

# Estimating the Incidence of Prolonged Turnover Times and Delays by Time of Day

Franklin Dexter, M.D., Ph.D.,\* Richard H. Epstein, M.D.,† Eric Marcon, Ph.D.,‡ Johannes Ledolter, Ph.D.§

**Background:** Prolonged turnover times cause frustration and can thereby reduce professional satisfaction and the workload surgeons bring to a hospital.

**Methods:** The authors analyzed 1 yr of operating room information system data from two academic, tertiary hospitals and Monte-Carlo simulations of a 15–operating room hospital surgical suite.

**Results:** Confidence interval widths for the mean turnover times at the hospitals were negligible when compared with the variation in sample mean turnover times among 31 hospitals. The authors developed a statistical method to estimate the proportion of all turnovers that were prolonged (> 15 min beyond mean) and that occurred during specified hours of the day. Confidence intervals for the proportions corrected for the effect of multiple comparisons. Statistical assumptions were satisfied at the two studied hospitals. The confidence intervals achieved family-wise type I error rates accurate to within 0.5% when applied to between five and nineteen 4-week periods of data. The diurnal pattern in the proportions of all turnovers that were prolonged provided different, more managerially relevant information than the time course throughout the day in the percentage of turnovers at each hour that were prolonged.

**Conclusions:** Benchmarking sample mean turnover times among hospitals, without the use of confidence intervals, can be valid and useful. The authors successfully developed and validated a statistical method to estimate the percentage of turnover times at a surgical suite that are prolonged and occur at specified times of the day. Managers can target their quality improvement efforts on times of the day with the largest percentages of prolonged turnovers.

This article is featured in "This Month in Anesthesiology." Please see this issue of ANESTHESIOLOGY, page 5A.

\* Director of the Division of Management Consulting, Associate Professor, Departments of Anesthesia and Health Management and Policy, § John F. Murray Professor of Management Sciences, College of Business, The University of Iowa. † Professor, Department of Anesthesiology, Jefferson Medical College, Philadelphia, Pennsylvania. ‡ President, Medical Data Applications, Ltd., Jenkintown Philadelphia, Pennsylvania. ‡ Associate Professor, Department of Industrial Maintenance, Jean Monnet University, Roanne Cedex, France.

Received from the Division of Management Consulting, Department of Anesthesia, The University of Iowa, Iowa City, Iowa. Submitted for publication December 2, 2004. Accepted for publication February 14, 2005. Support was provided solely from institutional and/or departmental sources. Presented at the Meeting of the American Association of Clinical Directors, Las Vegas, Nevada, October 24, 2004, and the Annual Meeting of the American Society of Anesthesiologists, Las Vegas, Nevada, October 27, 2004. Dr. Epstein is President of Medical Data Applications, Ltd., Jenkintown, Pennsylvania, which developed and markets CalculatOR software for operating room efficiency and staffing analyses. Dr. Dexter is Director of the Division of Management Consulting, which is a Division of the Department of Anesthesia. He receives no funds personally other than his salary from the State of Iowa, including no travel expenses or honorarium, and has tenure with no incentive program. After completion of this research, the methodology was added to the operating room efficiency consultation provided by the Department of Anesthesia. Steven L. Shafer, M.D., acted as Handling Editor for this article.

Address reprint requests to Dr. Dexter: Division of Management Consulting, Department of Anesthesia, The University of Iowa, Iowa City, Iowa 52242. Address electronic mail to: franklin-dexter@uiowa.edu. Individual article reprints may be purchased through the Journal Web site, www.anesthesiology.org.

OPERATING room (OR) information system data can be used to predict the impact of reducing turnover times on OR and anesthesia group staffing costs.<sup>1</sup> Although the direct effect of reducing turnover times on revenue is negligible at many hospitals,<sup>1</sup> indirect effects on revenue may be large. Long turnover times frustrate anesthesiologists and surgeons waiting to provide patient care,<sup>2</sup> may reduce professional satisfaction, and may reduce surgical workload if surgeons have a choice of facilities at which to do their cases. In addition, the perception of prolonged turnover times by surgeons, anesthesiologists, and administrators can result in substantial organizational costs resulting from multiple meetings and assessments of workflow. We evaluated the validity and usefulness of comparing (benchmarking) mean turnover times, without the use of confidence intervals, among hospitals.

