Surgeon Profiling

A Key to Optimum Operating Room Use

Suzanne M. Sokal, MSPH; Yuchiao Chang, PhD; David L. Craft, PhD; Warren S. Sandberg, MD, PhD; Peter F. Dunn, MD; David L. Berger, MD

Hypothesis: A high-efficiency Pod, composed of 3 parallel-processing operating rooms (ORs) and a dedicated 3-bed miniature postanesthesia care unit, can be filled with surgeons capable of converting time saved from parallel processing into incremental volume.

Design: Statistical and mathematical modeling.

Setting: Academic medical center with 52 serial-processing ORs, 1 parallel-processing OR, and a congested postanesthesia care unit.

Participants: Elective surgical cases (N=58,356) performed by a single surgical service without a preoperative intensive care unit bed request from April 1, 2004, through March 31, 2006.

Interventions: Results from our parallel-processing OR (n=1729) were extrapolated to all other cases (n=56,627) to estimate the duration of key process time intervals as if they were performed using parallel processing. Cases that could yield incremental throughput using parallel processing were labeled “good.” Total good case hours per week were then aggregated for each surgeon.

Main Outcomes Measures: Surgeons with 4.5 hours per week or more of good case time had a “profile” suitable for a 9-hour block in The Pod every 2 weeks.

Results: Of the 352 profiled surgeons, 30 had 4.5 hours per week or more of good case time, more than filling the 15 blocks per week.

Conclusions: The high-efficiency OR Pod can fill each of its 3 ORs with case/surgeon combinations that should yield additional throughput. Surgeon profiles based on stringent efficiency targets maximize the throughput potential of The Pod’s active ORs and more than compensate for the OR turned miniature postanesthesia care unit.

Arch Surg. 2007;142:365-370

Increasing demand for surgical services is creating capacity constraints in many surgical suites. In addition, the health care industry is experiencing financial pressure brought about by a combination of shrinking reimbursements and a cost structure characterized by a high proportion of fixed costs. In this environment, mission and profitability are aligned in the goal of increasing the number of patients cared for in the system. To that end, many organizations have begun to rethink patient flow in an attempt to create capacity and ease production pressure. Operating room (OR) suites are no exception. Several organizations have tried a technique in which anesthesia induction and OR turnover are performed concurrently (“parallel processing”), with mixed results.

The successful application of parallel processing balances increased costs derived from incremental staffing with a proportionate decrease in variable costs (ie, overtime reduction) and/or offsets the added expense with increased revenue achieved through additional case volume. Increasing the scale of the environment from a single-room to a multiple-room model creates a relative advantage because staffing costs are distributed over more cases. A second and equally important strategy is to selectively deploy the technique to surgeon-case combinations capable of translating saved time into additional cases per block because benefits of parallel processing do not accrue equally.

We created a model environment called The High-Efficiency Pod (The Pod), consisting of 4 rooms made up of 3 parallel-processing ORs and a dedicated 3-bed recovery room. Previous mathematical modeling demonstrated that there is more than enough case time that is sufficiently faster to allow more throughput in 3 par-
parallel-processing rooms than in 4 standard ORs (“good cases”) to fill The Pod each week. However, a second level of analysis is required to determine whether all 15 blocks can be assigned to individual surgeons. Consequently, it is important to profile surgeons to determine which surgeons have a sufficient number of good cases to fill a 9-hour block at least twice a month.

METHODS

Data for all elective surgical cases (including all cases with the following schedule codes: scheduled and wait-listed nonurgent) performed between April 1, 2004, and March 31, 2006, were extracted from the Nursing Perioperative Record. The Nursing Perioperative Record is an internally developed computerized system for perioperative documentation. This system captures milestone events in the perioperative course of care. These milestones are used to calculate critical time intervals. The accuracy and precision of the time stamps used in this study were validated by Sandberg et al.2 Data elements included time stamps for milestones in the perioperative process (Table 1) and provider (surgeons and anesthesia staff, including anesthesiologists and nurse anesthetists), procedure, administrative, and clinical data elements for each case. Because The Pod is designed for efficiency, the following case types were eliminated: cases that involved more than 1 surgical service (n=1653), cases with a preoperative intensive care unit bed request (n=5324), and cases for patients of another health care facility (n=60).

