

A Behavioral Study of Daily Mean Turnover Times and First Case of the Day Start Tardiness

Jihan Wang, MS,* Franklin Dexter, MD, PhD,† and Kai Yang, PhD*

BACKGROUND: Previous research has identified 2 psychological biases in operating room (OR) decisions on the day of surgery: risk attitude of the decision-maker at the OR control desk and decisions made by OR staff to increase clinical work per unit time during the hours they are assigned. Resulting decisions are worse than random chance at reducing overutilized time. To isolate the second bias from decisions at the OR control desk, previous studies of the second bias have analyzed decisions made in non-OR locations and on nights/weekends. Another way to isolate the second bias from decisions at the OR control desk is to study facilities with negligible overutilized OR time. We examined the second bias using data from such a facility.

METHODS: One year of data was collected from a 5-OR hospital. Allocated OR time that minimized the inefficiency of use of OR time was determined first to confirm there was virtually no overutilized OR time. A structural equation model was then built to evaluate the relations among variables while controlling for other correlations. We tested the hypothesis that nonoperative times were no longer on days with little versus relatively large workload.

RESULTS: The extra ORs were not cost efficient (i.e., the mean potential improvement varied among days from $21.1\% \pm 0.2\%$ [SE] to $38.9\% \pm 0.2\%$), resulting in very little overutilized OR time. However, conditioned on the preceding tactical decision of running extra ORs, the allocated OR time during the studied period was that which minimized the inefficiency of use of OR time. As the preceding results showed that the facility was suitable for the behavioral study, the behavioral study was performed, and the hypothesized relation confirmed. Each 1-hour decrease in the daily estimated (total) duration of elective cases resulted in a managerially unimportant decrease in the mean turnover times (0.41 ± 0.21 minutes, $P = 0.053$). Excluding turnovers when there were >2 turnovers occurring simultaneously, there was no significant decrease (0.17 ± 0.24 minutes, $P = 0.464$) in the mean turnover times per each 1-hour decrease in the daily estimated (total) duration. Similarly, after excluding prolonged turnovers (>60 minutes), there was no significant decrease (0.16 ± 0.16 minutes, $P = 0.315$) in the mean turnover times per each 1-hour decrease in the daily estimated (total) duration.

CONCLUSIONS: Previous experimental and observational studies found many clinicians maintained high clinical work per unit time during the hours to which they were assigned. We tested and confirmed a prediction of this bias as was applied during regularly scheduled OR hours among an entire surgical team. Overall, the staff worked just as quickly on days with few or many hours of cases. The OR staff did not slow down, thus filling the time. These results have important implications for the cost utility of information technologies to facilitate managerial decision-making on the day of surgery. (Anesth Analg 2013;116:1333–41)

Operating room (OR) scheduling systems list allocated OR time by service. The allocated OR time is used months before the day of surgery during staff scheduling.^{1,2} In addition, starting a few (e.g., 2) workdays before the day of surgery,³ to the extent feasible, cases are not scheduled outside of allocated OR time when they can be performed in allocated OR time.^{1,4} Doing so reduces both the “overutilized OR time” (cases beyond the allocated time) and the “underutilized OR time” (allocated but unused time). As the day of surgery approaches, the incremental

cost of an hour of underutilized OR time becomes negligible relative to the cost of an hour of overutilized OR time (e.g., because the staff have already been scheduled at one facility or another). Consequently, decision-making on the day of surgery to reduce labor costs should be focused on reducing the hours of overutilized OR time.^{1,4,5}

Contemporary OR information systems include electronic displays (“whiteboards”) showing information about surgical case progress to facilitate OR managers’ decision-making on the day of surgery (e.g., to reduce overutilized OR time). Decisions involving multiple ORs that are made using passive electronic displays are significantly *worse* than random chance.⁶ However, it has been shown that displays with *recommendations* enhance decision-making,⁶ and education increases trust in these recommendations.^{7,8}

The lack of benefit of presenting information alone is caused by 2 psychological biases.^{6,9} The first bias relates to whether or not the decision-maker at the OR control desk is psychologically risk-averse or nonrisk-averse. The decisions for add-on case scheduling, moving cases between ORs, etc., differ depending on the risk attitude of the decision-maker.⁹ The risk-averse decision-maker aims to have all the

From the *Department of Industrial and Systems Engineering, Wayne State University, Detroit, Michigan; and †Division of Management Consulting, Department of Anesthesia, University of Iowa, Iowa City, Iowa.

Accepted for publication December 10, 2012.

Supported by departmental funds.

Conflicts of Interest: See Disclosures at the end of the article.

Reprints will not be available from the authors.

Address correspondence to Franklin Dexter, MD PhD, Division of Management Consulting, Department of Anesthesia, University of Iowa, 200 Hawkins Dr., 6-JCP, Iowa City, IA 52242. Address e-mail to Franklin-Dexter@UIowa.edu or www.FranklinDexter.net.

Copyright © 2013 International Anesthesia Research Society
DOI: 10.1213/ANE.0b013e3182841226

ORs finish within the scheduled (allocated) OR hours (i.e., little to no chance of overutilized OR time). Such a decision-maker is relatively intolerant of overutilized OR time and discounts the possibility of subsequently scheduled "add-on" cases (i.e., neglects the fact that the allocated OR time is based on the workload from all performed cases, including add-on cases).^a The nonrisk-averse decision-maker considers the potential add-on cases. Decisions made by nonrisk-averse decision-makers result in less overutilized OR time caused by cases added later in the day. The mean inefficiency in use of OR time is larger for the risk-averse OR coordinators than the nonrisk-averse coordinators based on a study conducted by Stepaniak et al.⁹

The second bias relates to the decisions made by OR staff. The staff typically make decisions based on increasing clinical work per unit time during the hours they are assigned.^{6,10,11} This heuristic (rule-of-thumb) is logical for decisions involving single ORs, because the heuristic serves to reduce the expected hours of overutilized OR time. It can seem like the behavior exemplifies a good work ethic (see Discussion). The problem is that when applied to decisions involving multiple ORs, the decisions are (highly) suboptimal. For example, suppose there was 1 empty postanesthesia care unit bed for 2 ORs. The bed often would go to the first exiting OR with 1 hour expected underutilized OR time rather than to the second OR expected to exit 10 minutes later but with >1 hour expected overutilized OR time. Such decisions unnecessarily result in increased net overutilized OR time.^{6,11,b} As demonstrated by previous studies, the behavior is unaffected by education¹¹; rather, education increases trust in recommendations provided by displays.^{7,8} The behavior also is unaffected by changes to cases' classifications of medical urgency.¹⁰ The bias may be sustained by physician perception of team activity as being favorable.^{12,13}