Hospitals can reduce their incidence of prolonged turnovers.<sup>3,4</sup> Interventions to reduce prolonged turnovers can involve changing work hours of existing staff or the use of additional personnel. For example, between consecutive abdominal aortic aneurysm resections, an additional housekeeper can be assigned to help clean while an additional nurse assists anesthesia providers in preparing necessary intravenous fluids and supplies. If money were available for personnel who usually work elsewhere to be available part-time for the surgical suite, the time of day for them to be available should be when most prolonged turnovers occur. We describe and validate a statistical method to estimate the percentage of turnover times that are prolonged and that occur at specified hours of the day. The period of the day for which this percentage is the largest should be targeted for managerial improvement. We investigated whether this period can differ from the hours of the day with the largest percentage of turnovers that are prolonged.

## Materials and Methods

The OR room number and the date and times at which the patient undergoing elective surgery entered and left the OR were obtained for regular workdays in 2003 at two academic, tertiary surgical suites in the United States. Turnover times were considered to be the time from when one patient exited an OR until the next patient, if present, entered the same OR on the same day.<sup>5</sup>

**Table 1. Turnovers and/or Delays That Are Prolonged at Hospital A**

Hour-long Period of the Day	No. of Prolonged Turnovers*	% of All Turnovers during the 4-Week Period That Are Prolonged for Each Hourly Interval between 7 AM and 6 PM		
		Mean† of the 13 %s	SD of the 13 %s	99.38%‡ CI
7-7:59	2	0.0%		
8-8:59	5	0.1%		
9-9:59	83	0.9%	0.4%	0.5-1.3%
10-10:59	242	2.6%	0.9%	1.8-3.5%
11-11:59	279	3.0%	0.6%	2.5-3.5%
12-12:59	368	3.9%	0.6%	3.4-4.4%
13-13:59	441	4.7%	0.7%	4.1-5.3%
14-14:59	326	3.5%	0.9%	2.7-4.3%
15-15:59	198	2.1%	0.5%	1.6-2.6%
16-16:59	122	1.3%	0.5%	0.9-1.7%
17-17:59	0	0.0%		

\* The lower one-sided 95% confidence bound on the number of prolonged turnovers was at least 2 turnovers every 4 weeks provided there were at least 36 prolonged turnovers observed in thirteen 4-week periods. This lower bound was calculated from the second of the four steps of the method. The 7-7:59, 8-8:59, and 17-17:59 hour-long periods were not analyzed further because too few prolonged turnovers were observed (see second and fourth steps of the statistical method). † There were 9,344 turnovers during the thirteen 4-week periods studied. This table was created from the data from Hospital A. ‡ The 99.38% value equals 95% with Bonferroni correction for the multiple comparisons,<sup>7</sup> where 99.38% = [100 - 5 ÷ (8 comparisons)]%. The confidence interval (CI) was calculated using the Student *t* distribution. For example, the upper bound equals Mean + 3.308 × (SD) ÷ √13, where 3.308 is the *t* value with 13-1 degrees of freedom.

*Comparing (Benchmarking) Mean Turnover Times among Hospitals*

When estimating the mean turnover time and correlations among successive turnovers, values longer than 90 min were excluded because these typically included gaps in the OR schedule due to nonsequential case scheduling, not just cleanup and setup times.<sup>1</sup> Even for academic, tertiary surgical suites with reputations for slow turnover times, 90 min is more than 3 SDs longer than the mean (see Results). OR allocation and staffing analyses must consider nonsequential case scheduling, whereas benchmarking studies assess cleanup and setup times (including associated patient events, such as transport to the OR). Turnover times longer than a cut point (e.g., 90 min at tertiary surgical suite or 60 min at free-standing, outpatient facility) often are set equal to the cut point for purposes of OR allocation and staffing<sup>1</sup> but are excluded for purposes of benchmarking. For example, suppose that the first elective case of the day ended at 9 AM, and the second started at 2 PM. The calculated turnover time was 5 h. For purposes of assessing the impact of turnover times on OR allocation and staffing,<sup>1</sup> a value of 90 min could be used. For purposes of benchmarking in this article, we exclude large outliers such as the 5-h value. We do not know what portion of the 5-h period represented cleanup and setup. In contrast to the calculation of mean turnover time, no turnovers were excluded when determining the percentage of turnovers that were prolonged and occurred at an hour of the day. Calculation of the percentages was robust to the influence of outliers, unlike the calculation of the sample mean. The influences of both cleanup/setup times and nonsequential case scheduling are quantified in the percentages of turnovers that are prolonged. Managers must be able to compare the percentages to staffing by hour of

the day, because anesthesia providers are experiencing long periods between patient care and revenue.

