Mervyn Maze, M.B., Ch.B., Editor

Anesthesiology 2002; 96:1230–6

What Sample Sizes are Required for Pooling Surgical Case Durations among Facilities to Decrease the Incidence of Procedures with Little Historical Data?

Franklin Dexter, M.D., Ph.D.,* Rodney D. Traub, Ph.D.,† Lee A. Fleisher, M.D.,‡ Peter Rock, M.D., M.B.A.§

Background: **Better predictions of each case's duration would reduce operating room labor costs and patient waiting times. A barrier to using historical case duration data to predict the duration of future cases is the absence for some cases of previous data for the same scheduled procedure from the same facility. The authors examined sample size requirements for pooling case duration data from several facilities to create a 90% chance of having case duration data for almost all procedures.**

Methods: **Four academic medical centers provided data, totaling 200,401 cases classified by the scheduled Current Procedural Terminology codes.**

Results: **The 12% of cases in which procedures occurred once or twice accounted for 79% of procedures or combinations of procedures. When a procedure was being performed for the first time at a facility, that same procedure had been performed previously at least once at one or more of the other three facilities only 13–25% of the time. More than 1 million cases would be needed to have a 90% chance of having at least 3 cases for each procedure observed in the original 200,401 cases. However, with N = 200,401 cases in our initial data set, we observed less than one third of the estimated total number of possible procedures.**

Conclusions: **The lack of historical case duration data for scheduled procedures is an important cause of inaccuracy in predicting case durations. However, millions of cases probably would be required to provide historical case duration data for almost all procedures.**

DATA about the duration of surgical cases are needed to predict the time required for future cases. If case durations could be predicted more accurately, decision-support software could help to reduce operating room (OR) overtime labor costs,[1–3] patients' waiting time on the day of surgery,[3–5,6] the number of days patients must wait for surgery,[7] and the time patients wait in their surgeons' afternoon clinics.[8] Generally,[4] the strategy for predicting case durations uses historical data classified by surgeon, type of anesthetic, and scheduled procedure.[1–3,7–9] If surgeon-specific data are not available, case duration can be estimated from the type of anesthetic and scheduled procedure.[9] If information about the anesthetic is not available, historical data classified by the scheduled procedure alone can be used.[9,10] If none of this information is available, case duration data could be obtained from another facility. In this study, we address how many facilities would need to pool data for that approach to be useful.

Ideally, decision-support software built into the OR information system would help the OR manager to make better decisions. However, efforts to improve OR management by using decision support are hampered by the occurrence of cases involving uncommon procedures with little or no historical data. Importantly, by "case" and "procedure," we mean the following. When a patient enters and then leaves an OR, one *case* is performed. A case can involve one or more *procedures*. For approximately half of all surgical cases,[11,12] more than one procedure is performed. *In this article, we use "procedure" to mean one or a combination of procedures performed in a case.* For example, two procedures that are often performed during the same case are phacoemulsification and aspiration of cataract and insertion of intraocular lens. The combination of these two procedures would count as one procedure in this article.

We performed the initial work that identified the absence, for a surprisingly large percentage of cases, of previous data for certain procedures.[11] At an academic medical center, examples of uncommon procedures with limited historical data included anorectal myomectomy, excision of mandibular abscess, coccygectomy, partial ostectomy of the sternum, and intratemporal decompression of the facial nerve.[11] These are typical examples of uncommon procedures. They do not have easily identified analogs. They are procedures that are so

* Associate Professor and Director of the Division of Management Consulting, Department of Anesthesia, University of Iowa, Iowa City, Iowa. † Associate Professor, College of Business Administration, North Dakota State University, Fargo, North Dakota. ‡ Associate Professor and Clinical Director of the Operating Rooms, Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University, Baltimore, Maryland. § Professor and Vice-Chair, Departments of Anesthesiology and Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina.

