT-elective patients, however, we were initially forced to reject both hypotheses. Indeed, OT-elective arrivals have several characteristics that distinguish them from the other ICU applicants.

First, the timing of elective surgeries is deterministic in advance. A colonoscopy that results in the removal of polyps might be scheduled a week in advance, whereas breast reconstruction surgery might be scheduled months in advance. The surgeons choose an operation time and date in light of their own schedules, the patient’s condition, and what is known at the time about bed availability in the ICU. Second, the medical histories of the patients and the causes of their illnesses are well known. This simplifies the review-and-admission decision. Third, the medical conditions of these patients are ordinarily less severe, and normally require less intensive monitoring during the postoperative recovery period, than those of the other referrals with whom they may be competing for a bed. Most of these latter referrals require a full and active intensive care. Fourth, the average OT-elective patient requires less ICU time than does the average patient from any source other than A & E; more than 80 percent of OT-elective patients are discharged within 48 hours. Fifth, OT-elective surgeries are seldom scheduled for Saturdays or Sundays. Indeed, after eliminating weekends we no longer reject the hypothesis of Poisson-distributed OT-elective arrivals. Therefore we incorporate the Poisson-arrival assumption into the model and we only generate OT-elective arrivals on weekdays.

Last but not least, the mortality rates of OT-elective patients are one fourth of the rate of the others, and their paths to recovery also have unique characteristics that fall into three classes. Specifically, about 65 percent of these patients spend at most one night in the unit, another 20 percent spend between one and two nights in the unit, and the rest have quite open-ended lengths of stay. After eliminating the few outliers in our data, the “open end” is capped at fifteen days.

Theoretical considerations and a visual inspection of the data displayed in Figure 1 led us to dismiss the possibility that the random service times for OT-elective patients could be adequately modeled by, for example, a normal density, a lognormal density, or a beta density. The same considerations, however, did not apply to the Gamma density. We therefore hypothesized the service times in each of the three classes to be generated by an individual Gamma density tailored to that class [2]. The parameters of that density are determined from the empirical mean and variance for those patients. In no case was the resulting density Gamma-1, which is an exponential density. We then tested the hypothesis that the overall distribution of the service times can be modeled as a weighted average of the three densities. The weights are given by the percentages noted in the previous paragraph. A Chi-square test detailed in Table 2 fails to reject that hypothesis ($\alpha = 0.05$). Hence we introduce the tripartite Gamma density into our model and simulations.

The uniqueness of the OT-elective arrival and ICU service patterns, and the fact that last-minute cancellations of those surgeries is a very serious problem, encourages us to attempt to exploit the further fact that elective surgeries can be and always are scheduled well in advance. It is usually the case that the surgeons do the scheduling, subject only to the availability of the operating theatre and the current status of the ICU. The latter, however, will not necessarily reflect the status of the unit on the day for which the operation is scheduled, and therein lies the rub. What we propose here is to smooth the flow of elective surgeries by pooling all of the scheduled surgeries for some fixed time period, and then systematically allocating them over that period. We make that proposal because that flow can be controlled, and because the simplest means of doing so is to impose a quota
system that reduces the variance in the number of elective-surgery patients seeking admission to the unit on any given day. We shall refer to this reduction as the pooling effect.

THE QUOTA SYSTEM

Our data come from a 14-bed unit. In order to model a normal condition for the operations of the unit, our data analysis excludes patients whose ICU residency exceeded 15 days. Those patients comprised 5.89 percent of the total number of patients. To compensate for the exclusion of these data, we limit the unit in the simulations to only 13 beds, which reduces the available capacity by 7.14 percent. In each of the simulations that we run, the OT-elective arrivals are generated solely for Monday through Friday. To further emulate the actual situation, when an empty bed is not available in the ICU, a simulated OT-elective arrival does not join the queue. Instead, the arrival is recorded as a cancelled surgery.

