Proceedings of the 1999 Winter Simulation Conference
P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, eds.

DETERMINATION OF OPERATING ROOM REQUIREMENTS USING SIMULATION

Julie C. Lowery
Jennifer A. Davis

VA Center for Practice Management and Outcomes Research
P.O. Box 130170
Ann Arbor, MI 48113-0170, U.S.A.

ABSTRACT

In 1997 Brigham and Women's Hospital (BWH) in Boston initiated a construction project to renovate its existing surgical suite to include 32 operating rooms--two less than the current number. The new suite would be used for performing primarily inpatient cases; 95% of all outpatient cases would be moved to another facility. BWH administrators, planners, and clinicians wanted to be sure that the 32 rooms would be sufficient for accommodating projected increases in the inpatient surgical volume. In addition, they wanted to examine the possible effects of changes in the surgical schedule and in case times on the number of rooms required. A simulation model using MedModel simulation software was developed for examining these issues. The resultant model includes a number of assumptions that simplified model construction, yet still resulted in a valid model that met project objectives. The model showed that the projected changes in surgical workload could be accommodated in 30 operating rooms (or fewer) if scheduled block time were extended during the weekdays and Saturday blocks were added.

1 INTRODUCTION

In 1997 Brigham and Women's Hospital (BWH) in Boston initiated a construction project to renovate its existing surgical suite to accommodate primarily inpatient cases. Outpatient cases, which were being performed in the suite when the project began, were going to be moved to a separate ambulatory surgery facility. The new inpatient suite would include 32 operating rooms, which was two less than the number of rooms in the suite prior to renovation. Inpatient surgical volume was projected to increase; but with the elimination of the ambulatory surgery procedures, it was expected that the new suite would be able to accommodate the inpatient volume. However, BWH administrators, planners, and clinicians wanted to be sure that the 32 rooms would be sufficient. In addition, they wanted to examine the possible effects of changes in the surgical schedule and in case times on the number of rooms required. They selected simulation as the methodology for investigating these issues. The remainder of this paper describes the model and the results from its use.

2 MODEL DESCRIPTION

The model was developed using MedModel simulation software and consists of the following components:

Generation of the block schedule. Each block in the surgical schedule is represented by a separate entity (transaction), which has the following attributes: specialty, operating room, day of week, initial arrival time, and block start and stop times. The initial arrival time is the number of minutes after the beginning of the week when the entity arrives. For example, an entity representing a block with a Thursday 7:30 a.m. start time has an arrival time of 4,770 minutes (Thursday at 7:30 a.m. is 4,770 minutes after 12:00 a.m. Monday, which is the model starting time). Block start and stop times correspond to the number of minutes past midnight for the day that the block begins and ends. For example, the Thursday 7:30 a.m. - 8:00 p.m. block has a starting time of 450 (7:30 a.m. corresponds to 450 minutes after 12:00 a.m.) and a stop time of 1200 (8:00 p.m. corresponds to 1200 minutes after 12:00 a.m.).

Generation of cases. Once the entity representing a given block arrives in the system, it obtains a surgery time by sampling from the historical distribution of procedure times for its specialty and adding 35 minutes for case turnaround time (i.e., time for case clean-up and set-up). If adding the procedure time plus turnaround time to the current time exceeds the end of the block time by a certain amount of minutes (see discussion of "Downtime" below), the entity leaves the system and the block is completed for the day. Otherwise the case proceeds to the correct operating room (OR) where it remains for the sampled amount of procedure time plus turnaround time. After an entity completes its time in the OR, it leaves the model, another entity from the same block is generated, and the process is repeated.

Procedure time distributions. In addition to the start and stop times for each of the blocks, model input data include the procedure times for each of the specialties. The procedure times are obtained from the actual distribution of inpatient and outpatient surgery times in 1996, excluding 95% of the outpatient surgery cases, because it was expected that 95% of the outpatient cases would be performed outside the renovated OR suite. Thus, the procedure times associated with the remaining 5% of the outpatient cases are included in the procedure time database at a frequency consistent with their expected low rate of occurrence in the future. Cases performed after regularly scheduled hours were excluded from the procedure time distributions.

Separate procedure time distributions were generated for each specialty. In addition, if individual surgeons were allocated their own blocks in the schedule, separate distributions of procedure times were generated for these surgeons, and the surgeons' time data were removed from their specialties' distributions. Cases in blocks for surgeons who were added after 1996 (and, therefore, had no historical data) were sampled from the surgeon's specialty's distribution of procedure times.