From the Division of Management Consulting, Department of Anesthesia, University of Iowa, Iowa City, Iowa; the College of Business Administration, North Dakota State University, Fargo, North Dakota; the Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University, Baltimore, Maryland; and the Departments of Anesthesiology and Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina. Submitted for publication June 8, 2001. Accepted for publication January 11, 2002. Supported by University of Iowa Health Care, Iowa City, Iowa. Dr. Dexter and Dr. Fleisher have served as consultants to Picis, Inc., Arlington, Virginia. An abstract describing some this work was presented November 4, 2001 at the annual meeting of the Institute for Operations Research and the Management Sciences, Miami, Florida.

Address correspondence to Dr. Dexter: Division of Management Consulting, Department of Anesthesia, University of Iowa, Iowa City, Iowa 52242. Address electronic mail to: Franklin-Dexter@UIowa.edu. Reprints will not be available from the authors. Individual article reprints may be purchased through the Journal Web site, www.anesthesiology.org.

uncommon that little or no case duration data are available for them.

We were interested in these procedures that are so uncommon that little or no case duration data are available for them. The National Survey of Ambulatory Surgery showed that uncommon procedures occur nationally.[12] Twenty percent of outpatient surgery cases performed in the United States are of a procedure that is performed annually 1,000 times or less in the United States.[12] Thirty-six percent of outpatient surgery cases may be of a procedure that is performed less than once per facility per year.[12]

Procedures that are performed infrequently are important to OR decision support because they have a disproportionately large negative effect on OR efficiency. These uncommon procedures account for much of the dichotomy between scientific success in using historical data to predict case duration[1–3,7–9] and OR management frustration with cases that run late.[1] For example, when using case duration data to reduce OR overtime labor costs and provide suitable breaks for OR nurses and anesthesia providers, the objective is to predict the time required to complete a series of consecutive elective cases.[1,4,5] If as few as 15% of cases have limited historical data (half[12] the national figure) and each OR averages three cases per day, by random chance, more than one third of ORs would include at least one case with limited historical data.[1,11] One late-running case can adversely affect the entire day's schedule.

An important step to improve OR efficiency through decision support is to compensate for the cases that consist of uncommon procedures. This does not mean just determining (or guessing) an expected (average) duration for such procedures (*e.g.*, by asking the surgeon). OR decision support relies on the tails of the probability distributions (*e.g.*, the shortest and longest times the case could take).[4,7,8] A logical approach would be to pool case duration data among facilities, thereby increasing the sample size of historical cases sufficiently to have data even for the uncommon procedures.

For example, suppose a surgeon schedules an anorectal myomectomy at a hospital at which the procedure has not been recently performed. Data for this procedure's duration could be obtained from other hospitals. The resulting statistical analysis would involve calculating a naïve pooled estimate of the historical case durations, or using mixed-effects modeling to compensate for heterogeneity of case duration among facilities. Whatever statistical method is used, the strategy requires that if a procedure is performed infrequently or not at all at one facility, it must be performed at another facility.

In this study, we determine the necessary sample size for pooling case duration data among facilities to have at least three historical case durations for each procedure. If a relatively small sample size is adequate (*e.g.*, 50,000 cases), a surgical facility could partner with several other facilities within its healthcare system to obtain case duration data. However, if the required sample size is large (*e.g.*, 5 million cases), pooled perioperative data would need to come from scores of facilities because many cases would be of procedures that are very uncommon. To answer these questions, we used detailed surgical procedure data from four academic medical centers and less detailed data from a national survey.

Methods

Four academic medical centers provided data about the scheduled procedures for all cases performed at their main and ambulatory surgery facilities during specified time periods (table 1, fig. 1). Two facilities had 4 yr of data available in computerized format, one facility had 3.25 yr, and the fourth facility had 1.3 yr of data. This provided 200,401 cases (table 1).

Cases were classified by their scheduled[13] Current Procedural Terminology (CPT) codes and the presence or absence of an anesthesia provider.[9] If a procedure was designated by more than one CPT code, that combination of codes was considered to characterize a unique procedure. Combinations were considered the same regardless of the order in which the procedures were listed. CPT codes were adjusted to follow January 1, 1999 values. In this article, a procedure is defined as a unique combination of scheduled procedures with or without an anesthesia provider. Repeating the example from the Introduction, two procedures that are often performed during the same case are phacoemulsification and aspiration of cataract and insertion of intraocular lens. The combination of these two procedures performed with topical anesthesia would count as one procedure in this article. The combination of those two procedures with monitored anesthesia care would count as a different procedure in this article. We observed 26,829 different procedures (table 1).