Over the six-month period during which we gathered our data the highest bed utilization rate in the ICU was 83.22 percent, and the average rate was only 73.77 percent, which strongly implies that the administrator's problem is one of better managing the existing capacity as opposed to acquiring the funds to add to that capacity. A more efficacious scheduling of elective surgeries has the potential to help solve that problem.

Scheduling rules have the potential to reduce customers waiting times in a variety of operating environments [7] including surgical procedures in hospitals [8] and hospital outpatients departments [9]. Surgical demand scheduling has been a subject of particular interest [10]. That interest recently manifested itself in two studies devoted specifically to elective surgeries. Gerchak et al. [11] used stochastic dynamic programming to characterize the nature of the optimal reservation policy. Most closely related to our work, Ridge et al. [5] used a simulation model to explore the effects on an ICU of a rudimentary deferral rule for elective-surgery patients.

Our unique contribution to this growing literature is to propose the introduction of an elective-surgery quota system that supplements our previously proposed bed-reservation scheme for elective-surgery patients [2]. The first step in the process is to establish a scheduling window that will allow the simultaneous consideration of all of the demands coming from elective surgery for some fixed future period of time, and permit the allocation of those demands by day of the week. The introduction of such a planning horizon permits feedback between the ICU administrator, the surgeon, and the patient. In the interest of simplicity it would seem to make good sense to consider weekly time frames, and here we conduct our simulation experiments with one-week and two-week time frames.

Second, we need to establish a specific form of quota system, the simplest of which fixes a daily quota for Monday-through-Friday surgeries. It might, however, also make good sense to have variable daily quotas that seek to exploit the fact that elective surgeries are not scheduled over the weekend. In this regard, we evaluate one option that on the surface would seem to hold particular promise. That promise, it will be seen, is ultimately realized below the surface as well.

Suppose, for example, that we institute a one-week window and a one-surgery daily quota. The simulation procedure first generates one week of elective-surgeries data at a time for a Poisson distribution whose parameters have been determined from the empirical data. If five or fewer surgeries are generated, then one surgery is scheduled for each of the five weekdays successively, starting with Monday, continuing on to Tuesday, and so forth until all surgeries have been scheduled. If a bed is unavailable on the day of the scheduled surgery, the surgery is cancelled. If, however, seven elective surgeries are generated, then after filling the one-per-day quota with five surgeries, the remaining two surgeries are allocated to Monday and Tuesday, thereby giving two demands for ICU beds from elective-surgery patients on those days. If two beds are available on those days, the surgeries take place. Otherwise, one or more of the surgeries must be cancelled. Thus our quotas are not inviolable in that there is not a weekly quota. Rather, the quotas provide guidelines that the administrator will attempt to respect.

Fixed quotas

We first consider a fixed-quota system that imposes the same potentially violable quota for elective surgeries on weekdays. The system's operations are simulated for a twenty one-year window with one-surgery, two-surgery, and three-surgery quotas, and for one-week and two-week scheduling windows. The first year is cut off in order to eliminate the effect of the initial state of the simulation, and data is collected for the remaining twenty years. Each of the twenty years comprises $24 \times 365 = 8,760$ simulated hours, thus resulting in a total of $17,520$ hours of simulated data for each of the $3 \times 2 = 6$ initial experiments. The system's performance is then evaluated on a number of dimensions that are ordinarily of interest to the administrator, the physicians, and the surgeons. These dimensions include the average time that a patient spends in the queue and in the system, the number of cancelled surgeries, and the number of patients treated per year. By way of illustration, the average results for the simulations under a two-week scheduling window are presented in Table 3. Thus, for example, a two-week window and a one-surgery-per-day quota achieve a 23 percent reduction in the number of cancelled surgeries compared with the current system. The latter reduction is attained while slightly reducing the other patients’ average waiting times. The quota system improves the queuing measures for the other patients by reducing the demand fluctuations from OT-elective patients.