A limitation of the preceding studies^{6,11} is that most OR decisions are made during regularly scheduled (allocated) OR hours. However, the explanations for the anesthesiologists' decision-making behavior (i.e., the second bias) were studied in non-OR settings and in OR settings on nights and weekends.^{6,11} These periods were studied to isolate the behavior of clinicians throughout the surgical suite (i.e., the second bias) from the bias of the perioperative manager at the OR control desk with (or without) risk aversion (i.e., from the first bias).^{9,11} Additional research is warranted, because the work pace of service workers from patient transport services and cardiothoracic surgery was influenced by workload.¹⁴ Because of the second bias, OR staff's work pace in turnovers and first-case starts is hypothesized not to change significantly under different OR workloads. Simply stated, we hypothesized that OR staff do not slow down during turnovers and first-case starts on slow days nor speed up on busy days. As explained in the paragraph

below, we studied the OR staff's work pace during turnovers and at the beginning of the workday.

In the current study, we investigated a facility with 8 hours of allocated time in each OR, staff (e.g., anesthesiologists, certified registered nurse anesthetists, nurses, OR techs) scheduled for 8 hours, and negligible overutilized OR time (see below in Results).^c At such a facility, decisions at the OR control desk are not influenced by whether the decision-maker is risk-averse or nonrisk-averse, because there is very little chance of overutilized OR time and the staff working late.^{9,15} Consequently, the (second) bias of increasing clinical work per unit time during the hours to which each staff is assigned would be illustrated, overall, by OR staff maintaining a constant patient flow, regardless of the day's estimated (total) duration of elective cases (i.e., scheduled OR workload). Such behavior can be tested by estimating the correlations (1) between the scheduled OR workload and the mean turnover times, and (2) between the scheduled OR workload and the mean first case start tardiness. We hypothesized that the overall ensemble behavior of OR staff was decisions that would maintain clinical work per unit time, resulting in no managerially important correlations (e.g., on less busy days, if the mean turnover times were longer, then they would be so only by tiny amounts). If the hypothesis was proved true, then the second bias exists during regularly scheduled (allocated) OR hours. It would cause suboptimal decisions without recommendations from electronic displays. Thus, as above, the extent to which the hypothesis holds has implications regarding the value of providing electronic displays without evidence-based recommendations.

METHODS

The quality-improvement project was performed for the John D. Dingell VA Medical Center located in Detroit, Michigan.

The analysis was done in 2 phases. In the first phase, we checked that the allocated OR time during the studied date range exceeded that which minimized the inefficiency of use of OR time. The facility's use of 5 ORs for 8 hours should be less efficient than running fewer ORs for 8 hours (i.e., staff could reliably know that their workday would end in 8 hours regardless of their decisions). Satisfying this constraint was a prerequisite for the behavioral study, because otherwise decisions on the day of surgery to move cases or to change start times should have been based principally on reducing overutilized OR time.^{1,4,5} If the studied facility had extra allocated OR time, then there was little overutilized OR time; thus, there was virtually no decision at the control desk to reduce it (the first bias), and the explanation of OR staff's behavior would be the second bias. After this phase was completed, we studied the behavior of the

^a The risk-averse decision-maker moves cases and schedules add-on cases on the day of surgery based on Worst Fit Descending, which starts OR cases as early as possible. Conversely, the nonrisk-averse decision-maker uses Best Fit or Best Fit Descending to schedule OR cases.¹⁷

^b For another example, if an anesthesiologist has 2 ORs each with experienced certified registered nurse anesthetists ready to start, the anesthesiologist often would first do the quick adult induction in 1 OR with expected underutilized OR time and then do the longer pediatric induction with regional nerve block in the OR with expected overutilized OR time.¹¹

^c The OR nurses, surgical technologists, and certified registered nurse anesthetists were paid hourly. Most anesthesiologists were contract employees from neighboring institutes. When there was substantial underutilized OR time in all ORs, the OR manager would permit some OR staff to leave and the remainder of the OR staff would wait until the end of their workday in the expectation of add-on cases coming at the end of the day. When OR staff left early, they did not get paid for the hours that they were absent. However, OR staff were not required to leave early on the days with low workload if it was within the allocated OR time. We showed previously at 2 different hospitals, in non-OR and OR locations, that anesthesiologists did not vary their rate of clinical work based on whether they could leave when their last case of the day ended.¹¹

OR staff. Although the data were from a hospital, many of the surgical patients were outpatients. Furthermore, the OR conditions we studied are common for outpatient surgery centers. Thus, our results apply both to hospitals and to outpatient surgery centers. However, as explained above, the usefulness of testing our hypotheses is principally for hospitals with many ORs having >8 hours of cases (e.g., large general teaching hospitals).

Allocated OR Time

We used a full year of data, January 1, 2010 to December 31, 2010, with weekends and holidays excluded. We did not use a longer date range to avoid trends in OR workload over time and to avoid changes in the allocated OR time during the interval. Otherwise, the allocated OR time that minimized the inefficiency of use of OR time at the start of the period would be substantively different from the optimal allocated OR time later in the period.¹⁶ The use of the longest period possible gave us the advantage of the maximum possible sample size. This was important because our objective was to detect what we hypothesized to be small effects of OR workload on the mean turnover times and the mean first case start tardiness. It was a tradeoff between sample size and longer-term trends on OR workload.

