Fifteen Years of Research on Surgical Case Duration Prediction by Combining Preoperatively Available Service and Surgeon Data

Franklin Dexter, MD, PhD, FASA
Iowa City, IA

Richard H Epstein, MD, FASA
Miami, FL

We read with interest the article by Bartek and colleagues[1] examining surgical case duration prediction. They claim their "study is novel for its scope (using a large clinical dataset spanning 4 years and >47,000 cases), practical focus (limiting the data inputs for our models to only those that are available preoperatively), and approach of developing both service-specific and surgeon-specific models." We disagree that their study is novel for its scope, focus, or approach. The authors neglected to consider any of the multiple (>12) articles from the past 15 years that incorporated the most widely used Bayesian formulation for estimating operating room times[2-6] and to compare their method's performance.

Bartek and colleagues[1] refer to their method as "practical," but from Figure 1, they excluded 28.1% of cases (n = 15,222/54,102) and limited cases to "scheduled surgeries performed on weekdays for adult patients." In contrast, the Bayesian method has been tested and implemented at multiple hospitals for every case, elective or urgent, pediatric or adult.[2-6] That is feasible because the data used are historical estimates from surgeon/scheduler, current estimate from surgeon/scheduler, and historical durations by procedure. First, linear regression is used to exclude systematic bias (eg deliberate underestimation of case duration by schedulers) if present, by specialty. Second, the method relies on the adjusted estimate from surgeon/scheduler as having incorporated Bartek's multiple covariates. The proportional variability is modeled using historical data both by surgeon and procedure and by specialty. In 2013, we published a review article on its use for comparing case durations among hospitals.[4] Implementation can be in a spreadsheet (eg Excel worksheet) or in a database table, updated annually.[6]

Bartek and colleagues[1] refer to 4 years of data as representing a novel "scope." We used 3 years of data from each of 3 hospitals to compare qualities of surgeon/schedulers' estimates.[5] Whereas Bartek and colleagues used 38,800 cases, from Table 1 of reference 5, there were much larger numbers: 65,661, 72,279, and 159,778 cases, respectively. Therefore, the "scope" of Bartek and colleagues' study was not "novel."

Unlike Bartek and colleagues' approach, the Bayesian method can be used to estimate accurately the longest time cases will take (90th percentile) and the mean time remaining in ongoing cases.[2,3] The longest expected time helps to decide whether to perform a case between 2 previously scheduled cases without delaying the start of the original to-follow case. Also, appropriate gaps can be created between different surgeons' cases in the same operating room to reduce tardiness of start times, increasing the to-follow surgeon's productivity.[6] The longest case times are also useful to help provide price transparency to patients and insurers.[4] When an ongoing case is taking longer than originally estimated, the expected value (ie mean) time remaining, recalculated as needed, is the statistic useful for making decisions about moving a to-follow case to another room, and/or assigning add-on cases to the room.[3] We think that the method described in Bartek and colleagues' article does not address any of these issues, therefore substantively limiting its practicality.[1]

REFERENCES

1. Bartek MA, Saxena RC, Solomon S, et al. Improving operating room efficiency: a machine learning approach to predict case-time duration. J Am Coll Surg 2019;229:346–354.
2. Dexter F, Ledolter J. Bayesian prediction bounds and comparisons of operating room times even for procedures with few or no historical data. Anesthesiology 2005;103:1259–1267.
3. Dexter F, Epstein RH, Lee JD, Ledolter J. Automatic updating of times remaining in surgical cases using Bayesian analysis of historical case duration data and instant messaging updates from anesthesia providers. Anesth Analg 2009;108:929–940.
4. Dexter F, Epstein RH, Bayman EO, Ledolter J. Estimating surgical case durations and making comparisons among facilities: identifying facilities with lower anesthesia professional fees. Anesth Analg 2013;116:1103–1115.
5. Dexter F, Ledolter J, Tiwari V, Epstein RH. Value of a scheduled duration quantified in terms of equivalent numbers of historical cases. Anesth Analg 2013;117:204–209.
6. Dexter F, Bayman EO, Pattillo JCS, et al. Influence of parameter uncertainty on the tardiness of the start of a surgical case following a preceding surgical case performed by a different surgeon. PCORM 2018;13:12–17.

