Applications of Information Systems to Operating Room Scheduling

In this issue of the Journal, Wright and colleagues identify the problems in accurately predicting surgical case duration. They did three clinical studies. In the first study, surgeons estimated how long it would take them to perform an operation. Their predicted durations, and those of a commercially available computer scheduling system (Surgiserver, Serving Software, Minneapolis, MN) were compared with actual durations. Surgeons' estimates were more accurate than estimates made by the scheduling software. In the second and third studies, several statistical models that incorporated the surgeons' estimates, scheduling software's estimates, and patient specific data were used to improve the accuracy of predicted surgical durations. None of the patient information improved the accuracy beyond a model that used the scheduling software's and surgeons' estimates.

These results are very important. Accurate predictions of operating room times for surgical procedures are probably a prerequisite for matching operating room suite workload to capacity. Wright and colleagues showed that the use of surgeons' estimates can increase the accuracy of the commercial scheduling software to predict duration of surgery. However, the percentage improvements achieved by mathematical modeling versus either the surgeons' or scheduling software's estimates alone were not large (12% to 18%).

Improvements in predicting case duration are of interest to hospitals because they may be able to produce cost savings. Improved operating room scheduling could reduce costs by reducing underutilization (operating rooms empty earlier than expected) and/or by reducing unplanned extension of the workday (operating rooms occupied later than predicted). The benefits of better scheduling would also extend to patients (i.e., reduce unnecessary waiting time). However, we limit our discussion to the perspective of the hospital. Whether more accurate scheduling of cases can reduce costs requires understanding of the cost structure underlying operating room time. We measured the basic cost for 1 min of operating room time at Stanford University Medical Center, excluding incremental resources specific to a procedure. This $8.13 cost per minute can be divided into various parts (table 1). Using these data, if more accurate scheduling were to allow better arrangement of cases throughout the operating room suite (e.g., fewer rooms are empty unexpectedly early), then nurses and technicians (28% of costs) might be used more efficiently. Similarly, if the scheduling improvements were so good as to permit permanent closure of operating rooms, then direct fixed costs (clerks, housekeepers, etc.; 16% of costs) might be decreased. Alternatively, better scheduling may allow additional scheduled cases to be done during the day, which would increase revenue. Reducing indirect fixed costs (e.g., medical records) is important because they account for 36% of operating room costs. However, they are usually not under the direct authority or purview of anesthesiologists.

Wright and colleagues tried to identify confounding factors that might improve predictability of surgical durations. They used sophisticated regression methods. They examined the potential benefits of including gender, whether the surgery was going to be bilateral, rank of the resident assistant, whether diagnosis was known when the procedure was scheduled, and the surgeon's estimate of case difficulty. Such patient data did not improve the computer predictions. Just including a combination of the surgeon's and computer's estimates was as accurate a predictor as including any of these other factors. These insightful data may vastly simplify and focus future investigations of methods to improve predictability of case duration. However, the authors focus on predicting average case durations. For many problems in operating room management, the appropriate statistical goal is different. For example, if an operating room manager wanted to know whether there is a reasonable (e.g., 95%) chance that a case will be completed in a certain amount of time, different statistical methods would be needed.

The authors' results overestimate the benefit a hospital would obtain in using the commercial software to schedule its operating room suite. The scheduling software studied by Wright and colleagues stored duration and surgical procedure for each case. The software used these prior case times to predict the duration.
Table 1. Hospital Costs to Run an Operating Room for 1 Min

<table>
<thead>
<tr>
<th>Source of Operating Room Costs</th>
<th>Costs ($/min)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two nurses or technicians in each room</td>
<td>2.24</td>
<td>28</td>
</tr>
<tr>
<td>Basic supplies (e.g., linens)</td>
<td>1.69</td>
<td>21</td>
</tr>
<tr>
<td>Control desk clerks, housekeeping, etc.</td>
<td>1.27</td>
<td>16</td>
</tr>
<tr>
<td>Hospital administration, medical records, etc.</td>
<td>2.93</td>
<td>36</td>
</tr>
<tr>
<td>Total</td>
<td>8.13</td>
<td></td>
</tr>
</tbody>
</table>