The allocated OR time that minimized the inefficiency of use of OR time was determined separately for each day of week, because the OR workload varied among days of the week. During the 1-year period, there were 46 Mondays, 52 Tuesdays and Wednesdays, and 50 Thursdays and Fridays (holidays excluded).^d Five ORs were staffed on each workday for 8 hours, except for Wednesday with 1-hour late start for education. No ORs were allocated to individual specialties; rather, all specialties shared the time as if they were part of 1 large service. Nurses, anesthesia providers, etc., cared for patients of all specialties. After lists of cases had been scheduled, the lists were assigned whenever feasible to the designated OR(s) of the list's associated specialty. However, each specialty often did cases in other ORs. For each day, the OR workload (including turnover times) was grouped by surgeon, then the surgeons' lists of cases were scheduled into the ORs providing the earliest possible start time, thereby filling all first-case starts as if they were from a single service (i.e., surgeons' lists of cases were scheduled using Worst Fit Descending algorithm).^{4,17,18} Worst Fit matched the observed behavior of filling all the first-case starts.¹⁷ With 1 (typically) or 2 (occasionally) surgeons per OR per day and typically zero add-on cases, Worst Fit Descending minimized the expected inefficiency of use of OR time.⁴

Whenever a turnover time was longer than 90 minutes, we rounded the turnover time down to 90 minutes. We used 90 minutes as the maximum in part because it was the 90th percentile of the turnover times for the dataset.¹ There were 145 turnover times in our dataset longer than 90 minutes, just 10% of all turnovers. These turnover times longer than 90 minutes might be attributable to gaps in the OR schedule (e.g., cases

not scheduled sequentially).¹⁹ The prolonged turnover times might also be caused by reasons such as delays in housekeeping/setups. We substituted 90 minutes for these turnover times longer than 90 minutes because our study was of setup and cleanup times, and such (prolonged) turnover times are inconsistent a priori with minimizing the inefficiency of use of OR time on which the calculations were based.^{1,20}

The following calculation of allocated OR time was performed as described in previous studies.^{1,20} We explored the options of running 4 ORs, 5 ORs, and 6 ORs, each with the combination of 8 hours' and 10 hours' allocation for each day of week. Each surgeon's workload was assigned to an OR using the Worst Fit Descending scheduling algorithm (see above).^{17,18} For each combination of numbers of ORs and hours, we calculated the under- and overutilized OR time of each day. We then compared the inefficiency of use of OR time with the baseline inefficiency from the actual allocated OR time.¹⁵ The inefficiency of use of OR time was calculated as the daily sum of the underutilized OR time plus 1.75 times the overutilized OR time.^{1,16,20}

The mean potential improvement in the inefficiency of use of OR time and its standard errors were calculated by performing 1000 replications of cross-validation for each day of the week. For each replication, three-fourths of the days in the dataset were selected randomly and used as the training set and the other one-fourth as the testing set. The training set was used to determine the allocated OR time that minimized the inefficiency of use of OR time. Then, the best allocated OR time was applied to the testing dataset to calculate the inefficiency of use of OR time. This process was repeated 999 times. The analysis was conducted using MATLAB R2010a (The MathWorks Inc., Natick, MA).

Behavior of the OR Staff

The 3 slow days above (i.e., July 27, 2010, December 28, 2010, and December 6, 2010) were excluded from the behavioral portion of this study because no valid turnover times were available for those days. Thus, the mean turnover time could not be calculated for these 3 days. August 16, 2010 was also excluded, because the only turnover on that day was between an elective case and an add-on case. We did not include the 3.7% of turnover times that included an add-on case (47 of 1259) in the behavioral study because such cases were unexpected to the OR staff.^e The add-on cases would potentially have biased the calculation of mean turnover time, for reasons such as the shortage of supplies, preparation of equipment, and delays from other departments, which would make the turnover time longer than usual. Thus, the inclusion of such turnovers would have not correctly reflected OR staff's reaction to this unexpected workload. After excluding these 4 days, there were 246 days total for study.

^e The turnovers between elective and urgent cases were not excluded in the preceding first part of our analysis (i.e., the optimal allocated OR time). The reason for including the turnovers when calculating the OR allocations was that the turnover times were part of the OR workload. Excluding them would make the OR workload lower than what was actually observed, resulting in the optimal allocated OR time being smaller than needed. In contrast, for this second part of the analysis (i.e., mean turnover time calculation), the inclusion of such turnover times would have introduced outliers, causing the mean value to be inaccurate. We excluded these days to reduce the impact of potential outliers on our results.

^d We deliberately included 3 days with unusually low workload in the calculations of allocated OR time. Two workdays (July 27, 2010 and December 28, 2010) had no turnover times, because there was only 1 case performed in each OR. One workday (December 6, 2010) had just 1 turnover time, and it was longer than 90 minutes. Inclusion of these days (deliberately) biased results toward more overutilized OR time.

Table 1. Percentiles for Defined Variables in the Behavioral Study

	10%	25%	50%	75%	90%	Mean
Allocated operating room time (h)	35	40	40	40	40	39
Estimated (total) duration of elective cases (h)	12	15	20	24	27	19
Actual (total) duration of elective cases (h)	15	17	21	26	29	22
Estimated (total) duration of add-on cases (h)	0	0	0	2	3	1
Actual (total) duration of add-on cases (h)	0	0	0	1	3	1
Mean first case start tardiness (min)	0	2	5	12	17	7
Mean turnover times (min)	26	32	40	47	55	40

For each case, the data we used were OR, surgery date, time that patient entered the OR, time that patient left the OR, scheduled start time of the case, scheduled end time of the case, and elective (scheduled) or add-on (urgent). The phrase “add-on” referred both to the actual nonurgent add-on and to urgent cases. There were 4.4% add-on (urgent) cases (129 of 2906). From the procedure lists, most of these cases likely were known later in the workday (i.e., were unexpected). However, such temporal information could not be included in the behavioral analysis because (1) there were no reliable data available on when OR staff knew of each such case, and (2) there were too few such turnover times to have been able to detect any such effect. Importantly, all such cases were included in the preceding calculations of allocated OR time and hours of overutilized OR time.

For each workday, there was 1 independent variable: allocated OR time. There were also 6 correlated dependent variables: estimated duration of elective cases, actual duration of elective cases, estimated duration of add-on cases, actual duration of add-on cases, mean first case start tardiness, and mean turnover times. The first 4 were totals for all such cases during the workday. The latter 2 were means for all such cases during the day. The percentiles for the variables are summarized in Table 1.^f

The tardiness of each first case of the day start was calculated as the difference between the scheduled start time of the case and the time that the patient entered the OR.^{21,22} If the time the patient entered the OR was before the scheduled start time of the case (2.2% of first cases), the tardiness was considered to be zero.^{21,22}

The raw data had some missing information. For 1.0% of the cases (30 of 2906), there was no scheduled duration. We did not want to ignore these cases because they occurred on 10% of the workdays (25 of 246). We imputed the missing information from schedules before conducting the behavioral analysis.^g

^f The analyses could have been performed with the number of cases instead of with the estimated (total) duration of cases. By scatter plot and Pearson linear correlation coefficient ($r^2 = 0.72$), the 2 were interchangeable. To choose, we calculated the Kendall τ_b rank correlations with the mean first case start tardiness and mean turnover times. Because the correlations between the estimated (total) duration of elective cases and mean turnover times and mean first case start tardiness were larger than those with the number of cases, we used the estimated (total) duration of cases for our analysis. By doing so, we biased our results toward detecting a relationship between OR workload and turnover times (i.e., biased to be contrary to our hypothesis).