The "observed" number of cases for each procedure was determined from the data from the four medical centers. We calculated the percentage of cases that were of uncommon procedures, defined as one to three historical cases (see two paragraphs below). We also calculated the percentage of procedures that had few previous cases. We calculated standard errors for each of these percentages.

To further explore the effect of pooling data among facilities, we split the data into two sets. The most recent 1,500 cases from each facility were considered to be "new" cases. All preceding cases were used as "historical" data. We calculated the percentages of procedures among the new cases that were not in the historical data.

We hope that pooling the data will decrease the number of procedures with limited data. However, two concerns can limit the usefulness of the pooling of data. First, some procedures may still only appear in the

Table 1. Characteristics of the Dataset

Medical Facility	A	B	C	D	Pooled
Number of cases	23,818	40,975	65,661	69,947	200,401*
Number of procedures among the cases	6,021	7,160	11,288	10,066	26,829*
Dates of available cases	1/99–4/00	9/96–12/99	1/96–12/99	1/96–12/99	—
Percentage ± SE of cases† that were of a procedure with one case ("singletons")	17.2 ± 0.2	12.3 ± 0.2	11.1 ± 0.1	9.3 ± 0.1	9.3 ± 0.1
Percentage ± SE of cases that were of a procedure with two cases ("doubletons")	5.8 ± 0.2	3.1 ± 0.1	3.9 ± 0.1	3.2 ± 0.1	2.6 ± 0.04
Percentage ± SE of cases that were of a procedure with three cases	4.2 ± 0.1	2.0 ± 0.1	2.6 ± 0.1	2.4 ± 0.1	1.7 ± 0.03
Percentage ± SE of procedures† with one case ("singletons")	68.0 ± 0.6	70.5 ± 0.5	64.8 ± 0.4	64.6 ± 0.5	69.7 ± 0.3
Percentage ± SE of procedures with two cases ("doubletons")	11.4 ± 0.4	8.8 ± 0.3	11.4 ± 0.3	11.1 ± 0.3	9.8 ± 0.2
Percentage ± SE of procedures with three cases	5.5 ± 0.3	3.8 ± 0.2	5.1 ± 0.2	5.5 ± 0.2	4.2 ± 0.1

* SE = standard error of the percentage. The number of cases add with pooling data among facilities, but the number of procedures and combinations of procedures from the four facilities does not add to 26,829 because some procedures are performed at more than one facility. † Each use of the word "procedure" in this table refers to a unique procedure. The top three rows give the *percentages of cases* that were of a procedure with one, two, or three cases, respectively. The bottom rows give the *percentages of procedures* for which there were one, two, or three cases, respectively.

database once or twice. This is important because at least three historical cases of the same procedure are needed to provide good predictive accuracy.[11] Second, some procedures may be so uncommon that they do not appear in the database at all. We evaluated these two concerns, using the methods described in the next two paragraphs.

First, it is desirable to have a minimum of three instances of a given procedure in the database. However, some uncommon procedures appear only once or twice. We calculated the minimum number of additional cases needed to create a 90% chance of getting one or two *more* cases of each of the uncommon procedures.‖ This power analysis would provide at least three cases for each procedure that appears in the database.

Second, we addressed those procedures that do not appear in the database. The true number of different procedures will always exceed the observed number.[15] Increasing the sample size will decrease the underestimation of observed *versus* actual procedures. Statistical methods can decrease the effect of sampling error on the estimate of the number of different procedures. Let F_1 equal the number of procedures with only one observed case (*i.e.*, singletons), let F_2 equal the number of procedures with two observed cases (*i.e.*, doubletons), and so forth. S_{obs} is the total number of procedures observed. S_{abund} is the number of "abundant" procedures, defined, based on simulation of combinatorial statistics,[16] as those with 11 or more cases. The remaining procedures are considered in these statistical methods to be "rare," meaning that $S_{obs} = S_{abund} + S_{rare}$. Then, a conservative (lower) value can be derived for the total number of different procedures:

$$S_{obs} + F_1^2/(2F_2).^{15,17}$$

‖ We assume that x cases, of a specified procedure, have been observed in a random sample of size n from the population of all surgical cases, where x = 1 or 2. We want to calculate the smallest possible future sample size m required to provide for a 90% probability of obtaining y future cases of the specified procedure. For the value x = 1, the value of y = 2, and *vice versa*. By trial and error, we repeatedly performed the Fisher exact test to obtain the smallest value of m providing for $P \leq 0.1$.[14]