The win-win result of a one-surgery quota and a two-week scheduling window, offers a compelling argument for establishing some elective-surgery quota system. What
remains to be determined is whether further improvement can be achieved by coupling a quota system with an FBA scheme [2].

The ICU administrator faces a multiple-objective decision problem wherein there are tradeoffs to be made and evaluated between cancelled surgeries and waiting times. The noxious reality is that ICU patients often do have to wait for admission to the unit. Although we cannot measure accurately the clinical impact of additional waiting time on patients, it would not be a totally new experience for them, but rather an extension of what they currently endure. This choice among the tradeoffs is the administrator’s call and the administrator’s problem. As shown in [2], however, some form of FBA reservation scheme for elective-surgery patients may provide a better resolution to that problem. Under FBA the number of cancelled surgeries is reduced, but at the expense of longer average queuing times for the other patients. We now consider the impact of quotas imposed in conjunction with FBA. The results for a one-bed FBA policy and a two-week scheduling window quota system are summarized in Table 4.

Tables 5 and 6 summarize the effects on the number of cancelled surgeries and the overall waiting times, respectively, for both the one-week and the two-week scheduling windows. In addition, these tables provide p-values for one-tailed tests of the statistical significance of the average simulation results in pair-wise comparisons with the status quo. It is seen that in the overwhelming majority of cases, the differences are significant at virtually any reasonable level of statistical significance. Figures 2 and 3 present the cancelled-surgery data graphically, and Figures 4 and 5 display the overall waiting-time data graphically.

The simulation results reveal that a three-bed FBA scheme combined with a one-surgery-per-day quota achieves the greatest reduction in cancelled surgeries over the current system, with either a one-week or a two-week scheduling window. Table 5 shows that with a one-week window the reduction is almost 100(19.00 - 8.60)/19.00 = 55 percent. With a two-week window the reduction is almost 60 percent. Figure 4 shows that with a one-week window, the one-surgery quota also induces the smallest increase in waiting times that the other patients incur as a result of introducing a three-bed FBA scheme. It is seen from Figure 5 that with a two-week window the one-surgery and two-surgery quotas have essentially the same impact on the average waiting times.

Under either a one-week or a two-week scheduling window, the number of cancelled surgeries is the lowest with the one-surgery-per-day quota. The number of enforced cancellations increases as the quota increases from one to three, with a three-surgery quota system being particularly unproductive. A three-surgery quota system cannot improve the system's performance beyond the levels achieved with smaller quotas. This is so, because the average OT-elective arrival rate of 1.515 per day (7.575 per week) is not sufficiently large as to yield any pooling effect that can take advantage of a daily quota that exceeds two.

Figure 6 sketches the efficient frontiers that summarize the available tradeoffs between average waiting times and cancelled surgeries with some form of FBA scheme, with one-week and two-week scheduling windows, and with a one-surgery-per-day quota system. Our intention here, then, is not to suggest an optimal policy to an ICU administrator. Rather, the intent is to demonstrate how the administrator can be provided with the information upon which to best make an informed decision as to whether to impose a particular FBA scheme in conjunction with an elective-surgery quota system.

The efficient frontier presents our results in a readily understood visual format. The ICU should operate with a system corresponding to a point on the lowest frontier, that which lies furthest to the southwest and that derives from a two-week scheduling window. This is true regardless of which system is used in conjunction with the scheduling window. Each point on that particular frontier is associated with the smallest average overall waiting time for any given number of cancelled surgeries. The intermediate points between any two points that are derived from our experiments can always be achieved by selecting one of these pure strategies: percent of the time, and the other (1 - x) percent of the time. Assuredly, however, there might be other pure strategies with which we have not experimented that would perform even more efficiently than the mixed strategies. It is management's responsibility to choose the point that it prefers on the lowest frontier. That point will be determined once management has expressed its acceptable tradeoffs through the jagged indifference curves in the figure. The optimum is the point of tangency between the lowest frontier and lowest indifference curve.