^g We evaluated cases scheduled in the same OR by the same surgeon, looking for a hole in the schedule between 2 cases of an appropriate duration for the procedure with known missing scheduled duration but known assigned OR. If so, we considered the scheduled duration to be the duration of the hole, minus the minutes of scheduled turnover times. This approach was sufficient for 22 of the 30 cases. The other 8 cases were either the only case performed in the OR or with gaps from the previous cases. To impute the scheduled duration, we used the actual case duration and rounded it up or down to the nearest 30-minute interval, because the cases were scheduled in 30-minute intervals.

As reported previously, mean turnover times were correlated among days,¹⁹ and this was built into the analysis. In 1999, Dexter et al.²³ used structural equation modeling to model variability in underutilized OR time. We modified the path diagram to better study the psychological biases. The final path diagram used for parameter estimation is shown in Figure 1. The following 5 paragraphs give a detailed description of the path diagram and the reasons for the path selections. The development of the path diagram was a step in the structural equation modeling used to test our hypotheses.

On Wednesdays, when less OR time was allocated, fewer elective cases and add-on cases were scheduled. This is represented in the path diagram by the arrow going out from “Allocated OR time” to “Estimated duration of elective cases” and to “Estimated duration of add-on cases.” These paths make sense, because the availability of OR time can influence how many hours of case are added to the day’s schedule.²³

When less OR time was allocated, it might matter whether first cases of the day start on time, because the OR staff might perceive that a delayed start of the day might cause the staff to finish their workday late.²⁴ Thus, a path from “Allocated OR time” to “Mean first case start tardiness” was added.

The estimated (total) duration of elective cases linearly influenced the actual (total) duration of elective cases in the model (Fig. 1).²³ Likewise, the estimated (total) duration of add-on cases influenced the actual (total) duration of add-on cases.²³

The Kendall τ_b rank correlations did not show a significant relation between the mean first case start tardiness and the mean turnover times, nor between the estimated (total) duration of elective cases and the estimated (total) duration of add-on cases (both $P > 0.28$). Because the studied facility had copious underutilized OR time, we did not include these paths.

The other paths included were for hypothesis testing (i.e., the estimated and actual duration of cases to the mean turnover times and the mean first case start tardiness). We had 2 paths from the estimated (total) durations of both elective and add-on cases to the mean turnover times. OR staff’s perceptions of the workload might have been based on the estimated (total) duration of cases. As the day proceeded, staff behavior might also have been affected by the actual (total) duration of cases. For example, when a case ended much earlier than planned, there was more than sufficient time for the staff to set up the OR for the next case. Thus, they might work at a more leisurely pace. However, when a case took longer than scheduled, the staff might speed up to reduce the tardiness of start of the next cases in the OR.

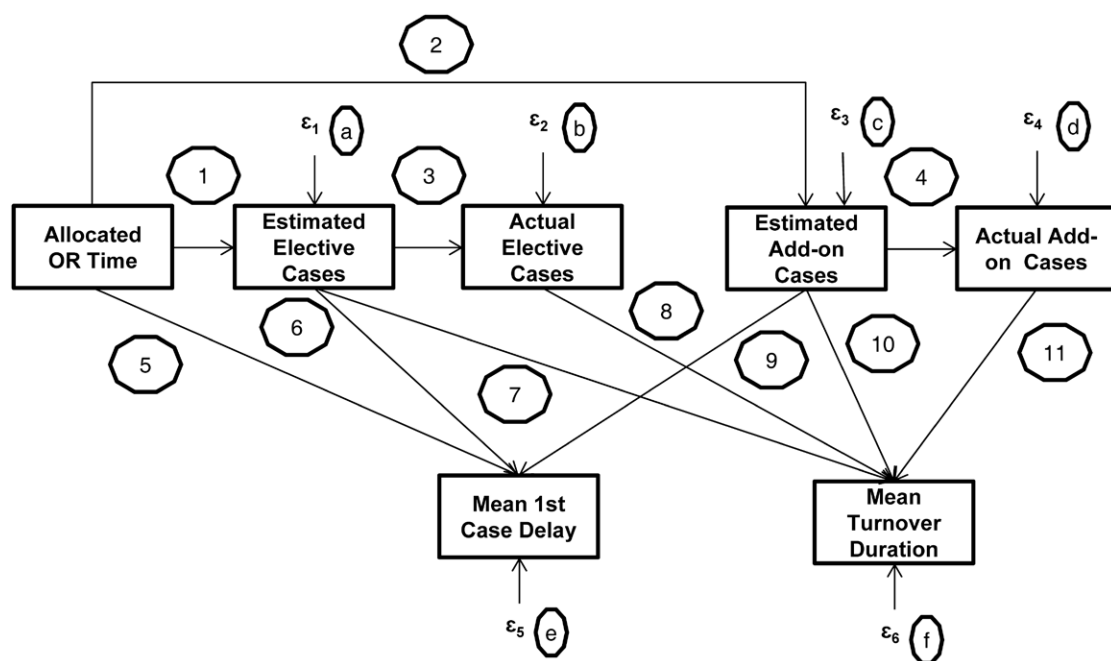


Figure 1. Path diagram of the structural equation modeling. The structural equation model estimates the values of 17 parameters: 11 path coefficients (indicated by circled numbers) and 6 variances of the random effects terms (indicated by circled alphabets). The single arrows represent hypothesized relationships among variables (i.e., they link the effects and causes). The box where each arrow starts is the cause variable, the change of which will change the effect variable where the arrow ends. OR = operating room.

The parameters of the structural equation model (Fig. 1) were estimated to emulate the observed covariance matrix among the defined variables.^{25,26} The distribution-free, weighted least-squares method was used for parameter estimation because several of the variables were not normally distributed.²⁷ LISREL 8.8 student version (Scientific Software International, Inc., Lincolnwood, IL) was used to generate the covariance matrix and estimate parameters.

Goodness of fit of the statistical model was assessed several ways, as recommended for structural equation models.²⁶ The χ^2 test along with its *P* value, root mean square error of approximation, and goodness-of-fit index (GFI)/adjusted goodness-of-fit index (AGFI) were the methods that we used to evaluate the model performance.²⁶ The χ^2 test and its *P* value indicate the differences between the observed and implied covariance matrices. Root mean square error of approximation estimates lack of fit compared with the saturated model (perfect model). The GFI/AGFI are similar to the R^2 /adjusted R^2 in regression. They measure the amount of variance in the observed covariance matrices predicted by the reproduced matrices. Criteria for a good fit based on each statistic are summarized in the Results along with the study findings.