Downtime. An important component of the model is downtime—i.e., that time during scheduled blocks when no activity is occurring in the OR (no surgery, set up, or clean up), usually due to delays (e.g., staff, patient, or equipment not available) or discrepancies between scheduled and actual surgery starting times. It is generally recognized that operating room utilization of between 80% to 85% (i.e., 15-20% downtime) is a realistic target in most operating rooms (Rinde and Blakely 1976).

It is critical that the *amount* of downtime be accurately modeled, because it has a direct effect on the number of cases that can be performed. However, the *timing* of when the downtime occurs in the model should not matter. Regardless of where it occurs in the model, it should have the ultimate effect of preventing the occurrence of any cases during that time. In reality, downtime is spread throughout the block. But rather than try to model downtime between cases by reconstructing an historical distribution of these times from the procedure time database (a not-so-simple task) and then sampling from this distribution between cases, we adopted the simpler approach of modeling all downtime at the end of the day.

Thus, downtime is modeled as the amount of time that is unused at the end of a block, which results from the application of the stopping rule described above under "Generation of cases"—i.e., if adding the procedure time plus turnaround time to the current time exceeds the end of the block time by a certain number of minutes, the entity leaves the system and the block is completed for the day.

The amount of downtime in the model was set by the hospital at approximately 20% (i.e., 80% utilization), based on their historical utilization rate. Therefore, the amount of time by which the last case can exceed the end of the block (the "stopping criterion") is varied until the desired utilization rate is achieved. The more time that is allowed past the end of the scheduled block, the more cases that can be scheduled and the higher the utilization that is obtained. The stopping criterion required for achieving a given utilization rate will vary depending on total number of blocks available, average procedure time, and desired workload. Thus, with each new scheduling or workload scenario that is modeled, the stopping criterion must be recalculated.

To the extent that the stopping criterion tends to eliminate the longer cases at the end of the day, the modeled average procedure times will be less than the actual average procedure times. When this occurred, the procedure time of the first case of the day in each block was increased to obtain the correct average.

Run length. Each replication of the model is run for 50 weeks (250 working days) to represent one year of operation. Elimination of two weeks from the year accounts for holidays and other light schedule days.

3 MODEL VALIDATION

Model validation consisted of comparing 1996 actual workload with model predictions, using 1996 procedure times and the 1996 block schedule as inputs. The model was run for ten replications to calculate an initial 95% confidence interval width for the performance measures of interest, total number of cases and total hours of cases performed. Table 1 shows the results of the final validation, including the difference between model prediction and actual number of cases, by specialty and overall, and the difference between model prediction and total hours, overall. A stopping criterion of 60 minutes past the end of the scheduled block time gave a utilization rate of 81.4%.

The results of the validation show fairly narrow 95% confidence interval widths of ±69.2 cases (0.3% of total predicted cases) and ±75.5 hours (0.1% of total hours), suggesting that additional replications are not needed to get a precise estimate of these performance measures. While neither of the confidence intervals includes the actual workload data, the difference between predicted and actual (-2.0% for number of cases and –2.5% for number of hours) was judged by the surgical staff to be sufficiently small that they considered the model to be a valid representation of actual operating room performance.

Table 1: Average Difference between Actual and Model Predictions of Number of Cases(1996 Data, 10 Replications)

	No. of Cases			
Specialty	**Model**	**Actual**	**Diff**	**% Diff**
General	1750	2090	-340	-16.3
Resident	623	502	121	24.1
Bone	67	66	1	1.5
Cardiac	1171	1091	80	7.3
Dental	100	53	47	88.7
Gyn	4340	4420	-80	-1.8
Neuro	664	850	-186	-21.9
Ophthal	408	354	54	15.2
Ortho	3501	3420	81	2.4
Oncol	1278	1334	-56	-4.2
ENT	544	590	-46	-7.8
Plastic	912	1228	-316	-25.7
Transpla	504	544	-40	-7.4
Thoracic	1377	1465	-88	-6.0
Urology	2436	2123	313	14.7
Vascular	814	773	41	5.3
TOTAL No.	20489[1]	20903	-414	-2.0
TOTAL Hrs.	52595[2]	53920	-1325	-2.5

[1]95% confidence interval = (20422, 20561).
[2]95% confidence interval = (52519, 52670).

The differences between model predictions and actual number of cases at the specialty level were much larger than at the overall level. However, because the complexities of actual block utilization by the individual specialties are not included in the model, the criteria for comparing predictions at the specialty level were not very strict. For example, not all specialties fully utilize their allocated block time on a regular basis, and other specialties routinely use more than their allocated block time. In the event where a specialty has not scheduled all of its allocated block time by 48 hours prior to the date of surgery, the remaining time is opened up to any other specialty who can use it.