The Runs Test was used to assess serial correlation among turnovers sorted by the date and time at which each turnover began. The Lilliefors test was used to test for normal distributions. Tests were performed using StatXact-6 (Cytel Software Corp., Cambridge, MA). *P* values were calculated using exact methods. Sample mean turnover times from 29 other hospital surgical suites in the United States were used for comparison.

*Four-step Statistical Method to Analyze Prolonged Turnovers*

A turnover time was considered prolonged if it was at least 15 min longer than the mean for the surgical suite.

First, for each combination of 4-week period and hour of the day, calculate the number of prolonged turnover times (table 1). The 4-week period is used, *versus* a shorter period, because otherwise there are hours of the day with no observed prolonged turnovers. We used thirteen 4-week periods (*i.e.*, 1 yr) when analyzing data from the two study hospitals and evaluated using between five and nineteen 4-week periods using Monte-Carlo simulation (below).

Second, for each hour of the day, calculate the lower one-sided 95% confidence bound on the mean number of prolonged turnovers within a 4-week period (see Discussion for rationale). Suppose that for a given 1-hr interval, there were a total of *n* prolonged turnovers during the *m* 4-week periods of data. Then, the lower bound for the mean number of occurrences during a 4-week period equals  $\chi^2(0.95, 2n)/2m$  per 4-week period.<sup>6</sup> For example, at Hospitals A and B, turnovers were studied for the 16 h of the day between 7 AM and 11 PM. For 8 and 14 h of the day, respectively, the mean number

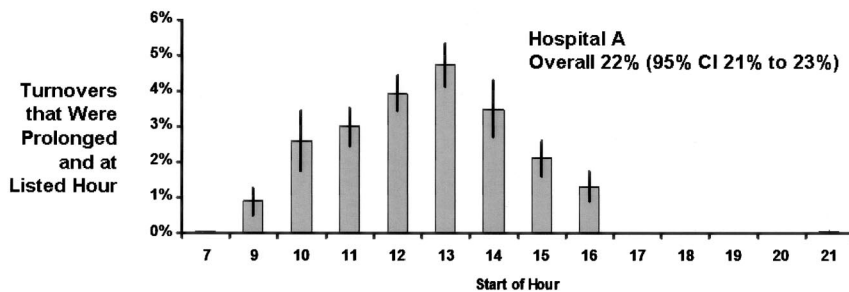


Fig. 1. Percentage of turnovers that were prolonged and occurred at a specified hour of the day for Hospital A. The time along the horizontal axis is the start of the hour of the day (e.g., 9 means that the patient exited the OR between 9:00 and 9:59). The vertical lines show 95% lower and upper confidence intervals (CIs), with corrections for multiple comparisons. The percentage of turnovers that were prolonged and occurred at a specified hour peaked in the middle of the workday, in part because most turnovers occurred in the middle of the workday (table 3, left side).

of prolonged turnovers was at least two every 4 weeks (figs. 1 and 2).

Third, for each 4-week period and hour of the day, calculate the proportion of prolonged turnovers. The numerator equals the number of prolonged turnovers during the 4-week period that occurred during the hour of the day. The denominator equals the total number of turnovers during the 4-week period.

Fourth, for each hour of the day, calculate the sample mean and SD of the proportions from the third step. Use the Student *t* distribution to calculate lower and upper confidence limits of the proportion. Maintain an overall 0.05 type I error rate by using a Bonferroni correction based on the number of comparisons.<sup>7</sup> The number of comparisons (in this fourth step) equals the number of hours of the day for which the lower 95% confidence bound for the mean number of prolonged turnovers during a 4-week period (from the second step) was at least 2. This requirement was necessary for the proportions to be estimated reliably (see Discussion).

#### Validity and Usefulness of the Statistical Method to Analyze Prolonged Turnovers

For development of the methodology, we used 1 yr of data (i.e., thirteen 4-week periods). From just a few 4-week periods, we could make an estimation error when we use the sample estimate as an estimate of the

population proportion. To investigate fewer (5–12) and larger (14–19) numbers of 4-week periods, we used a realistic discrete-event simulation of a 15-OR surgical suite (appendix). Simulation provided a known and unchanging mean turnover time for 1.3 million days. Different subsets of the data were analyzed statistically.