https://doi.org/10.1016/j.jamcollsurg.2019.09.006
ISSN 1072-7515/19

Disclosure Information: Dr Dexter reports to the Division of Management Consulting of the University of Iowa's Department of Anesthesia which provides consultations to corporations, hospitals, and individuals. This includes the topic of case duration prediction, as in the current letter. Dr Dexter receives no funds personally other than his salary and allowable expense reimbursements from the University of Iowa and has tenure with no incentive program. He and his family have no financial holdings in any company related to his work, other than indirectly through mutual funds for retirement. Income from the Division's consulting work is used to fund Division research.

Accurately Predicting Case-Time Duration

In reply to Dexter and Epstein

Matthew A Bartek, MD, MPH,
Rajeev C Saxena, MD, MBA, Stuart Solomon, MD,
Christine T Fong, MS, John D Lang, MD,
Bala G Nair, PhD
Seattle, WA

Lakshmana D Behara, MS,
Ravitheja Venigandla, MS,
Kalyani Velagapudi, PhD
Bellevue, WA

We thank Drs Dexter and Epstein for their response to our article titled, "Improving operating room efficiency: a machine learning approach to predict case-time duration."[1] We appreciate and recognize their numerous contributions to the literature on applying Bayesian methods to predict case duration. However, we would encourage the authors to have a more nuanced dialogue on how our different approaches, methods, and results can enhance operating room management—rather than a broad criticism.

Drs Dexter and Epstein criticize the novelty of our scope, approach, and practicality. Simply put: their previous work seeks to adjust surgeon estimates; our methods involve the prediction of case-time estimates completely independent of surgeon input. Philosophically, these are divergent ideas: if the case-time duration can be accurately predicted, as per our approach, we can rely less on the surgeon-as-an-estimator and more on the integration of robust data to improve models. Given the current advances in data accessibility and management, our approach is especially timely and applicable to any operating room.

We will address their criticisms in 3 points. First, we use a more comprehensive dataset as compared with their studies. A key aspect is that we included a variety of preoperative patient factors into our model to account for what Dexter and Epstein have termed "process variability."[2,3] While Dexter and Epstein cite previous studies using larger case volumes, the data available in their studies included only the procedure type, surgeon, and case-time, without considering patient data. Further, they considered it a limitation that we restricted our cases to only those performed in the operating rooms (ie not in the radiology or cardiology suites) and weekday cases (ie not after-hour or emergency cases). The exclusion of these cases was deliberate and was designed with an operating room manager in mind. In fact, their work seems to have taken the same approach.[3] Therefore, we defend our claim of using a novel dataset and scope.

Second, we applied multiple, modern machine learning algorithms to predict case duration and selected the best performing algorithm to build our final set of models. On the contrary, Dexter and Epstein have applied a Bayesian method to estimate case time durations. Traditional statistical methods are unable to derive patterns from large datasets. Dr Dexter's website bibliography lists 2 other papers that have used approaches similar to ours, both of which were published recently and neither of which included case volumes that approached that of our dataset.[4-6] Therefore, we defend our statement that ours was a novel approach.

Third, we have kept a keen focus throughout our work on implementation. We acknowledge that our machine learning models did not include predictions for longest expected time, anticipated gaps, or expected time remaining. However, these are certainly achievable with a machine learning approach combined with a robust dataset. We believe that our approach serves a different objective than that of Drs Dexter and Epstein. Our goal was to accurately predict case-time durations based on preoperative patient procedure and personnel data to optimize surgical scheduling. For that reason, and unlike several previous studies, we excluded any procedural data that were exclusively available intraoperatively or postoperatively, such as CPT codes. We designed both service-specific models and surgeon-specific models to generate predictions that govern the majority of cases in a hospital. We have already begun piloting a tool to improve the case-time estimates at our institution that relies both on accurate and appropriate data collection and on the machine learning models to interpret those data.

We intend for this work to be an early stepping-stone to realizing the dual potential of modern electronic medical records and modern "big data" computational