ECONOMICS

Dennis M. Fisher, M.D., Editor

Anesthesiology
1998; 89:1228–32

Method to Assist in the Scheduling of Add-on Surgical Cases—Upper Prediction Bounds for Surgical Case Durations Based on the Log-normal Distribution

Jinshi Zhou, Ph.D., Franklin Dexter, M.D., Ph.D.†*

Background: **A problem that operating room (OR) managers face in running an OR suite on the day of surgery is to identify "holes" in the OR schedule in which to assign "add-on" cases. This process necessitates knowing the typical and maximum amounts of time that the case is likely to require. The OR manager may know previous case durations for the particular surgeon performing a particular scheduled procedure. The "upper prediction bound" specifies with a certain probability that the duration of the surgeon's next case will be less than or equal to the bound.**

Methods: **Prediction bounds were calculated by using methods that (1) do not assume that case durations follow a specific statistical distribution or (2) assume that case durations follow a log-normal distribution. These bounds were tested using durations of 48,847 cases based on 15,574 combinations of scheduled surgeon and procedure.**

Results: **Despite having 3 yr of data, 80 or 90% prediction bounds would not be able to be calculated using the distribution-free method for 35 or 49% of future cases *versus* 22 or 22% for the log-normal method, respectively. Prediction bounds based on the log-normal distribution overestimated the desired value less often than did the distribution-free method. The chance that the duration of the next case would be less than or equal to its 90% bound based on the log-normal distribution was within 2% of the expected rate.**

Conclusions: **Prediction bounds classified by scheduled surgeon and procedure can be accurately calculated using a method that assumes that case durations follow a log-normal distribution. (Key words: Operating rooms; operating room information systems; prediction bound; staff scheduling; statistical interval.)**

* Postdoctoral Fellow.

† Associate Professor.

Received from the Department of Anesthesia, University of Iowa, Iowa City, Iowa. Submitted for publication May 1, 1998. Accepted for publication July 16, 1998. Dr. Zhou was supported by GM08442-05, U.S. P.H.S. Anesthesia Research Training Grant.

Address reprint requests to Dr. Dexter: Department of Anesthesia, University of Iowa, Iowa City, Iowa 52242. Address electronic mail to: franklin-dexter@uiowa.edu

"ADD-ON" surgical cases usually are scheduled individually. A problem that operating room (OR) managers face in running an OR suite on the day of surgery is to identify "holes" in the OR schedule in which to assign add-on cases. Often, this process necessitates knowing the typical and maximum amounts of time that the case is likely to require. For example, consider that at 1:00 PM anesthesiologists and nurses are available to staff an OR until 4:00 PM and a surgeon would like to perform a procedure. How does the OR manager determine whether the case would typically (*e.g.*, has a 50% chance) be finished by 3:00 PM or whether there is a reasonable (*e.g.*, 80% or 90%) chance that the case will be finished by 4:00 PM? The "upper prediction bound" is the predicted case duration for which there is a specified probability that the duration of the surgeon's next case will be less than or equal to the predicted duration.[1] For example, the functional definition of the case's "90% upper prediction bound" is that it is a value that will, with 90% degree of confidence, be equal to or greater than the duration of the next case.[2] Prediction bounds differ from confidence bounds in that the former refers to durations of future cases, whereas the latter refers to mean durations. If previous (*i.e.*, historical) case durations for a particular surgeon performing a particular scheduled procedure are stored in an information system, the 90% prediction bounds (or 50% or 80%, whatever is desired) can be calculated.

We[1] previously investigated how to calculate upper prediction bounds for case duration by using a distribution-free methodology[2] (*i.e.*, a nonparametric method). This approach does not assume that the statistical distribution of case duration follows any particular distribution. However, the method needs at least nine historical cases to calculate a 90% prediction bound, and optimally would need 19 historical cases. As shown herein, this requirement is often not satisfied when case durations are classified based on the combination of surgeon and scheduled procedure. Case durations have two characteristics that are representative of log-normal statistical distributions: (1) case durations are always greater than zero and (2) outliers tend to be very long in duration, not short. In this study, we therefore evaluate the accuracy of an alternative approach to calculating prediction bounds for surgical case durations, which is based on log-normal statistical distributions.[3]

Methods

Data Set

Case duration was defined to equal the time from when the patient enters an OR to when the patient leaves the OR. Case durations were obtained for all surgical cases performed during the 1994 through 1997 fiscal years at the OR suites of the University of Iowa Hospitals and Clinics. The 48,847 cases were classified based on the 15,574 combinations of scheduled surgeon and scheduled procedure. Procedures with the same combination of scheduled Current Procedural Terminology 1997 codes were considered the same procedure. Procedures were defined using the scheduled (*versus* actual) procedure code because (1) for the new case to which a prediction bound will be compared, only the scheduled procedure is known and (2) for some surgeons and scheduled procedures, actual procedures occasionally differed from scheduled procedures.