## Results

### Comparing (Benchmarking) Mean Turnover Times among Hospitals

At Hospital A, the sample mean  $\pm$  SD of turnovers was  $37 \pm 16$  min ( $n = 8,592$ ). There was significant positive serial correlation from one turnover to the next turnover ( $P < 10^{-4}$ ). When the average turnover was calculated for each workday, there was significant correlation from one daily average to the next daily average ( $P = 0.04$ ,  $n = 254$  workdays). When the average turnover was calculated for each week, there was significant correlation from one week to the next week ( $P = 0.03$ ,  $n = 52$  weeks). Averaging over 4-week periods was sufficient to eliminate this autocorrelation ( $P = 0.58$ ,  $n = 13$ ). The sample mean  $\pm$  SD of turnovers averaged over 4-week periods was  $37 \pm 1$  min ( $n = 13$ ). Because the distribution of the average turnover times during 4-week periods was consistent with a normal distribution (Lilliefors test,  $P = 0.91$ ), parametric confidence intervals were calcu-

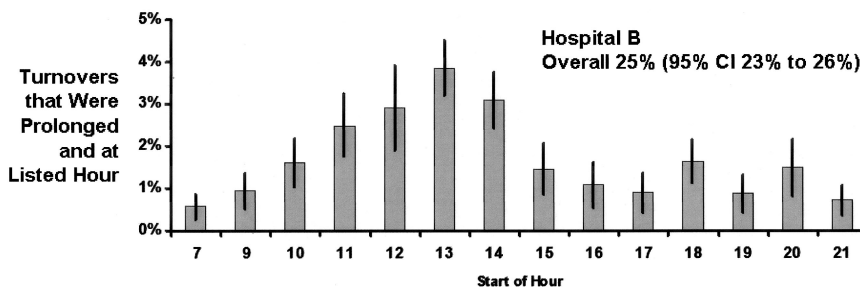


Fig. 2. Percentage of turnovers that were prolonged and occurred at a specified hour of the day for Hospital B. The time along the horizontal axis is the start of the hour of the day (e.g., 9 means that the patient exited the OR between 9:00 and 9:59). The vertical lines show 95% lower and upper confidence intervals (CIs), with corrections for multiple comparisons. The percentage of turnovers that were prolonged and occurred at a specified hour peaked in the middle of the workday, in part because most turnovers occurred in the middle of the workday (table 3, right side).

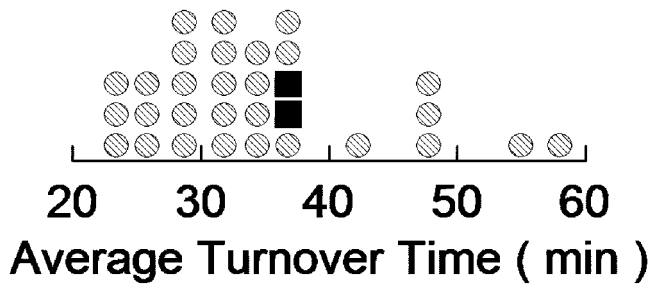


Fig. 3. Variation in average turnover time among 31 hospitals in the United States. The two studied hospitals are presented as filled squares. The bubble plot displays mean time in groups of 2–3 min.

lated. Using the SD of 1 min and the n of 13, the width of the 95% confidence interval for the mean implied by the Student *t* distribution was less than 1 min.

At Hospital B, the mean ± SD of turnovers was 36 ± 16 min (n = 5,930). There was significant positive correlation from one turnover to the next ( $P < 10^{-4}$ ). There was no correlation among daily averages ( $P = 0.80$ , n = 252 workdays). The mean ± SD of turnovers averaged by workday was 36 ± 4 min (n = 252). The distribution of the mean turnovers by day was consistent with a normal distribution by the Lilliefors test ( $P = 0.58$ ). Using the SD of 4 min and the n of 252, the width of the 95% confidence interval for the mean was again less than 1 min.

With a year of data, the process mean of turnover time was estimated reliably. Figure 3 shows the sample mean turnover times of 31 hospitals. The variability in the mean turnover times across hospitals far exceeded the confidence interval widths of less than 1 min. Therefore, the sample mean can be used safely for benchmarking, ignoring estimation uncertainty.