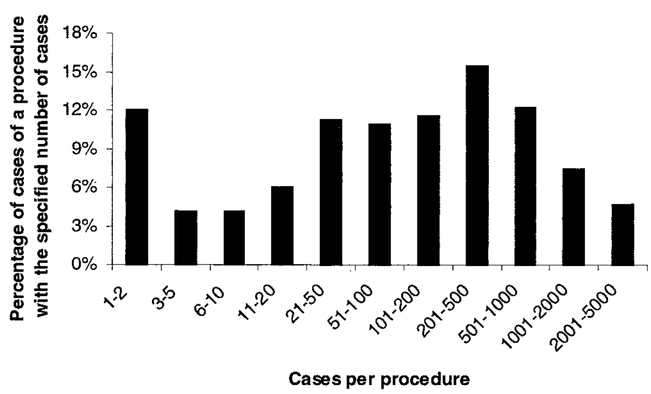

Fig. 1. Histogram of the percentage of cases that were of a procedure that occurred a specified number of times in the pooled case duration data from four medical facilities. When a patient enters and then leaves an operating room, one *case* is performed. The case can include one or more surgical procedures. To obtain the histogram, the 200,401 cases were sorted by procedure. The number of times that a patient entered an operating room and underwent a specified procedure was counted. The counts were sorted in ascending order. Each bar shows the sum of the counts divided by the total number of cases. The number of cases per procedure was then plotted along the horizontal axis. These numbers are reported as a range for each bar, rather than on a logarithmic scale, to indicate clearly the values for procedures performed only once or twice (the far left bar).

The corresponding variance of this estimator equals

$$F_2\Big((F_1/F_2)^4/4 + (F_1/F_2)^3 + (F_1/F_2)^2/2\Big).^{15,18}$$

This method is conservative because its calculation is performed without estimating the heterogeneity of the frequency of occurrence (F_k) of different procedures. This coefficient of variation of the F_k, referred to by γ,[2] can be estimated by taking the maximum of 0 and

$$-1 + \left(S_{rare} \sum_{k=1}^{10} k(k-1)F_k\right) \Big/ (N_{rare}[1 - F_1/N_{rare}][N_{rare} - 1]),$$

where N_{rare}, the number of cases that are of a rare procedure, equals

$$\sum_{k=1}^{k=10} kF_k.$$

The resulting more sophisticated estimator for the true number of different procedures in the population equals

$$S_{abund} + (S_{rare} + F_1\gamma^2)/(1 - F_1/N_{rare}).^{16}$$

Both of these estimators can underestimate the true number of different procedures when the sample sizes are sufficiently small that more than 30% of the different procedures are not observed.[19,20] Therefore, we used graphical methods to evaluate how many more procedures surgeons perform than we observed.

We pooled data among the four medical centers to create a histogram of the number of cases of each procedure, plotted on a logarithmic scale (fig. 2). Such plots with logarithmic scales typically yield normal distributions.[1,19,21] If our sample size was too small to observe many of the rare procedures, the plot would resemble a bell-shaped curve with an opaque card covering the left side of the curve with the most uncommon procedures.[21] The relatively common procedures, on the right tail of the histogram, would be detected.

We used an additional data set to assure that our expectation was correct for the shape of the underlying statistical distribution of the number of cases of each procedure (fig. 3). We used data from the United States National Center for Health Statistics' 1994 to 1996 National Survey of Ambulatory Surgery. We recently reported details and demographics of this survey.[12,22] We used Excel Visual Basic 6.0 (Microsoft, Redmond, WA) to analyze the raw data for the sample of 228,332 completed surgical cases with an anesthesia provider. The 24,084 different procedures were classified by up to six ICD-9-CM (International Classification of Disease) codes.

