Communicating Accurate Case Start Times for Surgeries

Robert Allen
Clemson University
Clemson, South Carolina

Kevin M. Taaffe, Ph.D.
Clemson University
Clemson, South Carolina

Abstract

Operating room (OR) rescheduling is the process of adjusting the surgery schedule when the current schedule is subjected to disruptions on the day of surgery. The decision to make a schedule adjustment will impact patient safety, patient satisfaction, hospital costs, as well as surgeon satisfaction. Of particular importance is when, and how frequently, to update the scheduling and tracking systems. These questions and their impact on maintaining schedule accuracy and minimizing room overtime are explored. Discrete event simulation was used to simulate surgical cases in the OR and to test different “right shifting” and case updating policies for their effectiveness. Results indicate that there is a window of delay within which an update should be made; otherwise staff satisfaction or schedule accuracy will suffer.

Keywords
Healthcare, rescheduling, discrete event simulation, day of surgery, right-shift

1. Introduction

During the course of a day in a hospital, surgery schedules set at the beginning of the day may undergo disruptions. These disruptions can include the addition of add-on elective, urgent or emergent cases, case cancellations, and deviations from scheduled case duration. This paper focuses on the latter. Deviations from scheduled case durations can be caused by unpredictable complications during surgery, patient health issues before surgery, surgeon availability and many other reasons. These deviations create a need for nurses and core coordinators to reschedule cases during the course of the day. In this paper we focus on the impact of right-shift rescheduling cases when disruptions occur. Right-shifting is the process of delaying cases by visually moving them to the right on the posted schedule. The decision that core coordinators will need to make are those related to when and by how much time to delay the rescheduled cases.

Currently the hospital will rarely reschedule a room because of case delay. They will, however, make updates to surgeries that have moved to another room. The information on the tracking boards is realistically only used to convey start of day schedules. Any changes made to the schedule are communicated through phone and pager to the involved staff (e.g. surgeon, pre-op coordinators, and nurses). If the tracking board could be kept more routinely and automatically updated, the loss of work due to unnecessary communication could be minimized.

These decisions on the day of surgery can impact OR utilization, equipment availability, surgeon availability, surgeon satisfaction, staffing levels, patient satisfaction, and costs (patient and hospital). OR managers are required to make rescheduling decisions every day. Interviews with Greenville Memorial Hospital (GMH) staff in Greenville, SC have confirmed that communication in the OR and across all activities within perioperative services are important to making good decisions. Hence, the motivation for this research comes primarily from research questions posed by perioperative management at GMH. They ask under what conditions they should update the tracking boards in the OR with a new schedule. There is a fine balance between providing useful information about case status and overwhelming the staff with too much information.

2. Literature Review

The timely incorporation of surgical equipment, hospital staff, surgeon groups, ORs, preop rooms, post anesthesia care unit rooms and the patients are all important. The scheduling of all of these interconnected parts creates a complex problem and falls under the general research area of OR scheduling. Although extensive research has been
performed under the umbrella of OR scheduling, there is comparatively few research papers and journal articles regarding decision making on the day of surgery.

Cardoen et al. [1] provide an extensive literature review and survey of OR planning and scheduling. A base knowledge of OR scheduling would include block scheduling, elective case scheduling, case duration estimation, and surgery capacity planning. Block scheduling is a scheduling system in which OR managers schedule “block time” to ORs, which belong to a specific surgical group or surgeon. A paper by Fei et al. [2] also discusses block scheduling as opposed to an open scheduling strategy.

Elective case scheduling is the act of scheduling an elective case. By definition, elective cases are those that are scheduled ahead of time and are not urgent or emergent in nature. Elective cases are usually scheduled in days or weeks in advance of the surgery, but it is possible for cases to be rescheduled to take place with very little notice (especially for inpatient cases) if the schedule permits. This is also known as scheduling of an add-on case. Add-on cases are added on to the day’s schedule in addition to the schedule posted at the beginning of the day [3]. In addition to elective add-on cases there are also non-elective add-on cases. Non-elective add-on cases are urgent or emergent cases. The scheduling of non-elective add-on cases would have higher priority than other elective cases already scheduled. The cancellation or postponement of an elective case is common when procuring OR time to accommodate non-elective case add-on surgeries [4]. Li et al. [5] show that we can improve OR schedules by rescheduling ORs before the day of surgery to minimize the staffing costs.