*Turnovers and/or Delays That Are Prolonged at Specified Hours*

Validity of the four-step statistical method was supported by the observation that the proportions obtained from 4-week periods followed approximate normal distributions. The Lilliefors test could not reject normality at the 5% significance level for 21 of the 22 h of the day studied (thirteen 4-week periods for each hour of the day; figs. 1 and 2).

Validity of the statistical method was also supported by results of the Monte-Carlo simulation (appendix). Because 95% confidence intervals were calculated, 5% of the thousands of assessments performed should have at least 1 h of the day for which one or more confidence intervals did not include the true percentage of turnovers that are prolonged. With thirteen 4-week periods (*i.e.*, 1 yr) of data, the type I error rate was accurate to within 0.5% (table 2).

Usefulness of the method depends on the ability to use different numbers of 4-week periods. The simulation

Table 2. Simulated Effect of Number of 4-Week Periods on Accuracy of the Statistical Method Using a Nominal Significance Level ( $\alpha$ ) of  $P < 0.05$

No. of 4-Week Periods	% of Assessments for Which There Was at Least 1 h of the Day When Its CI Did Not Include the True % of Turnovers That Are Prolonged and Occur at That Hour of the Day
5	5.0 ± 0.2%
6	5.1 ± 0.2%
7	5.3 ± 0.2%
8	5.3 ± 0.3%
9	5.4 ± 0.3%
10	5.4 ± 0.3%
11	5.5 ± 0.3%
12	5.6 ± 0.3%
13	5.5 ± 0.3%
14	5.4 ± 0.3%
15	5.2 ± 0.3%
16	5.6 ± 0.4%
17	5.3 ± 0.4%
18	5.8 ± 0.4%
19	5.6 ± 0.4%

If a test performed perfectly, 5% of comparisons would be detected as significantly different. Data are reported ± SE. The mean turnover time was 30 min, making turnovers prolonged that lasted longer than 45 min. This table was created using the simulated data described in the appendix.

CI = confidence interval.

results showed that the type I error rates were accurate to within 1% with five to nineteen 4-week periods (table 2).

Usefulness of the method was supported by its providing new, economically relevant information for managerial decision making. Using # to represent number and *turn* to represent turnovers, the proportion of turnovers that were prolonged and that occurred during the *t*th hour

$$\frac{\# \text{ turn (prolonged, } t\text{th hr)}}{\# \text{ turn}} = \left[ \frac{\# \text{ turn (prolonged, } t\text{th hr)}}{\# \text{ turn (} t\text{th hr)}} \right] \left[ \frac{\# \text{ turn (} t\text{th hr)}}{\# \text{ turn}} \right]$$

The first of the two components can be observed easily by clinicians. For example, at the two studied hospitals, clinicians working late in the workday may observe that at those hours, nearly half of turnovers are prolonged (table 3). The percentage of turnovers at each hour that were prolonged increased progressively over the workday ( $P < 10^{-4}$  by Kruskal-Wallis for both hospitals). However, there were few turnovers near the end of the workday at the two hospitals (table 3). That was why the prolonged turnovers later in the workday were readily observable. The percentage of turnovers that were prolonged and that occurred during each hour of the day peaked at both hospitals during the 1-hr period starting at 1 PM (figs. 1 and 2). This timing does not suggest causes such as lunch breaks, the ending of procedures lasting all morning, or something special in the middle of the day being the cause of most prolonged turnovers occurring in the middle of the day at the two hospitals. If so, the percentage of turnovers at each hour that were

Downloaded from http://aesa2.silverchair.com/anesthesiology/article-pdf/102/6/1242/359086/0000542-200506000-00026.pdf by guest on 20 April 2024

**Table 3. Exploration of Reasons for Findings at Hospital A in Figure 1 and Hospital B in Figure 2**

Start of Hour	Hospital A (Fig. 1)*			Hospital B (Fig. 2)*		
	Mean Turnover, min	% of Turnovers† during the 1-h Period That Are Prolonged‡	% of All Turnovers	Mean Turnover, min	% of Turnovers during the 1-h Period That Are Prolonged‡	% of All Turnovers§
8	26	2%	3%	26	14%	3%
9	32	8%	11%	31	9%	11%
10	35	14%	18%	32	11%	14%
11	36	18%	17%	35	17%	15%
12	39	25%	16%	36	19%	15%
13	41	32%	15%	39	25%	15%
14	41	34%	10%	39	29%	11%
15	42	36%	6%	38	29%	5%
16	44	44%	3%	42	48%	2%