The survey used probability sampling so that nationally representative results could be obtained without surveying every ambulatory surgery case in the United States.[4,12] The National Center for Health Statistics assigned each case a weight using statistical methods that considered the probability of selecting the case's facility, the probability of selecting the case among all cases at the case's facility, and the response rates of facilities and locations within facilities. For example, some cases had weights of 10 (*i.e.*, represented 10 outpatient cases nationally), and others had weights of 20,660 (*i.e.*, represented 20,660 outpatient cases nationally). We created a histogram showing nationally representative results by calculating, for each observed procedure, the sum of the weights of all cases of the procedure and then dividing by the sum of the weights of all cases. We were not able to apply the two nongraphical statistical methods to this

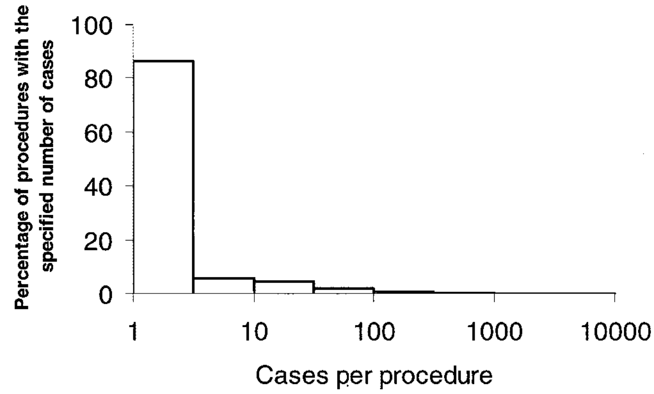

Fig. 2. Histogram of the number of procedures performed the specified number of times in the pooled case duration data from four medical centers. This plot was made using 26,829 procedures and 200,401 cases from the four facilities. When a patient enters and then leaves an operating room, one *case* is performed. The case can include one or more procedures. To obtain the histogram, the data were sorted by procedure. The number of times that a patient entered an operating room and underwent each specified procedure was counted. The graph is a histogram of these counts plotted on a logarithmic axis. Figure 3 shows the expected histogram with a larger sample size.[21]

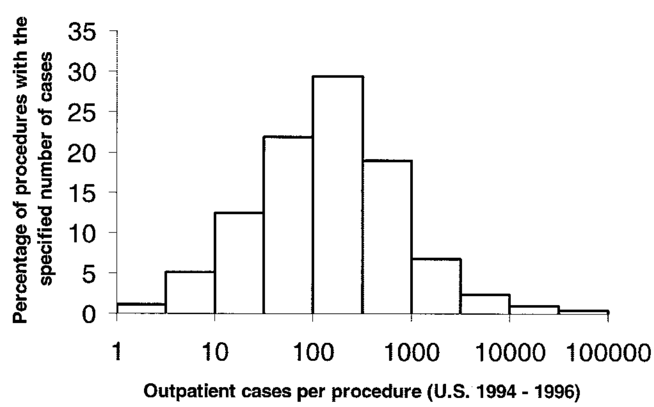

Fig. 3. Histogram of the number of procedures that were performed the specified number of times in the United States between 1994 and 1996. Data are from the National Survey of Ambulatory Surgery raw data released by the National Center for Health Statistics. When a patient enters and then leaves an operating room, one *case* is performed. The case can include one or more procedures. There were 228,332 cases in the survey.

Table 2. Percentage of Cases of a Procedure Not Previously Performed at One Facility That Was Performed at Another

	Facility	A	B	C	D
Demographics as in table 1	Number of cases excluding the most recent 1,500 cases from each facility	22,318	39,475	64,161	68,447
	Percentage ± standard error of the most recent 1,500 cases that was of a procedure that was not previously performed at the facility	18.1 ± 1.0	14.2 ± 0.9	12.0 ± 0.8	10.7 ± 0.8
Results	Percentage of these cases (with no historical data) for which two or more previous cases of the same procedure were available using data from other facilities	18.0 ± 2.3	16.4 ± 2.5	8.9 ± 2.1	17.4 ± 3.0
	Percentage of these cases (with no historical data) for which at least one previous case of the same procedure was available from another facility	25.4 ± 1.1	19.2 ± 1.0	12.8 ± 0.9	24.8 ± 1.1

data because the survey's use of probability sampling violated the assumptions of these statistical methods.