Calculation of the Prediction Bounds

Mathematic algorithms, formulas, and comments regarding the two methods to calculate prediction bounds are provided in the appendix. At least one, four, or nine previous case durations are necessary to calculate 50%, 80%, or 90% upper prediction bounds using the distribution-free method. At least two previous case durations are necessary to calculate any of the 50%, 80%, or 90% upper prediction bounds using the method based on a log-normal distribution of case durations.

Testing the Two Methods to Calculate Prediction Bounds

Separate analyses using the data set were performed for (1) each of the two methods to calculate prediction bounds; (2) 50%, 80%, and 90% prediction bounds; and (3) values of N ranging from 2 to 19 that are appropriate for the method to calculate prediction bounds. For each of the analyses, the following methodology was used:

1. A counter was set to zero.
2. The first surgical case in the data set was considered.
3. The date at which the case being considered was performed was read.
4. Did the scheduled surgeon perform the scheduled procedure at least N times before that date? If no, then the next step was skipped, and processing advanced to step 6.
5. The most recent N of the previous cases were used to calculate the prediction bound. If the duration of the case being considered was longer than the prediction bound, the counter was incremented by one.
6. Has the last of the 48,847 cases in the data set been considered? If not, then the next case in the data set was considered, and processing returned to step 3.
7. The proportion of cases exceeding the upper prediction bound for the method to calculate prediction bounds equaled the value of the counter divided by the number of prediction bounds that was calculated for the analysis (*i.e.*, the number of times that step 5 was performed).

All analyses were performed using SAS (SAS Institute, Cary, NC).

Results

Distribution-free Prediction Bounds

Prediction bounds cannot be calculated for all future cases. For example, the first column and row combination of table 1 specifies the requirement that two previous cases are available to estimate a prediction bound. The second and third columns reflect that at least two cases would be available to calculate a prediction bound for another case for 32% of the surgeon-and-procedure combinations or 78% of cases. Referring to the third and seventh rows of table 1, 80% and 90% prediction bounds could be calculated using the distribution-free method for 65% and 51% of future cases, respectively. Despite having a data set of 48,847 previous cases, 80% or 90% prediction bounds would not be able to be calculated for 35% or 49% of future cases (table 1). The finding that 90% prediction bounds cannot be calculated for 49% of future cases is the disadvantage to the distribution-free method that prompted this project.

Ideally, the 50%, 80%, or 90% prediction bounds would exceed the duration of the next case exactly 50%, 20%, or 10% of the time, respectively. Prediction bounds calculated using the distribution-free method generally ex-

Table 1. Method to Calculate Upper Prediction Bounds

Number of Previous Cases Used to Estimate Bound	% of Surgeon and Procedure Combinations for Which a Prediction Bound Can Be Calculated	% of New Cases for Which a Prediction Bound Can Be Calculated	% of Cases Exceeding the Specified Prediction Bound					
			Distribution-free Method			Log-normal Method		
			50% Bound	80% Bound	90% Bound	50% Bound	80% Bound	90% Bound
2	32	78	33	—	—	49	21	11
3	20	70	49	—	—	49	20	10
4	14	65	39	19	—	49	19	10
5	11	62	48	16	—	49	19	10
6	9	58	41	14	—	48	19	10
7	8	55	48	12	—	48	18	10
9	6	51	48	19	10	48	18	9
11	5	48	47	16	8	48	18	9
13	4	44	47	13	7	48	18	9
15	3	42	47	18	6	47	18	9
17	3	40	47	16	5	47	18	9
18	3	39	44	15	5	47	18	9
19	3	38	47	19	9	47	18	9

An increase in the number of previous cases used to estimate the upper prediction bound (column 1) causes a decrease in the percentage of new cases for which a prediction bound can be calculated (column 3), because fewer surgeon and procedure combinations (column 2) have the minimum number of previous cases. The standard errors of percentages (columns 4 to 9) ranging from 0 to 3%, 4 to 8%, 9 to 17%, and 18 to 78% were 0.1%, 0.1 to 0.2%, 0.2 to 0.3%, and 0.3 to 0.5%, respectively. If the prediction bounds worked perfectly, the percent of cases exceeding the 50%, 80%, and 90% bounds would equal 50%, 20%, and 10%, respectively. When the percent of cases exceeding these bounds is less than these specified (nominal) rates, the prediction bounds overestimated the desired value. The vacant entries in the fourth and fifth columns show that at least 4 and 9 previous case durations are needed to calculate 80% and 90% distribution-free upper prediction bounds, respectively.

ceeded durations of the next case by more than the specified (nominal) percent of cases to exceed the bound (table 1). Therefore, distribution-free prediction bounds tended to overestimate the desired value. There is generally a greater than 50%, 80%, or 90% chance that the duration of the next case will be less than or equal to the 50%, 80%, or 90% upper prediction bound calculated using the distribution-free method.