\* This table was created using the raw data from Hospitals A and B. † Column 3 shows the percentage of turnovers during a listed hour that were prolonged, whereas column 3 in table 1 shows the percentage of all turnovers between 7 AM and 11 PM that were prolonged and that occurred during the listed hour. ‡ A turnover time was considered prolonged if it was at least 15 min longer than the mean for the surgical suite. § The percentages in column 7 do not add up to 100%, because Hospital B had some turnovers between 4 PM and 11 PM (fig. 2).

prolonged would have also peaked in the middle of the day. Prolonged turnovers occurred predominantly in the middle of the day at the two hospitals because that was when most turnovers occurred (table 3).

## Discussion

This report provides two new insights with respect to the analysis of turnover times.

Benchmarking mean turnover times among hospitals can be insightful because, for sample sizes usually considered in such studies, the variability in mean turnover times across hospitals far exceeds the estimation error. The mean is an economically relevant measure because it is proportional to the total OR time devoted to setup and cleanup and quantifies the time that is not used for direct patient care. We showed that, for purposes of benchmarking, the sample mean can be reported without attaching to this estimate a confidence interval, because the latter is so narrow (*i.e.*, < 1 min) as to be unneeded.

We developed and validated a method to calculate the percentage of turnovers that are prolonged and that occur during each hour of the day. Managers can aim to reduce prolonged turnovers by focusing efforts on the times of the day with the most prolonged turnovers (figs. 1 and 2). For example, if money were available for two housekeepers who usually work elsewhere to be available part-time for the surgical suite, the time of day for them to be available could be when most prolonged turnovers occur. In addition, the percentage of turnovers that were prolonged and that occurred during each hour of the day can be compared to staffing by hour of the day. The comparison evaluates whether OR nurses and anesthesia providers can more closely match staffing to surgeons' workloads. For example, some surgical suites have frequent prolonged turnovers in the middle of the workday. Cases are not scheduled sequentially

into ORs. Surgeons want to operate in late afternoons after their offices close. Then, some ORs can be allocated to start later in the workday to match when surgeons want to work.

### *Choosing the Number of Hours of the Day to Analyze*

Clinicians and managers interested in routine monitoring of turnover times generally need a robust method that can be applied automatically during report generation, without requiring a formal statistical assessment as is performed during a research study. Simulation was particularly helpful in evaluating our method of spotting and omitting from the analysis hours of the day that had too few prolonged turnovers for valid study (table 2). We required that the 95% lower confidence bound for the mean number of prolonged turnovers during the hour of the day be at least two per 4-week period. This approach was implemented using the second and fourth steps of our method. The approach was reasonable in comparison with the following two alternative rules.

Suppose that when analyzing nineteen 4-week periods, we had simply required that every 4-week period have at least two prolonged turnovers. Using properties of Poisson distributions, the mean number of prolonged turnovers per 4-week period would have to be at least 9 to have a greater than 95% chance that all nineteen 4-week periods would have at least two prolonged turnovers. That mean number of 9 is so large as to result in exclusion of many hours of the day.

Suppose that when analyzing five 4-week periods, we had simply required that the observed mean number of prolonged turnovers at the hour of the day be at least two every 4 weeks. Using properties of Poisson distributions, if 10 prolonged turnovers are observed in five 4-week periods (*i.e.*, the mean is 2), there is a 21% chance that at least two of the five 4-week periods contain no prolonged turnovers. Such zero values violate assumptions of a normal distribution for the proportions,

making the confidence intervals for the turnovers inaccurate.

#### *Correlations among Successive Turnovers*

The focus of our article was consideration of turnover times for routine monitoring, not formal statistical assessment. Nonetheless, our finding that successive turnovers can be correlated has consequences for evaluations as to whether an intervention has successfully reduced turnover times. For the two hospitals that we studied, a two-sample *t* test could not validly be applied to individual observations of turnover times. Individual turnovers could not be pooled naively into two groups: before and after intervention. These results were expected, matching findings for OR staffing costs,<sup>1,8</sup> ORs in use at different times of the day,<sup>9</sup> and OR workload for purposes of OR allocation.<sup>5</sup> A simple and valid solution is simply to pool data, in this instance turnover times, by 4-week period.<sup>1,5,8,9</sup>

#### *Limitations*

Our conclusions likely can be generalized to other hospitals for three reasons.