Results

In this article, a procedure is defined as a unique combination of scheduled procedures with or without an anesthesia provider. Table 1 shows that after pooling 200,401 cases from the four academic medical centers, 11.9% of the cases were of a procedure that occurred only once (singletons) or twice (doubletons). Figure 1 shows this graphically.

The 12% of the 200,401 cases that were singletons or doubletons accounted for 79% of the procedures (table 1). This was larger than the percentages of singletons or doubletons for any one facility (table 1). That is, the sample size of 200,401 cases was sufficiently small that pooling the data among the four facilities did not decrease the incidence of rare procedures. For the first instance of a procedure being performed at a facility, 13–25% of the time, that procedure had been performed previously at least once at one or more of the other three facilities (table 2, last row). More than 1 million cases would be needed to have a 90% chance of having at least 3 cases for each procedure observed in the original 200,401 cases (table 3, third row of numbers).

A histogram of the frequency of each procedure from the National Survey of Ambulatory Surgery was bell shaped when plotted on a logarithmic scale (fig. 3). This curve is consistent with the expectations described in the Methods.[19,21] Procedures with a moderate number of cases were more numerous than procedures that had a small or exceedingly large number of cases. This contrasted with the pooled cases from the four facilities, in which more than half of the procedures were scheduled only once (fig. 2). As explained in the Methods, these graphical and inferential methods suggest that many procedures that surgeons perform were not observed in the 200,401 cases (*i.e.*, the sample size of 200,401 cases was too small to provide sufficient data).

We calculated a lower bound on the number of procedures that surgeons perform. The estimate included not only the procedures we observed, but also those that were sufficiently rare that we did not observe them (table 4). The conservative (lower) estimate was 93,340 procedures. The more sophisticated estimator for the estimated total number of different procedures was 96,497. Therefore, from 200,401 cases, we observed fewer than one third of the estimated total number of procedures (table 1).

Furthermore, 96,497 procedures is an underestimate of the true number of different procedures. As exemplified by the findings for each facility (table 4), these estimators underestimate the true number of different procedures when the sample sizes are sufficiently small that more than 30% of the different procedures are not observed.[19,20] The graphical and inferential results both suggest that this is true for the pooled data from the four facilities.

Table 3. Power Analysis (Reported to the Nearest 1,000 Cases)

Facility	A	B	C	D	Pooled
Number of cases from the facility	23,818	40,975	65,661	69,947	200,401
Number of additional cases needed for a 90% chance of increasing from one case of a procedure to at least two	214,000	369,000	591,000	629,000	1,803,000
Number of additional cases needed for a 90% chance of increasing from one case of a procedure to at least three	673,000	1,159,000	1,860,000	1,981,000	5,675,000
Number of additional cases needed for a 90% chance of increasing from two cases of a procedure to at least three	51,000	88,000	142,000	151,000	443,000

Table 4. Numbers of Different Surgical Procedures and Combinations of Surgical Procedures

Facility	A	B	C	D	Pooled*
Number of procedures among the cases	6,021	7,160	11,288	10,066	26,829
Conservative (*i.e.,* low) estimate of the number of procedures, including those we did not observe	18,243	27,440	32,034	28,928	93,340
95% Confidence interval for the conservative estimate	17,041–19,446	25,479–29,402	30,529–33,539	27,467–30,388	90,121–96,599
More sophisticated estimator for the number of procedures, including those not observed†	18,285	26,458	32,597	28,892	96,497

* The number of procedures from the four facilities does not add to 26,829 because some procedures are performed at more than one facility. † As described in the Methods, this estimator incorporates an estimate of the heterogeneity of the frequency of occurrence of different procedures.