To simplify model development, yet meet the desired objectives of the project, it was determined that it was not necessary to model the specialties' utilization of operating room time outside their allocated blocks. The key objective of the model is to determine how many blocks are required for accommodating total surgical workload, defined as total number of cases and total number of hours. If the model accurately predicts total surgical workload but not specialty workload, it is assumed that the modeled block schedule could indeed accommodate the predicted workload, as long as the following occurs: (1) blocks must be allocated to specialties in a manner that reflects the time required for their workload; and (2) specialties must be able to feasibly schedule cases within the designated time periods (i.e., the blocks must reflect the surgeons' schedules). To the extent that these criteria cannot be met, model predictions at the specialty level will be less accurate.

To further validate the model, surgical staff were asked if the differences at the specialty level shown in Table 1 accurately identified which specialties were over- or under-utilizing their block time. The surgical staff confirmed that all those specialties that showed a substantial negative difference between model predictions and actual workload were ones that were overutilizing their blocks (i.e., scheduling additional cases outside their blocks), and those specialties that showed a substantial positive difference were indeed underutilizing their blocks. Thus, the model was deemed sufficiently valid for its intended purposes.

4 DETERMINATION OF OPERATING ROOM REQUIREMENTS FOR FUTURE WORKLOAD

To determine the number of operating rooms required under different workload or operating assumptions, the block schedule is manipulated by trial and error until all of the required workload (for a year's period) is accommodated with the fewest number of blocks and the desired utilization rate. Additional time is added, or new blocks are created, for those specialties whose workload is projected to increase, and time (or entire blocks) are taken away from those specialties whose workload is projected to decrease. The revised model is run and if the total predicted workload is greater than or less than the projected future workload, the model is revised again. Similarly, if the predicted utilization rate is greater than 85% or less than 80%, the stopping criterion (defined in section 2) is changed.

The model was used to determine the number of operating rooms and blocks required for accommodating the projected 1998 workload. While total number of cases was projected to decrease significantly from 1996 to 1998 with the transfer of 95% of the outpatient cases to another facility, the total case hours for 1998 were projected to increase slightly. The increase in case hours was due to a projected increase in cases for several of the specialties with longer case times. As indicated in the introduction, the BWH surgical staff wanted to be sure that the main OR could accommodate this increase with two fewer rooms than were currently being used. To handle the increase, the staff were proposing extending the schedule for some of the rooms from 5:00 p.m. to 8:00 p.m., and adding Saturday blocks. In addition, the staff wanted to investigate the effect of a potential 7% reduction in case times (due to an expected increase in efficiency from staff and equipment) on room and block requirements.

Therefore, the following four scenarios were investigated for handling the projected 1998 caseload:

1. 7% reduction in 1996 case times, 6 Saturday blocks.
2. 7% reduction in 1996 case times, 10 Saturday blocks.
3. 1996 case times, 6 Saturday blocks.
4. 1996 case times, 10 Saturday blocks.

(The specialties of the Saturday blocks were defined a priori.)

The modeler was instructed to devise a new block schedule that used all of the new Saturday blocks, used the fewest number of rooms during the 7:30 a.m. to 3:30 p.m. shift, and met the following constraints for number of rooms running late: 24 rooms till 5:30 p.m., 12 till 8:00 p.m. Blocks were eliminated or shortened for those specialties whose workload decreased; and new blocks were added or lengthened for those specialties whose workload was projected to increase.

Each new model was run for ten replications, and model output (i.e., average total workload) was reviewed. If the modeled workload exceeded the projected workload, then blocks were eliminated or shortened. If the modeled workload was less than the projected workload, then blocks were added or lengthened. Once a block schedule was identified that accommodated the projected workload, the stopping criterion (for determining the last case in a block) was adjusted until a utilization rate of 80-85% was obtained. If adjusting the stopping criterion changed the projected workload, then the block schedule was readjusted. Adjustments continued to be made to the block schedule and the stopping criterion until the modeled workload was nearly equivalent to the projected workload and the utilization rate was between 80% and 85%. The resultant models for the four future scenarios ended up using a stopping criterion of 10 minutes past the end of the scheduled block time, to give an overall utilization rate of 80%.

An example of the format in which the final block schedule was presented for each of the four scenarios is shown in Table 2, which displays the schedule for scenario #1. In addition, the modeler prepared a detailed table of the specialties assigned to each of the blocks, so that BWH staff could review the schedule for feasibility.

For each of the scenarios, less than 32 rooms were needed to handle the projected workload. The results for each of the four scenarios can be summarized as follows:

1. 30 rooms with 20-23 rooms running till 5:30 and 5-12 rooms running till 8:00 (194.25 total blocks).
2. 29 rooms with 19-22 rooms running till 5:30 and 5-12 rooms running till 8:00 (192.0 total blocks).
3. 30 rooms with 21-22 rooms running till 5:30 and 10-16 rooms running till 8:00 (199 total blocks).
4. 29 rooms with 21-22 rooms running till 5:30 and 12-15 rooms running till 8:00 (198.25 total blocks).