First, we showed that the uncertainty in the mean turnover time for each hospital was small, relative to variability in mean turnover times among hospitals (fig. 3). We expect that our finding of narrow confidence intervals for the mean for data from hospitals will apply generally because it was a consequence of very large sample sizes for each 4-week period and the Central Limit Theorem. The theorem specifies that the mean of many independent identically distributed random variables is approximately normally distributed, even if the distribution of the random variables is not. The confidence intervals were calculated by using the means from each of several 4-week periods.

Second, we showed how to calculate confidence intervals for the percentages of turnovers during 4-week periods that are prolonged and that occur at an hour of the day, provided there are sufficiently many prolonged turnovers at the hour of the day. The reason that we expect this result to apply generally is again the Central Limit Theorem. The latter conditions would have been highly limiting in practice, had we not designed our statistical method to choose automatically hours of the day with sufficiently many prolonged turnovers for analysis. In addition, we used simulation to show the validity of the method for a wide range of numbers of 4-week periods.

Third, we limited our assessment of validity to a fourth step requiring that the 95% lower confidence bound for the mean number of prolonged turnovers for an hour of the day during a 4-week period be at least two. The choice of two was convenient, in that the results were not sensitive to the choice of other reasonable values, such as 1.9, 2.1, or so forth. We wanted to use as small

a mean number as possible to maximize the number of hours of the day that could be analyzed. However, if our choice were too small, the Student *t* distribution could not have been used to calculate the confidence intervals.

We investigated how to validly benchmark sample mean turnover times and estimate the percentage of turnover times at a surgical suite that are prolonged and occur at specified times of the day. We showed that these endpoints provide useful information. We did not investigate how best to intervene based on these results to change organizational behavior.

We did not separate our analysis according to service. There were three reasons why we chose to combine services.

First, the appropriate interpretation of results for each service can be unclear when different services' cases are performed in the same OR on the same day. A nonsampling error can result from the decision as to which service to attribute a turnover. For example, if myringotomy tube placement precedes a Whipple procedure, the setup time for the Whipple would be attributed to pediatric otolaryngology if turnovers were assigned to the preceding service. How turnovers are attributed negligibly affects resulting OR allocations and staffing.<sup>10,11</sup> However, it does affect the measured turnover times for each service as would be used for benchmarking. A potential solution to this problem is to exclude turnovers between cases performed by different services. However, because the validity of doing so likely varies among surgical suites, we think that validation for each suite would depend on performing the analysis without and with exclusion. Still, a method to do the latter has not yet been developed.

Second, our experience is that when the methods of this article are applied only to turnovers between cases of the same specialty, confidence intervals are wide because of small sample sizes. Apparent differences among specialties can represent random error.

Third, at least at the two studied hospitals, available management interventions to reduce turnovers are generally not specialty perspective. The housekeepers clean the ORs of several services. Logical interventions (*e.g.*, more staff) depend on the hour of the day.

## **Conclusion**

We successfully developed and validated a statistical method to estimate the percentage of turnover times at a surgical suite that are prolonged and occur at specified times of the day. Managers can focus their quality improvement interventions on the identified times of the day with the most prolonged turnovers.

## **Appendix**

Discrete-event computer simulation<sup>12</sup> using ARENA version 7.01 (Rockwell Software, Sewickley, PA) was used to represent the random flow of

patients from ORs through the postanesthesia care unit. Each of the 65,000 workdays was simulated independently of all other workdays.

Scheduled case durations for each of three services were described using different log-normal distributions.<sup>13</sup> Services with brief, moderate, and long durations were assigned mean scheduled durations of 1.0, 2.0, or 3.0 h, respectively, with a common SD of the logarithm of case duration in hours equal to 0.725.<sup>13</sup> After calculation, the scheduled durations were bounded between 0.3 and 1.9 h for the 1-h service, between 0.6 and 3.9 h for the 2-h service, and between 0.9 and 5.9 h for the 3-h service. Durations were rounded up to the nearest 5 min. Turnover times were generated randomly from a log-normal distribution with a mean  $\pm$  SD of  $0.50 \pm 0.25$  h. The 90th, 95th, and 99th percentiles were 0.8, 1.0, and 1.3 h, respectively. The time exceeded the mean by at least 15 min (*i.e.*, were "prolonged" by our choice used in the figures) for 13% of turnovers.