Missing values for each time interval were as follows: preoperative anesthesia preparation, 1231 (2.1%); OR anesthesia time, 510 (0.9%); surgical preparation, 1304 (2.2%); surgery, 0; OR emergence, 335 (0.6%); and turnover, 20 597 (33.3%). For induction, OR anesthesia time, and OR emergence, the mean durations for cases with the same anesthesia type and American Society of Anesthesiologists’ class were used. For surgical preparation times, the mean duration for the surgical service was used. The average room in our main OR is used for 2 to 3 cases per day. The last case of the day in a room does not turn over, so approximately 35% of cases do not have a turnover time associated with them. For those cases without a turnover time, the following logic was applied: cases consisting of a single procedure were assigned the mean turnover time for that procedure, whereas cases consisting of multiple procedures were assigned the service mean turnover time.

Next, the time savings from a single parallel-processing OR (1729 [3.0%]) was extrapolated to cases (n=56 627) completed in a standard OR using a regression model. Output from the statistical model consisted of projected durations of each time interval. The results of the statistical model were housed in a database along with historical performance on each of these patient flow milestones.

To generate surgeon profiles, each case in the database was first evaluated for “fit” in The Pod using a “goodness” ratio. The goodness ratio mathematically captures the essence of The Pod efficiency question: does the benefit of parallel processing in The Pod’s 3 active ORs offset the loss of an OR to the role of a recovery room? In short, good cases will have achieved sufficient time savings in a parallel-processing environment to allow a sufficient number of additional operations to occur, thus compensating for losing 1 OR. To calculate the goodness ratio, the time of each case performed in The Pod is compared with that of a traditional environment.

System throughput is given by the number of active ORs (3 for The Pod and 4 for the traditional configuration) divided by the duration of patient processing time. In a traditional OR, patient processing time is calculated using historical performance on key process intervals, except in the case of preoperative anesthesia preparation. An unknown portion of preoperative anesthesia preparation occurs simultaneous to turnover in a traditional environment. In this analysis, the fraction of time assumed to run in parallel is 50%. Therefore, the patient processing time is calculated according to the following formula:

\[
\text{Time}_{\text{traditional}} = \left(\frac{\text{Preoperative Anesthesia Preparation} \times 0.5}{\text{OR Anesthesia Time + Surgical Preparation + Surgery + OR Emergence + Turnover}}\right)
\]

When calculating the patient processing time for The Pod, historical performance was used for all intervals except turnover and OR anesthesia time, for which modeled performance was used. Because the goal of this analysis is to identify surgeons capable of optimally using a parallel-processing environment, the preoperative anesthesia preparation and OR anesthesia time intervals used were modified to compensate for the effect of the anesthesia provider (an anesthesiologist or nurse anesthetist) by using procedure-specific mean durations. Because of concurrent induction and turnover, patient processing time in The Pod is captured by the following formula:

\[
\text{Time}_{\text{pod}} = \text{OR Anesthesia Time + Surgical Preparation + Surgery + OR Emergence + Turnover}
\]