From 2008 through the studied year, the studied facility had no ongoing quality-improvement or monitoring program focused on reducing turnover times. However, there was a project focused on reducing first case start tardiness. The latter could negatively affect the generalizability of our findings to other facilities. However, there was no public incentive or financial reward for individuals or OR groups, nor measurement of individual performance, making the influence limited. Nonetheless, the estimated relationships for the mean first case start tardiness must be considered secondary to those for turnover time.

Besides the primary analysis, we ran 2 sensitivity analyses. There were only 2 housekeeping teams at the studied facility. When there were >2 cases finishing close to simultaneously, turnover(s) was delayed, waiting for the next available housekeeping team. This sharing of resources made the turnovers longer.²⁸ To isolate influences of >2 simultaneous turnovers, we first recalculated the mean turnover times, excluding turnovers including any 1-minute period with >2 simultaneous turnovers. Second, instead of excluding turnover times longer than 90 minutes for the calculation of the daily mean turnover time, we excluded those turnovers longer than 60 minutes.

RESULTS

The allocated OR time was significantly longer than that which minimized the inefficiency of use of OR time (Table 2), satisfying the condition for the behavioral study. Although 5 ORs were open daily, 4 ORs were sufficient (see Discussion), resulting in very little overutilized OR time each day of the week (from 0.01 ± 0.03 [SE] hours to 0.06 ± 0.24 hours per OR per workday) (Table 2). Thus, there was no need for decisions from the OR control desk to reduce overutilized OR time, and any such decisions made would not influence the (negligible) overutilized OR time. Given that the necessary condition was met, the findings of the behavioral study were related to the bias of OR staff to increase clinical work per unit time during the hours they were assigned.

The structural equation model demonstrated reasonable fit based on 4 criteria: χ^2 test *P* = 0.470, root mean square error of approximation 0.01 (unitless), goodness-of-fit index 0.99, and adjusted goodness-of-fit index 0.98.^h The coefficient parameter estimates and their *P* values are given in Table 3.

A decrease by 1 hour in the estimated (total) duration of elective cases caused the mean turnover times to *decrease*

Table 2. Difference in the Inefficiency of Use of Operating Room (OR) Time Between Actual and Allocated Hours Minimizing the Inefficiency of Use of OR Time

Allocation of OR time		Difference in the inefficiency of use of OR time, mean $\pm$ SE (%)	Overutilized OR time per OR per day while applying the optimal OR allocation, mean $\pm$ SE (h)
Day of week	Actual allocation of OR time		
Monday	5 ORs $\times$ 8 h	21.11 $\pm$ 0.24	0.06 $\pm$ 0.24
Tuesday	5 ORs $\times$ 8 h	37.87 $\pm$ 0.10	<0.01
Wednesday	5 ORs $\times$ 7 h	33.58 $\pm$ 0.24	0.03 $\pm$ 0.16
Thursday	5 ORs $\times$ 8 h	34.99 $\pm$ 0.27	0.01 $\pm$ 0.03
Friday	5 ORs $\times$ 8 h	38.87 $\pm$ 0.24	0.01 $\pm$ 0.07

The actual OR allocation was what had been implemented at the studied facility. The number before "ORs" represents number of ORs open for each day of week. The number before "h" is the staffing for each OR. The optimal OR allocation was calculated as described in the Methods section to achieve the minimum expected inefficiency of use of OR time.

Table 3. Outputs of Primary Structural Equation Modeling

		ORAll	EstELEC	EstADD	ActELEC	ActADD	1ST	TOT
ORAll	Coefficient	—	0.6932	0.0141	—	—	0.0047	—
	SE	—	0.1218	0.0238	—	—	0.0025	—
	P value	—	<0.0001	0.5520	—	—	0.0570	—
EstELEC	Coefficient	—	—	—	0.8785	—	0.0033	0.0068
	SE	—	—	—	0.0382	—	0.0013	0.0035
	P value	—	—	—	<0.0001	—	0.0084	0.0531
EstADD	Coefficient	—	—	—	—	0.7959	0.0014	0.0209
	SE	—	—	—	—	0.0371	0.0066	0.0296
	P value	—	—	—	—	<0.0001	0.8338	0.4793
ActELEC	Coefficient	—	—	—	—	—	—	0.0031
	SE	—	—	—	—	—	—	0.0034
	P value	—	—	—	—	—	—	0.3689
ActADD	Coefficient	—	—	—	—	—	—	-0.0053
	SE	—	—	—	—	—	—	0.0365
	P value	—	—	—	—	—	—	0.8853

For each path, the result is presented in 3 rows. The first row is the coefficient point estimate. The second row is the standard error of the coefficient, and the last row contains the P values.

ORAll = operating room allocation; EstELEC = estimated elective case duration; EstADD = estimated add-on case duration; ActELEC = actual elective case duration; ActADD = actual add-on case duration; 1ST = mean first case start tardiness; TOT = mean turnover time.

by 0.41 ± 0.21 minutes ($P = 0.053$) (Table 3). However, these changes to the mean turnover times were trivially small, which confirmed our hypothesis that there were negligible differences in OR staff's work pace during turnovers among different OR workloads (i.e., the staff maintained their clinical work per unit time). On the other hand, the association was close to statistically significant and in the opposite direction of what might be expected behaviorally. Decreases in the actual (total) duration of cases, either of elective cases or of add-on cases, did not result in significant decreases (changes) in the mean turnover times ($P = 0.369$ and $P = 0.885$, respectively, Table 3).ⁱ Furthermore,

2 sensitivity analyses' results matched the finding of no managerially relevant *increase* from decreases in the estimated (total) duration of elective cases. Observed decreases in the mean turnover times in these 2 analyses were 0.17 ± 0.24 minutes ($P = 0.464$) and 0.16 ± 0.16 minutes ($P = 0.315$), respectively (Table 4).^j