Cases were scheduled sequentially into each OR using an 8-h workday. There were two ORs for the 1-h service, five ORs for the 2-h service, and eight ORs for the 3-h service. Actual case durations were set equal to the scheduled case duration multiplied by a normally distributed random number with a mean of 1.00 and an SD of 0.25.<sup>14-16</sup> In addition, 1% of cases were cancelled at random, resulting in unused OR time.

The proportion of turnovers that were prolonged and that occurred at each hour of the day was estimated using all 1.3 million simulated days. Nonoverlapping 4-week periods of consecutive turnovers were formed. The four-step procedure in the Results section was used to calculate simultaneous confidence intervals for all hours of the day. If one or more of the simultaneous intervals did not cover the proportion for the hour of the day, this was counted as a failure. Table 2 shows the proportion of assessments that were failures. SEs were estimated by calculating Clopper-Pearson confidence intervals.<sup>17</sup>

## References

- Dexter F, Abouleish AE, Epstein RH, Whitten CW, Lubarsky DA: Use of operating room information system data to predict the impact of reducing turnover times on staffing costs. *Anesth Analg* 2003; 97:1119-26
- Vitez TS, Macario A: Setting performance standards for an anesthesia department. *J Clin Anesth* 1998; 10:166-75
- Overdyk FJ, Harvey SC, Fishman RL, Shippey F: Successful strategies for improving operating room efficiency at academic institutions. *Anesth Analg* 1998; 86:896-906
- Adams R, Warner P, Hubbard B, Goulding T: Decreasing turnaround time between general surgery cases: A six sigma initiative. *J Nurs Adm* 2004; 34:140-8
- Dexter F, Macario A, Qian F, Traub RD: Forecasting surgical groups' total hours of elective cases for allocation of block time: Application of time series analysis to operating room management. *ANESTHESIOLOGY* 1999; 91:1501-8
- Hahn GJ, Meeker WQ: *Statistical Intervals: A Guide for Practitioners*. New York, John Wiley & Sons, 1991, pp 123
- Soon G, Huque M: Confidence intervals associated with multiple comparisons. 2003 Proceedings of the American Statistical Association. Alexandria, VA, American Statistical Association, pp 4012-4
- Abouleish AE, Dexter F, Epstein RH, Lubarsky DA, Whitten CW, Prough DS: Labor costs incurred by anesthesiology groups because of operating rooms not being allocated and cases not being scheduled to maximize operating room efficiency. *Anesth Analg* 2003; 96:1109-13
- Dexter F, Epstein RH: Optimizing second shift OR staffing. *AORN J* 2003; 77:825-30
- Abouleish AE, Hensley SL, Zornow MH, Prough DS: Inclusion of turnover time does not influence identification of surgical services that over- and under-utilize allocated block time. *Anesth Analg* 2003; 96:813-8
- Epstein RH, Dexter F: Uncertainty in knowing the operating rooms in which cases were performed has little effect on operating room allocations or efficiency. *Anesth Analg* 2002; 95:1726-30
- Banks J, Carson J, Nelson B: *Discrete-Event System Simulation*, 3rd edition. Paramus, New Jersey, Prentice Hall, 2000, pp 1-40
- Dexter F, Macario A, Manberg PJ, Lubarsky DA: Computer simulation to determine how rapid anesthetic recovery protocols to decrease the time for emergence or increase the phase I post anesthesia care unit bypass rate affect staffing of an ambulatory surgery center. *Anesth Analg* 1999; 88:1053-63
- Kennedy, MH: *Bin-packing, Knapsack, and Change-constrained Approaches to Operating Room Scheduling* [dissertation]. Troy, New York, Rensselaer Polytechnic Institute, Department of Decision Sciences and Engineering Systems, 1992, pp 83-86, 90, 99
- Goldman J, Knappenberger HA, Shearon WT: A study of variability of surgical estimates. *Hosp Manag* 1970; 110:46-46D
- Dexter F, Macario A, O'Neill L: Scheduling surgical cases into overflow block time: computer simulation of the effects of scheduling strategies on operating room labor costs. *Anesth Analg* 2000; 90:980-6
- Newcombe RG: Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Stat Med* 1998; 17:857-72