OR scheduling techniques include a wide range of solution methodologies. Such methodologies predominately involve mathematical programming techniques or simulation [1]. Such mathematical programming techniques include mixed integer programming and column generation to solve these problems [2, 6]. For example, the objective may be to minimize the total staff in a hospital OR suite where the variables may be levels of experience among staff [6]. In a paper by Fei et al. [2], they use mixed integer programming to minimize the amount of idle time between surgeries. Branch-and-price strategies have been used to minimize the number of staff needed over a particular work-day to meet coverage constraints [7]. There are many articles that discuss linear programming or threshold-based statistics as a means to schedule OR cases [8, 9].

Simulation has also been a widely used tool in OR scheduling. Discrete event simulations have been used in many incidents to analyze management policies, determine OR schedules, increase OR utilization and many others [10-12]. Monte Carlo simulations based on Markov Decision Processes and discrete event simulations have also been used to generate policies for accommodating elective and non-elective surgeries [4].

To thoroughly review methodologies in OR scheduling and rescheduling, it is important to look past methodologies proven in healthcare and also look at other industries such as manufacturing or project management. The stochastic job shop scheduling problem in manufacturing is similar in nature to the surgery scheduling problem in healthcare. They are similar in that jobs are like the surgeries to be scheduled and that both problems have stochastic processing times. In a paper by Jones and Rabelo [13], solution methodologies to the job shop scheduling problem include mixed-integer programming, dynamic programming, branch and bound, lagrangian relaxation, discrete event simulation, neural networks, and a variety of meta-heuristics. Vieira et al. [14] provide a framework of strategies, policies and methods for rescheduling manufacturing systems. They discuss practical rescheduling methods that include right-shift rescheduling, partial rescheduling, and complete regeneration. The similarities between these two problems will prompt further research into which solution methodologies may be able to bridge the gap to the healthcare scheduling problem.

With regards to OR rescheduling, there is comparatively little literature amongst the general research in OR scheduling. One paper addresses the human factors element on the day of surgery, which can include how visual presentation of the OR status can affect decision making [15]. The literature on day of surgery case scheduling focuses mostly on how to optimally accommodate add-on cases, whether they are elective or non-elective [4, 6], while some literature focuses on case duration as a driver for the scheduling process [16,17]. There is some literature that discusses decision making on the day of surgery [15]. Dexter et al. [16] has also provided advantages and disadvantages to moving a case to another room at the end of the day to minimize OR over-utilization costs. Van Essen et al. [18] discuss rescheduling of ORs due to case delay and addition of emergency surgeries. They employ an integer linear program to minimize the deviation from the preference of the stakeholders (e.g. surgeon, hospital, patient). Van Essen et al. [18] found that the “preferences mainly lead to shifting a surgery and scheduling a break between two surgeries.”
Examples of using any type of quantitative method (math model, simulation, or fixed policy based on model findings) in real time to replace gut-feeling decisions, judgment calls, and experience-based decisions on the day of surgery are non-existent. Quantitative methods being used in real time have been discussed briefly in a few instances of the literature, but have not been implemented nor documented [19]. There appears to be an opportunity to investigate the value of using such methods for making decisions on the day of surgery.

3. Methodology
In this section, we present the OR rescheduling problem as well as introduce the simulation model used to generate new schedules throughout the day. We will use the model to test varying parameters and to create generalized policies. We use 30 simulation replications to obtain the results that are discussed in section 4.