Each 1-hour decrease in the estimated (total) duration of elective cases caused the mean first case start tardiness to *decrease* by 0.2 ± 0.1 minutes ($P = 0.008$). Again, the difference was so small as to be practically unnoticeable (12 seconds), confirming our hypothesis. However, the difference was statistically significant and in the opposite direction expected behaviorally. At the studied facility, 69% of the workdays (170 of 246) were of 5 ORs (i.e., first cases of the day). The mean estimated (total) duration of elective cases on the days of 5 first cases was 21.8 ± 0.4 hours versus 16.1 ± 0.5 hours for the days of 4 first cases. There were fewer numbers of first cases of the day when there was less workload. Thus, there was no fifth OR to wait for an anesthesiologist with 2 preceding inductions, resulting in smaller mean tardiness.²⁹

^h A small P value suggests a poor model fit.²⁶ A value of ≤ 0.05 indicates a good fit, and a value of 0.08 represents an acceptable model.^{23,26} AGFI/AGFI that is > 0.95 suggests a good model.²⁶

ⁱ In our structural equation modeling, add-on case factors were included (Table 3), and none of them was statistically significant with respect to the mean turnover time and the mean first case start tardiness. Results were insensitive to the heterogeneity among days in the hours of add-on cases (i.e., OR staff behavior was not significantly influenced by the add-on cases). One likely reason is that most turnovers take place during the morning,²⁸ whereas most add-on cases are performed during the afternoon. Another reason is that for surgical suites with allocated OR time that minimizes the inefficiency of use of OR time, the substantial OR workload from add-on cases has already been included in the allocated OR time; thus, the possibility of substantial overutilized OR time caused by add-on cases is not high. At the studied facility with an extra first case of the day start, this was even more so.

^j In total, there were 178 turnovers between 60 and 90 minutes (not including 60 and 90 minutes), among which 118 (66%) overlapped with at least another 2 turnovers (i.e., likely waiting for housekeeping teams).

Table 4. Key Outputs of the Sensitivity Analyses of the Structural Equation Modeling

		Exclude simultaneous TOTs >2		Exclude TOTs >60 min	
		1ST	TOT	1ST	TOT
EstELEC	Coefficient	0.0033	0.0029	0.0032	0.0027
	SE	0.0012	0.0040	0.0013	0.0027
	P value	0.0074	0.4639	0.0146	0.3149
EstADD	Coefficient	0.0018	0.0208	0.0027	0.0063
	SE	0.0067	0.0356	0.0068	0.0211
	P value	0.7881	0.5587	0.6931	0.7660
ActELEC	Coefficient	—	0.0038	—	0.0036
	SE	—	0.0038	—	0.0030
	P value	—	0.3192	—	0.2253
ActADD	Coefficient	—	−0.0056	—	0.0111
	SE	—	0.0429	—	0.0285
	P value	—	0.8964	—	0.6959

The left 2 columns include coefficients of the sensitivity analysis that excluded the turnover times when there were >2 simultaneous turnovers. The right 2 columns are for the sensitivity analysis that excluded prolonged turnovers between 60 and 90 minutes.

1ST = mean first case start tardiness; TOT = mean turnover time; EstELEC = estimated elective case duration; EstADD = estimated add-on case duration; ActELEC = actual elective case duration; ActADD = actual add-on case duration.

DISCUSSION

Allocated OR Time

To achieve a reduction in costs, the management of the studied facility could have run 1 fewer OR daily and changed allocated OR time.^{1,20} The way to optimize OR cost efficiency was not to reduce turnover times or ensure on-time start of the workday, but to reallocate OR hours and to reduce the hours of underutilized OR time.¹ However, randomized clinical trials have shown that when there are 4 ORs with different surgeons each with at least 8 hours of cases, productivity can be increased by running 5 ORs.^{30,31} When running 3 ORs with >8 hours of cases, running 4 ORs also can increase productivity.³² Furthermore, running more ORs increases first-case starts, which reduces tardiness from scheduled start times for surgeons.^{7,21,22} Therefore, although the choice of 5 ORs reduced the efficiency of use of OR time, the tactical decision of running an extra OR might have been rational. Regardless, the number of ORs open was a tactical decision made before the OR time was allocated, and was thus incorporated in our calculation of allocated OR time.^{1,33}

Behavior of OR Staff

We studied the overall (ensemble) behavior of OR staff at a facility with virtually no overutilized OR time. As hypothesized, the mean turnover times were negligibly affected by the estimated (total) duration of elective cases. OR staff kept a constant work pace for nonoperative times, except for a slight slowing when there were >2 simultaneous turnovers. The staff overall did not slow down to fill the time when less busy. This negative finding was not caused by there being only small differences in the estimated (total) duration of elective cases among days, because this variable varied markedly (e.g., 10th percentiles 12 hours versus 90th percentiles 27 hours; Table 1). Rather, the behavior was consistent with the bias of most staff to maintain their clinical work per unit time during the hours to which they were assigned.^{6,11} As summarized in the introductory text, knowing that this bias applies commonly overall among many individuals is important for different facilities with some underutilized and some overutilized time, because the consequence is that electronic displays providing

information without evidence-based recommendations will result in decisions that are *worse* than decisions made at random.^{6,22,24}

The fact that the study was made of the overall behavior of the population (community) of OR nurses, anesthesiologists, etc., is important because many managerial decisions are distributed spatially and temporally. Changing the behavior involves the use of multiple displays (e.g., electronic whiteboards and pagers). Our results likely do not depend on the type of facility studied, because the psychological bias has been observed in other facilities in previous studies.^{10,11} Our new results confirmed that the bias occurs during regularly scheduled OR hours.

Sensitivity Analyses

The 2 sensitivity analyses focused on the managerially irrelevant, but close to statistically significant ($P = 0.053$), change in the mean turnover times. As described previously,²⁸ we calculated the percentage of days with >2 simultaneous turnovers. On days with more cases scheduled (i.e., more turnovers), there were more resulting delays in turnovers because there were only 2 available housekeeping teams (Fig. 2). Waiting for an available housekeeping team made the turnover times longer than usual.²⁸ When we eliminated the influence of the simultaneous turnovers, the impact of the estimated (total) duration of elective cases on the mean turnover times became even less. As hypothesized, from another perspective, OR staff's work pace was not influenced by OR workload.