Prediction Bounds Based on Log-normal Distribution

The 50%, 80%, or 90% prediction bounds could not be calculated using this method for 22% of future cases (table 1). When 80% or 90% prediction bounds are desired, the method based on the log-normal distribution can calculate them for more cases than the distribution-free method.

Ninety percent prediction bounds based on a log-normal distribution underestimated durations of the next case 1% more often than the expected (nominal) percent of cases to exceed the bound when there were two previous cases (table 1). Therefore, if OR managers are particularly interested in using the method to be nearly (*i.e.*, 10%) certain that a case will be performed within the time specified by the prediction bound, then they may only want to consider prediction bounds based on log-normal distributions when there are at least three previous cases.

The prediction bounds based on the log-normal distribution overestimated the desired value less often than did the distribution-free method (table 1). The chance that the duration of the next case would be less than or equal to its 90% bound was within 2% of the expected (nominal) rate of 10%.

Discussion

Examples of How an OR Manger Can Use Upper Prediction Bounds for Case Duration

In the introduction we considered the application of upper prediction bounds to scheduling of add-on cases at the end of the regularly scheduled operating day. There are at least two other applications of upper prediction bounds to OR management.

1. A case in an OR is running late, and the OR manager wants to move the next case in that OR to a different OR. The anesthesiologist who would care for the patient has to leave by 5:00 PM. Using the 80% or 90% upper prediction bound, the OR manager can assure the anesthesiologist that there is an 80% or 90% chance that she will be finished by 5:00 PM.

2. A surgeon wants to repair an inguinal hernia. An OR is available. However, another surgeon will arrive to start a case in 2 h. Should you put the case into the OR? On average, the surgeon can complete the procedure (including induction and emergence) within the 50% upper prediction bound. However, the case could necessitate more than average time. The 80% or 90% upper prediction bound can provide information to help direct the decision.

Log-normal Distribution

The results in table 1 provide evidence that log-normal distributions can be adequate to describe the statistical distributions of case durations for a particular surgeon and procedure. This result was expected because (1) case durations are always greater than zero and (2) outlier cases tend to be very long in duration, not short. Both of these characteristics are hallmarks of the log-normal distribution, and are not satisfied by the normal distribution.

Earlier we reported that surgical case durations for common procedures, classified based on procedure only, do not reliably follow either a normal or a log-normal distribution.[1] We therefore used a distribution-free method to calculate prediction bounds,[1,3] because they can be used without concern regarding the appropriate statistical distribution. Table 1 shows that prediction bounds based on a log-normal distribution are accurate for case durations classified by surgeon and procedure. Results from table 1 can explain the results that we obtained in our previous study. In our previous[1] study, case durations were classified based on procedure. If case durations classified by surgeon and procedure follow a log-normal distribution (table 1), then case durations classified by procedure alone may not follow a log-normal distribution. Instead, they will follow a statistical distribution that represents a mixture of different log-normal distributions weighted by the number of procedures performed by each surgeon. For small numbers of different surgeons, this mixture of different log-normal distributions may differ significantly from a log-normal distribution. Based on this observation, we recommend that OR managers do not use prediction bounds based on the log-normal distribution for case durations defined in a different manner without completing an analysis similar to the one that we present in this article.

This study shows why it has been difficult to determine the appropriate statistical distribution for surgical case durations. The clear approach to this problem would be to test whether case durations for each combination of scheduled surgeon and procedure follow a log-normal statistical distribution.[4] To do so, ideally there would be approximately 200 case durations for each combination. However, in our data set of 48,847 cases there were 15,574 combinations of scheduled surgeon and procedure. These data are similar to those reported by Strum *et al.*[4] for "a large teaching hospital," at which 46,317 cases had 5,122 different procedures (not surgeon and procedure) after excluding procedures with more than one current procedural terminology code.[4] The consequence of these many combinations is that only 3% of combinations have at least 19 case durations (table 1), far from the 200 case durations that would be desired. Even if 19 case durations were adequate to test a statistical distribution, the results would then apply only to 3% of the combinations.