Table 2: Proposed Block Schedule, 1998 Projections, Scenario #1

	No. of Rooms in Operation, by Day of Week					
Time	M	Tu	W	Th	F	S
7:30 a.m.-3:30 p.m.	30	30	30	30	30	6
3:30 p.m.-5:30 p.m.	20	22	22	23	20	0
5:30 p.m.-8:00 p.m.	10	12	12	5	8	0

The model was also used to investigate the effects of a change in turnaround time (from 35 to 40 minutes) and a decrease in the length of the third shift (from 8:00 p.m. to 7:00 p.m.). (The results are not presented here.)

5 DISCUSSION

The simulation model described in this paper was able to provide BWH staff with the information they sought regarding the effects of changes in workload and scheduling on number of operating rooms required. The information from the model justified the hospital's decision to proceed with the renovation, because it provided reassurance that 32 rooms would be sufficient for accommodating the projected workload. In addition, as with all simulation projects, the modeling process served to quantify a number of important assumptions and expectations of the surgical staff, which is an important component of the planning process that does not always occur.

Unfortunately, due to financial constraints, the renovation project was suspended and the outpatient cases were not moved from the current facility. So, while the feedback on the model was positive, we do not know whether the model projections were accurate. Of course, the accuracy of model projections is affected by many factors that are not related to the validity of the model (e.g., accuracy of the 1998 workload projections, feasibility of implementing the final block schedule, etc.); but it is still informative to compare actual system performance with model predictions whenever possible, to better understand a model's limitations.

While the model is an extremely useful tool for determining operating room and block scheduling requirements, one of its limitations lies in the trial and error approach necessary for determining these resource requirements. This approach is unavoidable, because the variables of interest--number of operating rooms and the block schedule--are model inputs rather than outputs. Adding and eliminating blocks requires some judgment on the part of the modeler, who, in this case, tried to set them to minimize changes to the current schedule, while meeting the workload requirements and constraints set by the surgical staff. Ultimately each of the proposed schedules must be carefully reviewed by the surgical staff to determine whether they are feasible to implement. (For example, if new blocks were added, someone must review these blocks to be sure that they do not conflict with the surgeons' clinic schedules.)

This model is a good example of the "keep it simple" rule in simulation (Salt 1993; Lowery 1996). It includes several important assumptions that significantly simplified the model development process, yet produced a valid model that met project objectives. For example, a more realistic model might include variable lengths of downtime between cases. Or, rather than discarding long cases at the end of a block, the model could continue to pull cases from the distribution of procedure times until it identifies one that fits; and then it could schedule the "discarded" cases in a later block. Or the model could include the utilization of block time by specialties outside their own blocks. However, potentially substantial, additional effort would be required to include these features in the model. Yet it is unlikely that the added complexity provided by these features would significantly improve the model's validity. It is important to remember that a simulation model is a representation of a system, not an exact replica. Efforts should be made to meet project objectives while keeping the model as simple as possible.

REFERENCES

Lowery J.C. 1996. Introduction to simulation in health care. In *Proceedings of the 1996 Winter Simulation Conference*, ed. J.M. Charnes, D.J. Morrice, D.T. Brunner, and J.J. Swain, 78-84. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Rinde A., Blakely T. 1976. *Operating room resource utilization: Chicago area survey findings and recommendations.* Chicago: Chicago Hospital Council.

Salt, J.D. 1993. Keynote address: Simulation should be easy and fun. In *Proceedings of the 1993 Winter Simulation Conference*, ed. G.W Evans, M. Mollaghasemi, E.C. Russell, and W.E. Biles, 1-5. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

AUTHOR BIOGRAPHIES

JULIE C. LOWERY is Associate Director of the Center for Practice Management and Outcomes Research, Department of Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, MI. She received her PhD in Health Services Organization and Policy from the School of Public Health, University of Michigan, where she specialized in operations research and information systems in health care. She has a BS in Microbiology and a Master's in Health Services Administration, also from the University of Michigan. Her research interests, in addition to simulation, include telemedicine and shared decision-making between patients and providers.

JENNIFER A. DAVIS is a research associate at the Center for Practice Management and Outcomes Research. She received her Master's in Health Services Administration and her BA in Psychology from the University of Michigan. In addition to developing simulation models, Ms. Davis manages databases and conducts statistical analyses for a variety of research projects on outcomes of medical care, including diabetes care and reconstructive surgery.