Good Work Ethic

Some readers may have considered it obvious from their experience that most clinicians would work each day nonstop, rather than slowing down when there were fewer cases. For example, in a Swedish qualitative research study, perceptions of efficiency included "doing what must be done to achieve good workflow" and "working with preserved quality of care as fast as possible."³⁴ Readers may consider the possibility that the constant work pace would be the result of OR managers creating a culture that contributes to a good work ethic. Such prior expectations reinforce the strength of our statistical findings. The perception likely is

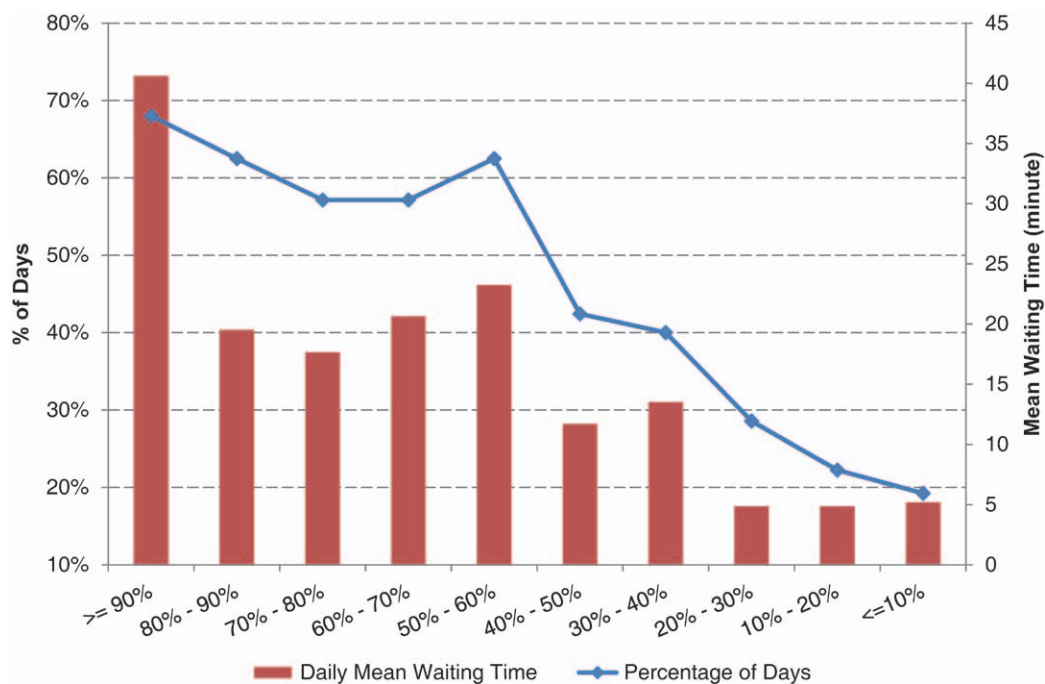


Figure 2. Percentage of days with >2 simultaneous turnovers and mean waiting time from such turnovers for each percentile of the estimated (total) duration of elective cases. The horizontal axis shows the percentiles of the estimated (total) durations of elective cases for the day. The left vertical axis shows the percentage of days with >2 simultaneous turnovers, with the lines connecting values for each percentile. The right vertical axis shows the daily mean wait time due to >2 simultaneous turnovers, and the columns represent values for each percentile.

correct for staff working inside individual ORs. However, the behavior results in increased patient and surgeon waiting on nights and weekends because anesthesiologists do not run all of the allocated ORs.^{10,11,35} The behavior results in increased overutilized OR time during regular workdays because the workforce is distributed equitably among ORs, with attention given to all ORs rather than preferentially to those ORs expected to have overutilized time.⁶ Such prior expectations also highlight the striking importance of the behavior's consequences. First, paying anesthesiologists substantially more to work late does not result in their working slower and thereby increasing compensation (see footnote c).³⁶ Making a similar observation from a different hospital was important. Second, for a facility like the one studied with little or no overutilized OR time, the use of electronic information displays throughout a surgical suite would not provide any return on investment. That would not be true if OR staff often increased nonoperative time on days with fewer estimated (total) duration of cases. Third, for a different facility with >8 hours of cases regularly in some ORs and different hours of cases performed in different ORs, the behavior is (highly) suboptimal without recommendations.³⁷ Again, the reason is that although clinicians' behavior to increase clinical work per unit time when assigned is reasonable for decisions involving 1 OR at a time, that does not apply to decisions involving multiple ORs (e.g., anesthesiologists supervising several nurse anesthetists, moving of cases, and scheduling of add-on cases).^{6,11,22,24}

In summary, we explored the OR staff's behavior at a facility with allocated OR time that was optimal conditioned on a predetermined number of ORs. Overutilized OR time was rare. The staff was scheduled to work for at least 8 hours daily. Given such a system, staff behavior did not respond

to the change in workload. The staff did not increase nonoperative time on days with fewer scheduled hours of cases. The OR staff worked as diligently on light days as on heavy days. The results show that the predicted psychological bias that OR staff work overall to increase clinical work per unit time during the hours they are assigned applies also during regularly scheduled OR hours.

RECUSE NOTE

Dr. Franklin Dexter is the Statistical Editor and Section Editor for Economics, Education, and Policy for the Journal. This manuscript was handled by Dr. Steven L. Shafer, Editor-in-Chief, and Dr. Dexter was not involved in any way with the editorial process or decision. ■

DISCLOSURES

Name: Jihan Wang, MS.

Contribution: This author helped design the study, conduct the study, collect and analyze the data, and write the manuscript.

Attestation: Jihan Wang approved the final manuscript. Jihan Wang reviewed the original study data and data analysis. Jihan Wang is the archival author.

Conflicts of Interest: The author has no conflicts of interest to declare.

Name: Franklin Dexter, MD, PhD.

Contribution: This author helped design the study, conduct the study, analyze the data, and write the manuscript.

Attestation: Franklin Dexter has approved the final manuscript. Franklin Dexter reviewed the data analysis.

Conflicts of Interest: The University of Iowa, Department of Anesthesia, Division of Management Consulting, assisted

the analyses described in this article for hospitals. Franklin Dexter has tenure and receives no funds personally from such activities. Income from the Division's consulting work is used to fund its research.

Name: Kai Yang, PhD.

Contribution: This author helped analyze the data and write the manuscript.

Attestation: Kai Yang approved the final manuscript. Kai Yang reviewed the original study data and data analysis.

Conflicts of Interest: The author has no conflicts of interest to declare.