Maximum Number of Previous Case Durations to Use in Calculating Prediction Bounds

Surgeon case durations for specific procedures may change progressively,[1] for example, as a result of subtle changes in demographics of a patient population. When OR managers calculate upper prediction bounds, they may not want to include all previous case durations available for the specified surgeon performing the specified procedure. The more previous cases that the OR manager includes in calculating a prediction bound, the greater the chance that the surgeon may have become faster or slower progressively.[1] Therefore, there is an advantage to including as few previous cases as possible when calculating upper prediction bounds, yet ensuring that the prediction bounds do not systematically overestimate or underestimate the desired value. Based on this criterion and the results of table 1, OR managers may want to consider limiting calculation of prediction bounds based on the log-normal distribution to the most recent nine previous case durations, when more than nine previous case durations are available.

Summary

The upper prediction bound specifies with a certain probability that the duration of the surgeon's next case will be less than or equal to the bound. Prediction bounds classified by scheduled surgeon and procedure can be calculated accurately using a method that assumes that case durations follow a log-normal distribution. Such prediction bounds can be calculated from only two previous cases and can be printed on the OR schedule to assist managers in the scheduling of cases.

Appendix

Distribution-free Prediction Bounds

The N previous case durations for the same scheduled surgeon and procedure are sorted in ascending order. We let t_1 refer to the shortest case duration, t_2 to the second shortest, and so forth to t_N, the longest. The upper prediction bound is one of the previous case durations, which we call t_L, $1 \leq L \leq N$. Letting $(1 - \alpha) = 0.5$, 0.8, or 0.9 to obtain 50%, 80%, or 90% prediction bounds, then[2] L equals the smallest integer, up to N, greater than or equal to $(1 - \alpha) \times (N + 1)$. Because 80% of $(4 + 1) = 4$ and 90% of $(9 + 1) = 9$, $N \geq 4$ and $N \geq 9$ previous case durations are needed to calculate 80% and 90% upper prediction bounds, respectively.

The even values of N in table 1 are 2, 4, 6, and 18. Prediction bounds calculated using the distribution-free method are expected to overestimate future case durations more often than specified for these N. For example, we consider the 50% prediction bound calculated using $N = 2$. The median of two numbers equals the mean. However, L equals the smallest integer, up to $N = 2$, greater than or equal to 50% of $(2 + 1)$ or the maximum of the two numbers. Because upper prediction bounds calculated using the distribution-free method can overestimate future case durations more often than specified[2] and the method to calculate prediction bounds based on the log-normal distribution performed well (table 1), we chose at the University of Iowa to use the method based on the log-normal distribution. An alternative strategy would be to interpolate between adjacent case durations when $(1 - \alpha) \times (N + 1)$ does not resolve to an integer. For example, when $(1 - \alpha) = 0.9$ and $N = 13$, $L = 13$ because $(1 - \alpha) \times (N + 1) = 12.6$. Instead, of using t_{13} as the 90% upper prediction bound, $t_{12} + 0.6 \times (t_{13} - t_{12})$ could be used. The limitation to using interpolation is that we are not aware of a statistical theory for the technique, and therefore successful testing of the method based on data from the University of Iowa may not apply to other OR suites.

Prediction Bounds Based on Log-normal Distribution

The $100(1 - \alpha)$% prediction bound equals[3] $\exp(\bar{T} + s \cdot t_{1-\alpha}(N - 1) \cdot \sqrt{1 + 1/N})$, where $\bar{T}$= the mean of the natural logarithms of the N previous case durations, s = the standard deviation of the natural logarithms of the N previous case durations, and $t_{1-\alpha}(N - 1)$ = the $100(1 - \alpha)^{th}$ percentile of the t distribution with $(N - 1)$ degrees of freedom. The $\sqrt{1 + 1/N}$ term is the appropriate factor when calculating a prediction bound, and differs from a $\sqrt{1/N}$ term that would have been appropriate if a confidence bound had been desired. Calculation of the standard deviation requires at least two data. Therefore, to calculate a prediction bound based on a log-normal distribution of case durations, an OR manager must have at least two previous cases.

References

1. Dexter F: Application of prediction levels to operating room scheduling. AORN J 1996; 63:607–15

2. Hahn GJ, Meeker WQ: Statistical Intervals. A Guide for Practitioners. New York, John Wiley & Sons, 1991, pp 31, 92–6

3. Dahiya RC, Guttman I: Shortest confidence and prediction intervals for the log-normal. Can J Stat 1982; 10:277–91

4. Strum DP, May JH, Vargas LG: Surgical procedure times are well modeled by the log normal distribution. Anesth Analg 1998; 86 (suppl):S47