Table 1. Description of Surgical Milestones and Critical Time Intervals*

<table>
<thead>
<tr>
<th>Time Interval Name</th>
<th>Milestone</th>
<th>Time Interval Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preoperative anesthesia preparation</td>
<td>Patient in room</td>
<td>Anesthesia provider notes event in anesthesia information system, signaling the beginning of billed activities, to patient crosses OR door into OR</td>
</tr>
<tr>
<td>OR anesthesia time</td>
<td>Patient in room</td>
<td>Patient crosses OR door into OR to patient is ready for surgeon to begin positioning and skin preparation</td>
</tr>
<tr>
<td>Surgical preparation</td>
<td>Patient ready for surgeon</td>
<td>Patient is ready for surgeon to begin positioning and skin preparation to surgical instrument contacts the patient’s skin</td>
</tr>
<tr>
<td>Surgery</td>
<td>Surgery start</td>
<td>Surgical instrument contacts the patient’s skin to dressing applied after the procedure or drapes removed if no dressing was used</td>
</tr>
<tr>
<td>OR emergence</td>
<td>Surgery finish</td>
<td>Dressing applied after the procedure or drapes removed if no dressing was used to patient out of OR</td>
</tr>
<tr>
<td>Turnover</td>
<td>Patient out of room</td>
<td>Time from the previous patient out of OR to time of the next patient in OR for sequentially scheduled cases</td>
</tr>
</tbody>
</table>

Abbreviation: OR, operating room.

*Milestone and critical time interval definitions adapted from Sandberg et al.2

†Anesthesia provider refers to an anesthesiologist or a nurse anesthetist.
A case is flagged as good if its throughput in The Pod is greater than or equal to the throughput in the traditional environment, which is equivalent to a goodness ratio as follows:

\[ \frac{\text{Time}_{\text{Pod}}}{\text{Time}_{\text{Traditional}}} \geq 0.75. \]

The goodness ratio identified each case as either good or bad. Surgeon profiles were generated by summarizing hours for good and bad cases to yield the average total number of good case hours per week for each surgeon. A surgeon with a minimum of 4.5 good case hours per week had an acceptable profile for The Pod because the surgeon would be capable of filling a 9-hour block every 2 weeks.

A "profile" was generated for every surgeon (N = 352) who operated during the study period. Of these surgeons, 169 completed at least 50 cases during this time frame, accounting for 56.167 or 96.2% of the 58356 cases included in this analysis.

Next, for surgeons with an acceptable profile, surgeon-case combinations were examined. For each surgeon-case combination, the ratio of good-total case hours and the mean surgical times for good and bad cases were determined. Given that surgeons would be asked to select appropriate cases to schedule into The High-Efficiency Pod, the ability of surgeons to accurately predict good cases based on information that is available prospectively is important. A regression analysis was performed on a subset of surgeon case types to test whether the inclusion of readily available patient information would improve the ability to predict which cases would fit well in The Pod.

### RESULTS

Total and good case time, as measured in hours per week, was calculated for each surgeon. First, cases were stratified on the basis of their goodness ratio, such that cases with a ratio of 0.75 or less were classified as good and cases with a ratio greater than 0.75 were labeled as "bad." Next, we added the projected surgical (surgical preparation, surgery, and OR emergence) and turnover times for all cases and all good cases performed by each surgeon during the study period. The resulting total case time and total good case time was divided by the number of weeks included in the data set. Results for each surgeon are expressed as a weekly average for total and good case times. In addition, the percentage of cases for that surgeon that are labeled good are included (Table 2).

Every block in The Pod should be filled by 1 surgeon per room per day to avoid delays induced by changing surgeons. Consequently, a surgeon must be able to fill at least one 9-hour block every other week to be included in The Pod. Thus, surgeons would be considered good candidates if they routinely have at least 4.5 hours of good case volume per week. Of the 352 surgeons profiled, 30 meet or exceed the minimum number of good hours per week. Within this cohort, 7 surgeons have 9.0 hours or more of good cases per week, 9 have between 7.0 and 8.9 hours per week, and the remaining 14 have between 4.5 and 6.9 good case hours during an average week. In aggregate, the 30 surgeons with acceptable profiles account for 23,904 (41.0%) of the cases included in the analysis.

Among surgeons with an acceptable profile, the percentage of good case hours ranges from a high of 64.2% to a low of 19.7% of their total case hours per week (Figure). This finding calls into question the ability of an individual surgeon to routinely select cases that fit into the Pod environment. We then tried to establish a more clear set of good case selection criteria by quantifying the proportion of specific procedures classified as good for each surgeon. Unfortunately, for any given procedure, there is often a high degree of variability in surgical times for a particular surgeon and among surgeons performing a specific procedure.