3.1 Data Input
The input for our simulation model was developed from 18 months of case data over 28 ORs from GMH and was used to create distributions for scheduled and actual case durations. Empirical distributions were created from the data for estimated durations of cases and schedule gaps. Actual durations of milestones were fit to a mixture of Beta, Erlang, and Lognormal distributions depending on the lowest square error. Case start time data was fit to a normal distribution because of the possibility of starting early (negative sample). Although the reasons for case delays are known by the coordinating staff, they are difficult to see in case data collected at the hospital. Therefore, we can only make decisions on the times certain events occur during the course of the day. Since we are not considering add-on cases (e.g. urgent and emergent cases), our model data was taken from historical data on elective outpatient cases only. Case delay is determined by the difference in estimated case duration and actual case duration (sum of individual milestones). We are mimicking how the hospital handles case delays by simply using offsets, but there may be a better way estimate how case delay affects remaining case time. In our analysis, we are assuming that case duration is not affected by rescheduling.

3.2 Model development
This simulation model has been developed to study the rescheduling problem for one OR. ARENA modeling software was used for input analysis as well as for generating the random schedules. A random schedule is generated by sampling case times and scheduled gaps from probability distributions. Once the day has a full schedule, the cases are then played out according to distributions created from the historical data. The actual case durations have been divided up into 5 parts and are sampled from individual distributions for each part. The five parts are the times between each of the following milestones: begin setup, patient in room, procedure start, procedure finish, patient out of room, and clean up end. As the day progresses, cases begin to get ahead or fall behind. If a case has not yet started and is running late by more than the allowable amount, a reschedule event is triggered. During this reschedule event, the remaining cases in the room are adjusted by the reschedule amount. Then the process continues and the rest of the day is played out. It is possible that multiple reschedule events will happen on the same day.

During the reschedule event, the current schedule is right-shifted or delayed by the reschedule amount. After a case has been delayed by a certain number of minutes, the amount that the schedule is right-shifted is dependent on whether or not there is a gap in the schedule between the delayed case and the potentially affected case. As can be seen in Figure 1, if there is no gap between the cases then the next case is rescheduled by the reschedule amount. At 9:30am, case 1 was scheduled to complete (as well as case 2 was scheduled to begin). However, case 1 has not finished by 9:30am, which means case 2 will have to wait as long as case 1 is delayed (in this case, 30 minutes). 10am is when the reschedule event occurs, during which only the end time of case 2 is rescheduled (right-shifted by 30 minutes). Case 3 and 4 are not rescheduled because the gap between cases 2 and 3 is sufficient enough to absorb the disruption. In this case, there is one reschedule event and one rescheduled surgery.

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>8am</th>
<th>9am</th>
<th>10am</th>
<th>11am</th>
<th>12pm</th>
<th>1pm</th>
<th>2pm</th>
<th>3pm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Room Schedule</td>
<td>Case 1</td>
<td>Case 2</td>
<td></td>
<td>Case 3</td>
<td></td>
<td>Case 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Room Reschedule</td>
<td>Case 1</td>
<td>Case 2</td>
<td>Case 3</td>
<td></td>
<td>Case 4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Example Scenario 1
However, if there is a scheduled gap between cases and the first case is delayed. The gap will be used up before delaying the start of the next case. As can be seen in Figure 2, Case 3’s duration was one hour longer than scheduled. The 30 minute gap is used up and case 4 was only delayed a total of 30 minutes. Case 4 was delayed by 30 minutes; hence the case will be right-shifted during the reschedule event by 30 minutes. The gap between cases 3 and 4 was not sufficient enough to absorb the disruption.

Figure 2: Example Scenario 2

<table>
<thead>
<tr>
<th>Scenario 2</th>
<th>8am</th>
<th>9am</th>
<th>10am</th>
<th>11am</th>
<th>12pm</th>
<th>1pm</th>
<th>2pm</th>
<th>3pm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Room Schedule</td>
<td>Case 1</td>
<td>Case 2</td>
<td>Case 3</td>
<td>Case 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Room Reschedule</td>
<td>Case 1</td>
<td>Case 2</td>
<td>Case 3</td>
<td>Case 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3 Model Output
During the course of a simulation run, the model records important information as the room’s cases are played out. First, we look at the average start time offsets and average end time offsets. Start time offsets refer to the difference between the actual start time and scheduled start time of an individual case, where end time offsets refer to the difference between the actual end time and scheduled end time. Second, we look at the number of reschedule events and the number of surgeries affected by rescheduling. This is important because there is a tradeoff between communicating accurate case start and end times and overwhelming the users of the hospital tracking boards with information. Last, we consider accuracy of shift end predictions. It is important to accurately schedule and manage nurses in perioperative services, and having accurate case end times (and, thus, the room completion time) is an advantage for the core coordinators.