REFERENCES

- McIntosh C, Dexter F, Epstein RH. The impact of service-specific staffing, case scheduling, turnovers, and first-case starts on anesthesia group and operating room productivity: a tutorial using data from an Australian hospital. *Anesth Analg* 2006;103:1499–516
- Dexter F, Wachtel RE, Epstein RH, Ledolter J, Todd MM. Analysis of operating room allocations to optimize scheduling of specialty rotations for anesthesia trainees. *Anesth Analg* 2010;111:520–4
- Dexter F, Shi P, Epstein RH. Descriptive study of case scheduling and cancellations within 1 week of the day of surgery. *Anesth Analg* 2012;115:1188–95
- Dexter F, Traub RD. How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. *Anesth Analg* 2002;94:933–42
- Dexter F, Epstein RH, Traub RD, Xiao Y. Making management decisions on the day of surgery based on operating room efficiency and patient waiting times. *Anesthesiology* 2004;101:1444–53
- Dexter F, Willemsen-Dunlap A, Lee JD. Operating room managerial decision-making on the day of surgery with and without computer recommendations and status display. *Anesth Analg* 2007;105:419–29
- Wachtel RE, Dexter F. Curriculum providing cognitive knowledge and problem-solving skills for anesthesia systems-based practice. *J Grad Med Educ* 2010;2:624–32
- Wachtel RE, Dexter F. Review article: review of behavioral operations experimental studies of newsvendor problems for operating room management. *Anesth Analg* 2010;110:1698–710
- Stepaniak PS, Mannaerts GH, de Quelerij M, de Vries G. The effect of the Operating Room Coordinator's risk appreciation on operating room efficiency. *Anesth Analg* 2009;108:1249–56
- Ledolter J, Dexter F, Wachtel RE. Control chart monitoring of the numbers of cases waiting when anesthesiologists do not bring in members of call team. *Anesth Analg* 2010;111:196–203
- Dexter F, Lee JD, Dow AJ, Lubarsky DA. A psychological basis for anesthesiologists' operating room managerial decision-making on the day of surgery. *Anesth Analg* 2007;105:430–4
- Shapiro LE, Alfille PH, Sandberg WS. Robots with a social memory. *Anesth Analg* 2010;111:19–20
- Masursky D, Dexter F, Isaacson SA, Nussmeier NA. Surgeons' and anesthesiologists' perceptions of turnover times. *Anesth Analg* 2011;112:440–4
- Kc DS, Terwiesch C. Impact of workload on service time and patient safety: an econometric analysis of hospital operations. *Manag Sci* 2009;55:1486–98
- Strum DP, Vargas LG, May JH, Bashein G. Surgical suite utilization and capacity planning: a minimal cost analysis model. *J Med Syst* 1997;21:309–22
- Epstein RH, Dexter F. Statistical power analysis to estimate how many months of data are required to identify operating room staffing solutions to reduce labor costs and increase productivity. *Anesth Analg* 2002;94:640–3
- Dexter F, Macario A, Traub RD. Which algorithm for scheduling add-on elective cases maximizes operating room utilization? Use of bin packing algorithms and fuzzy constraints in operating room management. *Anesthesiology* 1999;91:1491–500
- Galambos G, Woeginger GJ. The on-line bin packing: a restricted survey. *Math Methods Oper Res* 1995;42:25–45
- Dexter F, Epstein RH, Marcon E, Ledolter J. Estimating the incidence of prolonged turnover times and delays by time of day. *Anesthesiology* 2005;102:1242–8
- Dexter F, Epstein RH, Marsh HM. A statistical analysis of weekday operating room anesthesia group staffing costs at nine independently managed surgical suites. *Anesth Analg* 2001;92:1493–8
- Wachtel RE, Dexter F. Influence of the operating room schedule on tardiness from scheduled start times. *Anesth Analg* 2009;108:1889–901
- Wachtel RE, Dexter F. Reducing tardiness from scheduled start times by making adjustments to the operating room schedule. *Anesth Analg* 2009;108:1902–9
- Dexter F, Macario A, Lubarsky DA, Burns DD. Statistical method to evaluate management strategies to decrease variability in operating room utilization: application of linear statistical modeling and Monte-Carlo simulation to operating room management. *Anesthesiology* 1999;91:262–74
- Dexter EU, Dexter F, Masursky D, Garver MP, Nussmeier NA. Both bias and lack of knowledge influence organizational focus on first case of the day starts. *Anesth Analg* 2009;108:1257–61
- Raykov T, Marcoulides GA. *A First Course in Structural Equation Modeling*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000
- Schumacker RE, Lomax RG. *A Beginner's Guide to Structural Equation Modeling*. 3rd ed. Danvers, MA: Taylor & Francis Group, 2010:74–7
- Lomax RG. Covariance structure analysis: extensions and developments. In: Thompson B, ed. Vol 1. *Advances in Social Science Methodology*. Greenwich, CT: JAI, 1989:171–204
- Dexter F, Marcon E, Aker J, Epstein RH. Numbers of simultaneous turnovers calculated from anesthesia or operating room information management system data. *Anesth Analg* 2009;109:900–5
- Epstein RH, Dexter F. Influence of supervision ratios by anesthesiologists on first-case starts and critical portions of anesthesiology. *Anesthesiology* 2012;116:683–91
- Marjamaa RA, Torkki PM, Hirvensalo EJ, Kirvelä OA. What is the best workflow for an operating room? A simulation study of five scenarios. *Health Care Manag Sci* 2009;12:142–6
- Torkki PM, Marjamaa RA, Torkki MI, Kallio PE, Kirvelä OA. Use of anesthesia induction rooms can increase the number of urgent orthopedic cases completed within 7 hours. *Anesthesiology* 2005;103:401–5
- Hanss R, Buttgerit B, Tonner PH, Bein B, Schleppers A, Steinfath M, Scholz J, Bauer M. Overlapping induction of anesthesia: an analysis of benefits and costs. *Anesthesiology* 2005;103:391–400
- Dexter F, Macario A. Changing allocations of operating room time from a system based on historical utilization to one where the aim is to schedule as many surgical cases as possible. *Anesth Analg* 2002;94:1272–9
- Arakelian E, Gunningberg L, Larsson J. How operating room efficiency is understood in a surgical team: a qualitative study. *Int J Qual Health Care* 2011;23:100–6
- Dexter F, Epstein RH, Marsh HM. Costs and risks of weekend anesthesia staffing at 6 independently managed surgical suites. *AANA J* 2002;70:377–81
- Masursky D, Dexter F, Garver MP, Nussmeier NA. Incentive payments to academic anesthesiologists for late afternoon work did not influence turnover times. *Anesth Analg* 2009;108:1622–6
- Marcon E, Dexter F. An observational study of surgeons' sequencing of cases and its impact on postanesthesia care unit and holding area staffing requirements at hospitals. *Anesth Analg* 2007;105:119–26