If one looks at intrasurgeon and intersurgeon performance for a specific procedure among surgeons with a Pod-friendly profile, for a common procedure such as total hip replacement (Table 3), it becomes clear that it will be difficult for some surgeons to select cases that would work in The Pod. For this particular procedure, some surgeons have a profile such that 38.7% of cases are good while other surgeons' cases are acceptable only 18.9% of the time.

In an attempt to better differentiate good from bad cases, logistic regression was used to try to predict good cases. Total hip replacement is presented as an example. Before using logistic regression, there were 797 cases among the 5 surgeons, and 413 (51.8%) of these cases were good. Only patient factors routinely known in advance of surgery at our organization were considered. The factors were American Society of Anesthesiologists’ class, anesthesia type, age, and sex. The model excluded 251 (60.8%) of good cases and improved the proportion of good cases to 62.3%. In the end, the patient factors used in this analysis did not add much predictive value to the model.

### COMMENT

The successful application of parallel processing minimizes fixed costs and distributes those costs over the maximum possible number of patients. Because salary expense accounts for most fixed costs, structuring the environment in a manner that minimizes staff requirements is prudent. Many applications of parallel processing have used a single OR suite, thus almost univer-
it is a function of the absolute time savings per turnover and the inherent duration of the surgical procedure. Sandberg et al.\(^2\) highlighted the types of efficiency gains that surgeons can realize from parallel processing as follows: none, ending the day earlier, reducing or eliminating overtime, and/or completing additional cases. To that end, surgeon profiles were generated in an attempt to determine if a sufficient number of surgeons and/or surgeon-case combinations capable of completing more cases per block in a parallel-processing environment could be identified to fill the fifteen 9-hour blocks each week.

Preliminary analysis identified 30 surgeons with a minimum of 4.5 good hours per week or a 9-hour block every 2 weeks. No surgeon had a profile in which 80% or more of their cases achieved the maximum benefit of parallel processing. To that end, a second analysis stratifying each surgeon’s cases into good and bad cases by procedure was performed. Even after controlling for procedure, there was great variability as to the proportion of good cases within that procedure type, such that for some surgeon-case combinations the proportion of good cases was in excess of 87% while for others it was as low as 5%. Further attempts to better predict good cases using logistic regression with prospectively known clinical factors did not prove valuable.

This study has several limitations. The mathematical model only flagged cases as good when the total time in the OR was 25% less using parallel processing compared with historical performance. This stringent cutoff, while analytically practical and mathematically pure, introduces several potential limitations into the accuracy of the surgeon profiles. A strict cutoff does not take into consideration the degree of goodness or badness of an individual case nor does it put that case in the context of a surgeon’s typical daily operative schedule. While 1 or even 5 minutes in case length may be relatively unimportant in practice, it was critically important in the model.

The statistical model introduced a potential limitation to the analysis. Specifically, the experience from a single parallel-processing OR was extrapolated to all cases.\(^7\) To the extent that cases performed in the single parallel-

---

Table 3. Surgeon Profiles for Total Hip Replacement\(^*\)

<table>
<thead>
<tr>
<th>Surgeon No.</th>
<th>Cases</th>
<th>Procedure Times§</th>
<th>% Good‡</th>
<th>Good</th>
<th>Bad</th>
<th>Total†</th>
<th>% Good‡</th>
<th>Total†</th>
<th>% Good‡</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>230</td>
<td>162</td>
<td>392</td>
<td>58.7</td>
<td>135</td>
<td>170</td>
<td>58.7</td>
<td>135</td>
<td>170</td>
</tr>
<tr>
<td>2</td>
<td>147</td>
<td>116</td>
<td>263</td>
<td>55.9</td>
<td>142</td>
<td>181</td>
<td>55.9</td>
<td>142</td>
<td>181</td>
</tr>
<tr>
<td>3</td>
<td>15</td>
<td>30</td>
<td>45</td>
<td>33.3</td>
<td>167</td>
<td>215</td>
<td>33.3</td>
<td>167</td>
<td>215</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>28</td>
<td>40</td>
<td>30.0</td>
<td>185</td>
<td>209</td>
<td>30.0</td>
<td>185</td>
<td>209</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>43</td>
<td>53</td>
<td>18.9</td>
<td>164</td>
<td>195</td>
<td>18.9</td>
<td>164</td>
<td>195</td>
</tr>
</tbody>
</table>