**Variable quotas**

Elective surgeries are not scheduled on weekends. It thus seems reasonable to expect more ICU beds to be available on Mondays than otherwise, with or without FBA. Figure 7 depicts the patient arrival data for all four sources by day of the week. Nothing in the data suggests that a greater number of arrivals from the other sources will make up for the deficiency in weekend OT-elective arrivals. The question thus arises as to whether one might not seek to exploit that expectation by relaxing the one-surgery quota for Monday or possibly some other day, and if this is done perhaps imposing a stiffer quota on some other day.

In particular, since on average a little more than seven elective surgeries per week were scheduled during our sample period, it would seem a priori reasonable for a variable-quota scheme to permit seven scheduled surgeries per week. Further, no such surgeries are scheduled for weekends, and more than 80 percent of elective-surgery patients are discharged from the ICU within 48 hours. The most plausible variable-quota candidate would therefore seem to be that which permits two scheduled surgeries on both Monday and Friday, with the one-per-day quota maintained for the other three midweek days. Thus we
explore the performance of the simulated system under the latter (2, 1, 1, 1, 2) scheme, with and without FBA policies, and for one-week and two-week scheduling windows. Table 7 summarizes these results in apposition to those for the fixed-quota system. Figure 8 shows the efficient frontiers that emerges from these explorations.

As with the fixed-quota system, there is a variable-quota system that, when used in conjunction with a three-bed FBA scheme, is more effective in reducing cancelled surgeries than is a stand-alone quota system, or one that is used with either a one-bed or a two-bed FBA scheme. Or, at least, that is what our simulation results imply. Specifically, with a two-week scheduling window the variable-quota system reduced the number of cancelled surgeries by 100(17.25-6.15)/17.25 = 64 percent from the current system. This is modestly bigger than the 60 percent reduction achieved through the best-performing fixed-quota system. Still further, Table 7 shows that when compared to the three-bed FBA alternatives this (2, 1, 1, 1, 2) variable-quota system accomplishes that reduction with the smallest increase in waiting time (1.37 - 0.30 = 1.07) for the other patients, relative to the current system with a quota. The p-values from the statistical tests show that the differences are significant between the quota-only systems and those with the FBA schemes for most cases. This indicates that an FBA scheme will have significant impacts on the performance of a quota system in achieving further reduction in the number of cancelled surgeries, but the waiting times will also increase due to the reservation of some beds. Although not specifically detailed in Table 7, statistically significant differences also exist between, for example, a one-bed FBA scheme and a three-bed FBA scheme, regardless of the quota system and the scheduling window. Moreover, it may easily be inferred from the table that once the optimal FBA scheme has been established, the differences among the various quota systems are not significant for the waiting times, but they impact on cancelled surgeries. Particularly profound impact is shown on a three-bed FBA scheme in that the number of cancelled surgeries is affected by the quota systems and the scheduling window whereas the waiting times change little.

The superiority of the variable-quota system under a two-week scheduling window to that system under a one-week scheduling window might have been anticipated. Thanks to the pooling effect, a two-week arrival period is more likely to produce a distribution of elective-surgeries that can be mapped into the (2, 1, 1, 1, 2) variable-quota allocation than is a sequence of two one-week arrival periods. As is seen by comparing the first and third blocks of Table 7, with a one-week scheduling period and FBA 2, the (2, 1, 1, 1, 2) variable-quota scheme produces the exact same results as those of the fixed-quota scheme. With FBA 1 and FBA 3 there are tradeoffs. Comparing the second and fourth blocks of the table reveals the clear if modest dominance of the variable-quota system with a two-week scheduling window. For example, with a one-bed FBA scheme, the fixed quota results in 11.95 cancelled surgeries and an average waiting time of 0.44 hours, whereas the variable quota results in 11.80 cancelled surgeries and an average waiting time of 0.42 hours. These results might be explained as follows. On a week when there are seven elective surgeries, the one-per-day quota would assign the sixth and seventh surgeries to Monday and Tuesday, whereas the (2, 1, 1, 1, 2) system would assign the sixth and seventh surgeries to Monday and Friday. Since no elective surgeries are scheduled over the weekend, scheduling two elective surgeries on Friday reduces the probability of having an empty reserved bed on a Saturday or Sunday. But with most elective-surgery patients spending less than two days in the ICU, there is also likely to be two empty reserved beds on Monday, ceteris paribus, even when two surgeries have taken place on Friday.