Discussion

Implications for Predicting Case Durations

We previously showed that more than one third of outpatient surgery cases may be of a procedure performed less than once per facility per year.[12] In this article, we report pooling case duration data among four facilities to try to increase the number of previous cases for each scheduled procedure. However, we failed to decrease the percentage of procedures that were performed only once or twice. To obtain data for almost all procedures, more than 1 million historical cases are needed (tables 3 and 4).

Our results provide insight into appropriate pooling of procedure-specific surgical data among facilities. Pooling such data may be important because data about surgical case duration are needed to minimize OR overtime labor costs, patients' waiting times on the day of surgery, and patients' waiting times in surgeons' afternoon clinics.[1–5,7,8] However, we showed in this study that informal arrangements for pooling data among several small- to moderate-sized facilities within a healthcare system are unlikely to be sufficient to improve the accuracy of case duration predictions. Instead, to increase the accuracy by simply pooling case duration data, larger databases will be needed.

We focused on the sample size required to obtain at least three historical cases for each procedure. When estimating the average duration of a case, three historical cases is a sufficiently large number that variability in case duration is affected more by intrinsic variability in case duration for the procedure than by uncertainty in the value of the true mean.[11] However, when predicting the shortest or longest times needed to complete a case (*e.g.*, for calculating optimal patient arrival times[4] or scheduling breaks between cases,[6] respectively), more cases are needed for the effect of uncertainty in parameter estimates, from the small sample size, to have a smaller effect than intrinsic variability in case duration.[4] As such, our focus on obtaining three previous cases had the deliberate effect of underestimating the necessary size of a case duration database.

Inferential and Graphical Methods of Estimating the Number of Different Procedures

We know that surgeons schedule at least 26,829 different CPT codes and combinations of CPT codes because we observed that many. The true number of different procedures must be higher. To estimate the true diversity of procedures, we used statistical methods developed for other scientific fields.[15–21] The less and more sophisticated methods estimated 90,368 and 96,497 different procedures, respectively (table 4). We do not know which one is closer to being correct, nor do we think that it matters. The point is that despite the large effort required to acquire 200,401 cases, we did not come close to observing all of the procedures.

We used the graphical method to address the same question: Was our sample size of 200,401 even close to being large enough to detect all surgical procedures? Clearly it was not, because if it had been, the histogram of the number of procedures with a certain number of observed cases (fig. 2) would have more closely resembled its true statistical distribution (fig. 3).[21]

Alternative Strategies to Pooling Data among Facilities

The OR manager can take some steps to decrease the effect that uncommon procedures have on OR efficiency. We previously showed that when choosing the sequence of elective cases performed by the same surgeon in the same OR on the same day, the OR manager's primary nonmedical criterion could be to avoid limitations in equipment or personnel that would result in staffed but unused OR time.[7] This approach not only serves to minimize hospital and anesthesia costs, but also to maximize surgeon and patient convenience.[7,23] When there are no such restrictions, the criterion for sequencing a series of elective cases by the same surgeon on the same day could be to minimize the time patients wait at the surgical suite on the day of surgery.[4,7] Scheduling cases with common procedures before a case of an uncommon procedure will decrease the expected difference between scheduled and actual case start times. This will benefit the patients and surgeons both.

Limitations

A limitation to our work is that we may have pooled data from the "wrong" four medical centers. Hypothetically, out of the thousands of different surgical facilities worldwide, we may have chosen the four facilities with the highest incidence of rare procedures. However, previous results from the National Survey of Ambulatory Surgery suggest that this did not happen.[12]

Another limitation to our work is that we considered combinations of procedures to be new, unique procedures. We did this because we are not aware of algorithms that can accurately predict the time needed for a case that is a combination of procedures from the case duration of each of the procedures separately. If such a methodology could be developed, the sample sizes that we calculated could be decreased. However, the absence of current literature describing this approach does not reflect a lack of research effort, but the failure, to date, of that work.

The relatively high frequency of very uncommon procedures is not a "trick" of the CPT system. Procedures with easily identifiable analogs are generally not uncommon. For example, coronary artery bypass with three venous grafts (CPT code 33512) is not uncommon. Coronary artery bypass with four venous grafts (CPT 33513) is also not uncommon. These two procedures differ only in the last digit of the CPT code. Both procedures are less common than if the CPT system combined them in one code. Nevertheless, neither is so uncommon as to be absent. Otherwise, the CPT system would not distinguish between them.