3.4 Parameters
We generalize the historical case data by looking at two general case lengths or scenarios. The two scenarios were considered because they both represent routine case schedules with elective outpatient surgeries scheduled in the room. Two scenarios were considered for this paper: 1) The room schedules were generated with cases between 60 and 90 minutes in length; 2) The room schedules were generated with cases between 90 and 120 minutes in length. These two scenarios are used to show the effect of the policies on rooms with differing average case durations.

In addition to the two scenarios proposed above, we are proposing many different experiments. The experiments conducted explore the effects of changing two parameters. The first parameter is the reschedule criterion which denotes the amount of time that a case must be running late in order for a reschedule event to occur. The second parameter is the reschedule amount which denotes the amount that each remaining case must be right-shifted. Please note that while it is common that the criterion amount and reschedule amount are traditionally the same value, they do not have to be. Therefore, we also explore combinations in which the reschedule amount is less than the criterion amount. For example, if a case is running an hour behind, and a reschedule event is triggered, we could only reschedule the remaining cases in the room for 30 minutes later instead of the normal 1 hour. If we reschedule a particular case for an hour later, it is possible that the leading or prior case could actually finish early, allowing the rescheduled case to actually begin at its originally scheduled time. Situations like this are what motivate the experimentation we propose.

Table 1 shows the different combinations of parameter values for each scenario. The criterion amount can be one of four values. It does not make sense to include a reschedule amount greater than the criterion amount, so those combinations of parameter values were not tested.
4. Results
Based on the many different possible combinations of criterion and reschedule amounts, we looked for trends among the results. First we looked at the most simple scenario in which the criterion amount was equal to the reschedule amount. On one hand, we found that the number of reschedules decreases as we increase the values of each parameter. But on the other hand, the start time offsets and end time offsets also increase. This was the same for both scenarios (60-90 minute cases and 90-120 minute cases). The results for this experiment can be seen in Figures 3 and 4 below. In both scenarios, the number of reschedule events becomes very large as you decrease the parameter values, but the start and end time offsets only get marginally smaller.

![Scenario 1 - Equal Parameter Values](image)

**Figure 3: Scenario 1 with equal parameter values**
In conclusion, depending on the cost of the reschedule event and the cost of not starting a case on time, the optimal parameter values can be found somewhere on this curve. Generally speaking, managers at GMH believe that the reschedule event will have a very low cost since the tracking boards can be updated automatically when a reschedule event occurs. The more subjective cost to a reschedule event is the cost to the stakeholders of the information. The more often we reschedule cases the more burdening it is for the surgeons and nurses to keep up-to-date with the latest information.

Scenario 1 and Scenario 2 were compared using paired t-tests for each parameter value. The differences in start and end time offsets were statistically significant with a 95% confidence interval. The number of reschedule events and rescheduled cases however were not statistically significant across all parameter values shown. The likelihood of the resulting difference being statistically significant decreased as we increased the value of the parameters. This is likely due to reaching a lower bound of zero for the number of rescheduled events and rescheduled cases.

The second experiment that we considered was the measuring of the effect of differing parameter values on start and end time offsets, the number of reschedule events and the number of surgeries rescheduled. Our findings during this experiment were three fold.

1) As we increase the criterion amount the number of reschedule events and number of rescheduled cases decrease in both scenarios.

2) If we keep the reschedule amount the same and increase the criterion amount ((15,15), (30,15), (45,15), and (60,15)), we find that the start and end time offsets increase. This is similar to the first results we showed with the parameter values being equal.