In sum, with a one-week scheduling window whether the fixed-quota system or the variable-quota system outperforms the other depends upon the number of reserved beds. With a two-week scheduling window the variable-quota system, however, dominates the fixed-quota system in all instances. Moreover, the variable-quota system becomes more effective with a longer scheduling window, because the longer scheduling window permits additional pooling that allows the administrator to more readily produce the variable-quota pattern.

**CONCLUSION**

Modeling has a time-honored past in contributing to the improved management of health services in general [12, 13]. This paper seeks to add to that tradition by its unique proposal of a coordinating mechanism that effects the partial integration of the operations of two administratively independent hospital facilities, the Surgery and the ICU. That mechanism, a quota system for scheduling elective-surgery patients for ICU admissions, is evaluated via a computer simulation that reflects the actual operations of a very typical ICU and whose parameters are determined from the unit's historical data.

The last-minute cancellation of an elective surgery can be a traumatic event for the patient and the patient’s family. The cancellation also means a waste of both a surgeon’s valuable time and that of the supporting cast. Still further, a cancelled surgery results in the enforced idleness of one of the hospital's most valuable resources, the operating theatre in which it was scheduled to occur. Because the demands made on a hospital's services are random demands, and because the services are provided with random service times, it is not altogether possible to avoid cancelled surgeries and their negative consequences. It is, however, incumbent upon the hospital's administration to attempt to reduce those cancellations and to lessen their negative consequences. To do so commonly involves tradeoffs between improving the service provided to elective-surgery patients and imposing lengthier waiting times on the other patients seeking entry to the ICU. That is, the ICU administrator faces a multiple-objective decision problem with conflicting objectives. Nonetheless, as we have shown here, one can take advantage of the fact that elective surgeries are scheduled in advance by introducing an elective-surgery...
quota system into the ICU admissions process. In particular, we have shown that by using a quota system in conjunction with a flexible bed allocation scheme it is possible to achieve significant reductions in cancelled surgeries. Moreover, this might even be accomplished without impacting adversely on the average waiting times endured by the other patients seeking admission to the ICU.

From a traditional managerial perspective, scheduling elective surgeries through a quota system effects the formal vertical integration of the upstream patient sources and the downstream ICU server. The four departments that the unit serves constitute the upstream processes to the downstream ICU server. In order to improve the performance of an ICU effectively and efficiently, it is necessary to alter the processes as well as the server. We have shown the beneficial effects of formally linking the one controllable upstream process, the scheduling of elective surgeries, to the downstream ICU admission process. We are confident that similar benefits can be enjoyed through such a linkage in any ICU that operates along the lines modeled here, as so ICU units throughout the world.

There are a number of paths for future research. It would be worthwhile to investigate the possibility of full integration to include all the four upstream departments in the ICU admission process by, for example, developing a priority system for different groups of patients. Furthermore, the admission process needs to be linked to the discharge process by establishing the criteria for expedited discharge or regular discharge. It will also be useful to develop a prediction mechanism of calculating a patient's expected time to stay in the ICU when he/she is admitted. This will enable the management to determine the expected date of vacancy for beds, and further improve the allocation of surgery quotas in the variable-quota system.

REFERENCES


FIGURES AND TABLES

Available upon requestt