Conclusions

The lack of historical case duration data for scheduled surgical procedures is an important cause of inaccuracy in predicting case durations. For the strategy of simply pooling case duration data among facilities to provide data for almost all procedures, databases with millions of cases probably will be needed. Therefore, many hospitals would need to pool data if this strategy were used.

References

1. Dexter F, Traub RD, Qian F: Comparison of statistical methods to predict the time to complete a series of surgical cases. J Clin Monit Comput 1999; 15:45–51
2. Dexter F, Macario A, O'Neill L: A strategy for deciding operating room assignments for second-shift anesthetists. Anesth Analg 1999; 89:920–4
3. Dexter F: A strategy to decide whether to move the last case of the day in an operating room to another empty operating room to decrease overtime labor costs. Anesth Analg 2000; 91:925–8
4. Dexter F, Traub RD: Statistical method for predicting when patients should be ready on the day of surgery. Anesthesiology 2000; 93:1107–14
5. Dexter F, Macario A, Traub RD: Optimal sequencing of urgent surgical cases: Scheduling cases using operating room information systems. J Clin Monit Comput 1999; 15:153–62
6. Dexter F, Traub RD, Lebowitz P: Scheduling a delay between different surgeons' cases in the same operating room on the same day using upper prediction bounds for case durations. Anesth Analg 2001; 92:943–6
7. Dexter F, Traub RD: Sequencing cases in operating rooms: Predicting whether one surgical case will last longer than another. Anesth Analg 2000; 90:975–9
8. Zhou J, Dexter F: Method to assist in the scheduling of add-on surgical cases: Upper prediction bounds for surgical case durations based on the log normal distribution. Anesthesiology 1998; 89:1228–32
9. Strum DP, Sampson AR, May JH, Vargas LG: Surgeon and type of anesthesia predict variability in surgical procedure times. Anesthesiology 2000; 92:1454–67
10. Macario A, Dexter F: Estimating the duration of a case when the surgeon has not recently performed the procedure at the surgical suite. Anesth Analg 1999; 89:1241–5
11. Zhou J, Dexter F, Macario A, Lubarsky DA: Relying solely on historical surgical times to estimate accurately future surgical times is unlikely to reduce the average length of time cases finish late. J Clin Anesth 1999; 11:601–5
12. Dexter F, Macario A: What is the relative frequency of uncommon ambulatory surgery procedures in the United States with an anesthesia provider? Anesth Analg 2000; 90:1343–7
13. Dexter F: Application of prediction levels to operating room scheduling. AORN J 1996; 63:607–15
14. Thatcher AR: Relationships between Bayesian and confidence limits for prediction. J Royal Stat Soc Series B 1964; 26:176–210
15. Colwell RK, Coddington JA: Estimating terrestrial biodiversity through extrapolation. Phil Trans R Soc Lond 1994; 345:101–18
16. Chao A, Ma MC, Yang MCK: Stopping rules and estimation for recapture debugging with unequal failure rates. Biometrika 1993; 80:193–201
17. Chao A: Non-parametric estimation of the number of classes in a population. Scand J Stat 1984; 11:265–70
18. Chao A: Estimating the population size for capture-recapture data with unequal catchability. Biometrics 1987; 43:783–91
19. Schmit JP, Murphy JF, Mueller GM: Macrofungal diversity of a temperate oak forest: A test of species richness estimators. Can J Bot 1999; 77:1014–27
20. Chao A, Hwang WH, Chen YC, Kuo CY: Estimating the number of shared species in two communities. Stat Sinica 2000; 10:227–46
21. Preston FW: The commonness and rarity of species. Ecology 1948; 29:254–83
22. Dexter F, Traub RD: The lack of systematic month-to-month variation over one-year periods in ambulatory surgery caseload: Application to anesthesia staffing. Anesth Analg 2000; 91:1426–30
23. Dexter F, Macario A, Traub RD, Hopwood M, Lubarsky DA: An operating room scheduling strategy to maximize the use of operating room block time: Computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time. Anesth Analg 1999; 89:7–20