3) In every case as the reschedule amount increases to become equal with the criterion amount (e.g. (60,15), (60,30), (60,45), (60,60)), the start and end time offsets also decrease.

Therefore, it is best to have the criterion amount and reschedule amount the same. Results can be found below in Figures 5 and 6. These results were also compared using paired t-tests. The results of the paired t-tests were comparable to those shown for scenarios with equal parameter values. Differences in start and end time offsets were always significant, while reschedule events and rescheduled cases were less likely to be significant. Differences were less likely to be significant as we increased the parameter values.
Last we consider end of day prediction accuracies. We first looked at keeping the criterion amount and reschedule amounts equal. We confirmed our intuition that the prediction accuracy increases as we wait till later in the day to make our prediction. Figure 7 shows that the predictions are less variable as the time of day the reschedule events occur becomes later. Each data point on the chart marks a reschedule event. In addition we found that cases were routinely under-posted, meaning that the scheduled case duration was not usually long enough to accommodate the actual case. This is why in Figure 7; we find the initial predictions (points along the y-axis, t=480 minutes) to
average below the line marking 100%. As the criterion amount and reschedule amounts increase, our mean end of
day predictions becomes marginally worse. Figure 7 marks this result using linear trend lines.

Figure 7: Prediction accuracy graph for scenario 1 with equal parameter values

When considering differing parameter values we find more varied results. Figure 8 shows a criterion amount of 60
minutes, but with varying reschedule amounts between 15 and 60 minutes. We discovered that the larger the
difference between the criterion amount and reschedule amount, the worse our end of day prediction becomes. For
example, looking at (60, 15) and (60, 30), case predictions becomes worse as the room progresses (the trend lines
have a negative slope). The trend seen here also applies to different criteria amounts (i.e. 30 and 45 minutes).

For the sake of brevity (and the fact that the key takeaways for scenario two are the same as scenario one), we omit
the full presentation of results from scenario 2. The main difference was that scenario 2 (the longer cases) had more
pronounced under-posting of cases. We can also confirm that it is best to have the reschedule criterion and
reschedule amount equal based on all the results tested. We found that to minimize the start time offsets, end time
offsets, and increase the accuracy of the end of day predictions the criterion amount should be as low as possible.
However, the cost of a rescheduling event should also be considered.
In conclusion, we have developed a discrete event simulation model that was used to simulate regular days in an OR suite to test different rescheduling policies. We explored different right-shift rescheduling policies by changing two parameters, criterion amount and reschedule amount. We confirmed the idea that as we increase the criterion amount, we have fewer reschedule events, larger start and end time offsets, and worsen our end of day prediction. We also discovered that using differing values for criterion amount and reschedule amount and when using a reschedule amount smaller than the reschedule criterion negatively affects the system. The larger the difference between the two parameters the worse our start and end time offsets and end of day predictions become. These key takeaways held true for both scenarios proposed (60–90 minute scheduled cases and 90–120 minute scheduled cases).

In order to strengthen the validity of our simulation model, a more thorough approach to input analysis will be completed with future research. This could be achieved by conducting a more thorough statistical approach to fitting distributions including the use of goodness of fit tests. In addition, conditionally probability models or other methods could be used to provide a more accurate estimation for our remaining time in each case. We also acknowledge the fact that more statistical analysis could be used to strengthen our results for end of case prediction times. We also consider generating new data sets that more closely adhere to estimated case duration times. This would be used to explore how the model responds to a smaller number of under-posted cases on the schedule.

In the future, we plan to explore more realistic scenarios including consideration of two or more rooms, the addition of surgeon constraints, and the addition of add-on cases. In addition, we would like to begin testing more policies including left-shifting, partial regeneration, and complete regeneration of the surgery schedule. As our simulation
model becomes more realistic and more advanced, we will be able to use this modeling approach to gather performance data on different hospital rescheduling policies. In addition, if we can specify all of the costs of the system, we would be able to test different policies against each other rather than comparing them subjectively.

References