The scheduling of the same surgery. However, the scheduled operation may not be the same as the actual procedure. For some operations, there are rarely differences between the scheduled and actual cases (e.g., carpal tunnel release). But, for other procedures, scheduled operations often differ from performed procedures (e.g., Whipple). A way around this problem is for hospitals to record, for each operation, the scheduled procedure, the actual operation, and the actual case duration. Future case durations would be predicted using actual case durations classified by the scheduled operation. For example, the predicted duration of a Whipple procedure would not be calculated from prior durations of Whipples, but from cases that were scheduled as a Whipple.

A premise of Wright and colleagues' analysis is that the duration of one case in an operating room can be predicted from the durations of other cases in the room. We expect that this assumption would hold at some hospitals but not others. For example, if a first case of the day, which was supposed to last 3 h, took longer, the surgeon's next scheduled 3-h case might take less time than expected, because the schedule is running late. Some surgeons may speed up to get to their office or home, but others may not. If the surgeon were to speed up, the authors' results would overestimate the benefit the commercial scheduling software would obtain.

In addition, Wright and colleagues' study does not give much insight into how well statistical analysis of prior case durations alone can do to predict future case durations. We should not interpret their study as evidence in support of or against the value of computer scheduling of operating room cases. Their analysis is based on a commercial scheduling software program. The accuracies of its predicted durations depend on the validity and reliability of the statistical method the software used.

We evaluated the accuracy and precision of this software's statistical method. To predict case duration, this software used the last two prior case durations, discarded the longest and shortest case, and then calculated the mean. We compared this method to using the median of 39 prior durations. To compare these two methods of analysis, we used a database of actual case durations that were described elsewhere. Computer simulation assessed the accuracy and precision of the predicted case durations. The methodology was described in detail elsewhere, and details of the results are available from the authors.

Performance of the median of 39 prior cases was better than the performance of the truncated mean of 10 prior cases. The absolute accuracy of the median exceeded that of the trimmed mean. Also, for all operations, the trimmed mean overestimated the duration of the next case. Therefore, the trimmed mean was a biased estimator of case duration. It predicted reliably that case duration would be longer than it really was. Results were similar for precision of the estimates. For all surgical procedures, precision was better for the median than for the trimmed mean.

These simulation results do not, and are not meant to, suggest that the median of 39 prior durations is the preferred method to predict case duration from prior times. We hope that, in the future, superior predictors of case duration will be developed. But, the implication should be clear for the appropriate interpretation of Wright and colleagues' study. The performance of the commercial scheduling software they used should not
be extrapolated to all statistically based methods to schedule operating room cases. Wright and colleagues’ research gives important lessons about how to predict case duration. Combining surgeons’ estimates and prior case duration data may be the most successful method. Their work does not tell us about the relative importance of these two types of information.

Franklin Dexter, M.D., Ph.D.
Assistant Professor
Department of Anesthesia
University of Iowa, College of Medicine
200 Hawkins Drive
Iowa City, IA 52242
Electronic mail: franklin-dexter@uiowa.edu

Alex Macario, M.D., M.B.A.
Assistant Professor
Department of Anesthesia
Stanford Medical Center, H3580
Stanford, CA 94305-5115
Electronic mail: macario_a@hosp.stanford.edu

References
1. Wright III, Kooperberg C, Bonar BA, Bashein G. Statistical modeling to predict elective surgery time: Comparison with a computer scheduling system and surgeon-provided estimates. Anesthesiology 1996; 85:1235-1245
2. Dexter F. Application of prediction levels to operating room scheduling. AORN J 1996; 63:607-615

Anesthesiology, V 85, No 6, Dec 1996