*Defined as operating room scheduling code 8151S. †Reflects all cases performed during the study period. ‡Reflects the percentage of good cases performed during the study period. §Reflect mean times for cases flagged as good or bad performed during the study period.
processing OR are not representative of process requirements for all operative cases, the time estimates used in this analysis may prove inaccurate. However, this room performs a diverse mix of open and laparoscopic cases with block allocations to general surgery (60%), gynecology (20%), and urology (20%). We also estimated the extent to which preoperative anesthesia preparation occurred in parallel to turnover to be 50% of the total time for that interval. While we understand that the actual proportion varies widely around this estimate, this figure was thought to be a fair estimate, arrived at based on discussions with surgeon, anesthesia, and nursing leadership.

By standardizing the anesthesia-controlled times within a given case type, the influence of the anesthesiologist was not considered. Moreover, numerous systems and workflow improvements planned as part of this project were not included in the model. While this makes sense in terms of profiling surgeons, it minimizes the potentially significant effect of team performance on outcomes. We expect to see significantly better performance in practice but wanted to present a conservative estimate of the effect of this workflow redesign. Furthermore, only 1 potential approach to better predicting good cases was explored, namely, logistic regression. More general nonlinear prediction techniques from the field of machine learning should be explored. It is also possible that surgeons may be able to better predict good cases; however, this was not explored.

To solve the capacity and profitability problems facing health care, organizations will have to be more creative in their approach to resource allocation. Parallel processing of parts of the perioperative process is a promising tool for more efficiently and effectively allocating resources. Like any tool, it does not work in every situation. In fact, parallel processing will likely work best for repetitive cases of short duration (ie, 30 minutes to 2 hours). The maximum benefit of parallel processing occurs when the technique is deployed to select surgeons and/or surgeon-case combinations (ie, those whose practice allows additional cases to be performed without expanding scheduled hours) and when economies of scale are achieved.

As this analysis highlights, even surgeons with the most Pod-friendly profile have many cases that are better suited to a traditional environment. An efficiently functioning OR suite must acknowledge this fact and adjust accordingly, allocating different resources to balance productivity. In the case of The Pod, surgeons will have time in and out of The Pod each week to allow for case stratification. By more effectively matching surgeon-case combinations with the appropriate resources and an appropriately designed environment, significant improvements in productivity can be achieved.

Accepted for Publication: November 1, 2006.
Correspondence: David L. Berger, MD, Department of Surgery, Massachusetts General Hospital, 15 Parkman St, Wang 465, Boston, MA 02114.

Author Contributions: Study concept and design: Sokal, Sandberg, and Berger. Acquisition of data: Sokal and Berger. Analysis and interpretation of data: Sokal, Chang, Craft, Sandberg, Dunn, and Berger. Drafting of the manuscript: Sokal, Chang, Craft, and Berger.

REFERENCES

4. Torkki PM, Marjamaa RA, Torkki MI, Kallio PE, Kirvela OA. Use of anesthesia induction rooms can increase the number of urgent orthopedic cases completed within 7 hours. Anesthesiology. 2005;103:401-405.
5. Smith MP, Helfand R, Schubert A. QuEST operating room system decreases turnover time over 50% for hip and knee replacements [abstract]. Anesthesiology. 2005;103:A1181.

DISCUSSION

Orlando Kirton, MD, Hartford, Conn: The question I have is as follows: you specifically address the need to be very selective; you have to be selective of the surgeon and you have to be very selective of the anesthesiologist to make this work. Are you concerned that you could potentially create hostilities, create favorite sons and daughters? Could you end up penalizing the rest of the OR by selecting out the best anesthesiologists, best nurses into these Pods, leaving the rest of the surgeons with less than optimal support?

Dr Berger: What I would ultimately envision when building an operating suite is that you would have several sections. One area that is a better fit for higher-efficiency surgeons, a Pod for cases that take a longer period of time, and finally a larger area for extended-length cases. The resource allocation is good for everyone. That is what you have to get across: that everybody gets what they need. If you are going to do 1 case a day, whether the start time is at 7:45 or 8:00 makes no difference, but if you are going to do 5 cases in 1 day in 1 room, if your turnover is 45 minutes vs 30 minutes, it makes a huge difference because you cannot get it done. The key is communication and getting everybody to work together as a team to understand that everyone is going to get their resources and it is going to be done in a fair manner. But you cannot take a surgeon who has a capability of doing “X” number of cases and give [him or her] the same resources as a surgeon who is going to do 1 case. It just does not make sense from an efficiency standpoint.

Randolph Reinhold, MD, New Haven, Conn: I presume you have modeled this idea compared to giving the “busy hand” surgeon with the 8 cases 2 rooms on a given day and have him move back and forth rather than remodel the existing OR.
Dr Berger: I am a big believer in not having 2 rooms, so I do not think that is a good thing to do. I do not think it really complies with all the appropriate standards, and we could have an argument about that but, no, this is all modeled to 1 surgeon 1 room. We have a certain number of ORs, and we have to make the best with what we have from 1 room. It is not modeled that way at all.

Dr Reinhold: When the length of the case equals the length of outpatient reasonable induction time, then you really are in a situation where the surgeon in fact moves from room to room but still does not have more than any 1 patient in an OR at any moment.

Dr Berger: I agree with you and that is 1 area. In hand surgery, if you have a hand surgeon who can perform a case in 15 minutes, that person is the 1 person that I can figure who can do 2 rooms who is actually present from the beginning to the end of their cases in both rooms. Hand surgeons actually do not qualify to be in The Pod because they go too fast and it does not make sense to have them in an environment like this.

Thomas Tracy, MD, Providence, RI: Can you characterize the specialties, or the largest groups of specialties, those that are the “left-hand side” of your analysis as well as the most efficient ones?

Dr Berger: To give you an example of how hard this actually is to do, this has been 3 years in the making. Last year, I presented the preliminary model and we still have not built it yet. The construction is going to start soon and we hope to start in February. In terms of what groups can go into this, the best groups are your consistent surgeons, so the person who does it the same way every time, your total hip people, your total knee people, colectomy surgeons. Any medium-length operation fits into this very well when there is volume and reproducibility. A distal pancreatectomy fits in here very well. A Whipple is a little bit long. You can name the cases. Carotids fit in here very well. Some pediatric surgery.

Hardy Hendren, MD, Boston, Mass: I congratulate you on trying to do something about the logjam in the OR. When I started in the OR in 1952, the big problem was anesthesia turnover time, getting the anesthesia department not to assign the same young resident to case after case after case. When I left the OR 52 years later (2 years ago), the same problem existed, unchanged. My entire career was done in 2 hospitals, the MGH [Massachusetts General Hospital], where I have a 50-year pin, and Children’s Hospital Boston, where I have a 25-year pin. The turnover problem was the same in both hospitals when I started, and it was the same in both when I ended. However, during that half century, we surgeons applaud and recognize the immense improvement in the safety of anesthesia, which has occurred. Does your Pod system attack the turnover delays, which, in my experience, were the number one problem that we surgeons faced?

Dr Berger: Absolutely. The Pod is formed by anesthesia, nursing, and surgery working together. The key to doing any of this is having everybody buy in and work on the same team. When you talk about parallel processing, which I did not really want to get into too much because I presented that before, the idea is to overlap anesthesia and turnover so that everybody is moving at the same time rather than proceeding linearly, which is the way the system was built.

Financial Disclosure: None reported.