# Clinical Prediction of Postoperative Respiratory Failure

CLINICAL research in Anesthesiology increasingly focuses on prediction of myriad perioperative complications. This research has been driven by development of databases that capture various measures of patient health, anesthetic care, surgical procedure, and outcomes. One of the strengths of this approach is that it allows analysis of outcomes that were previously understudied due to their rarity or complexity. Postoperative respiratory failure is a perfect target for such researches because it is relatively rare and has a significant impact on healthcare costs[1] and patient mortality.[2,3] Previous investigators have developed several screening tools for the prediction of postoperative respiratory failure,[2–4] but most have focused on 30-day outcomes, a period that may well exceed the immediate influence of anesthesia technique and perioperative respiratory management. For example, the first 24 h after surgery represent the highest risk of unanticipated respiratory failure due to opioids,[5,6] whereas postoperative hypoxemia has been shown to peak by the third night after major surgery.[7–10] In this issue of the Journal, Eikermann *et al.* report the development and validation of a Score for Prediction of Postoperative Respiratory Complications (SPORC) focusing on the early postoperative period of 3 days after surgery.[11]

The investigators identified several independent predictors for reintubation such as planned postoperative hospital admission, preoperative history of congestive heart failure, chronic pulmonary or cerebrovascular disease, emergency surgery, American Society of Anesthesiologists score of 3 or more, and high-risk surgical service. By using a weighted point system, the SPORC yielded a calculated area under the receiver operating characteristics curve of 0.84–0.87, with a step-wise increase in the odds for reintubation with increasing number of risk factors. As previously reported, the development of respiratory failure was associated with a large increase in 30-day mortality.[2] The SPORC tool is, thus, a
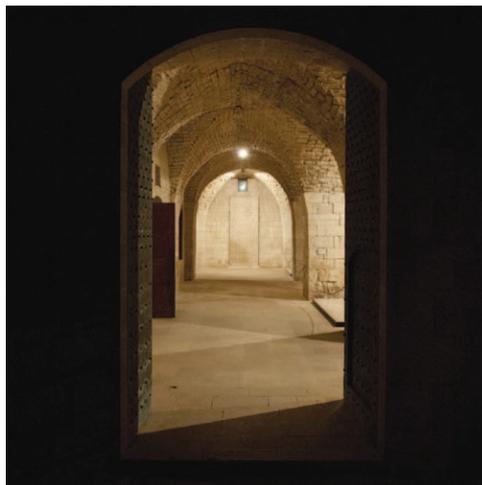
"[In developing this tool,] Eikermann et al. … have lit a path of discovery toward greater efficiency of screening for postoperative respiratory failure."

simple way to identify high-risk patients in future studies and prospectively evaluate the effectiveness of interventions in preventing or reducing the incidence and severity of postoperative respiratory failure. For example, success of continuous positive airway pressure therapy in patients recovering from major abdominal surgery suggests that screening tool such as SPORC may have a role in identifying the patients who are likely to benefit from this therapy.[12] There are, however, several limitations to the authors' approach to screening.

Bayes Theorem describes the relation between the prevalence of a disease and the accuracy of prediction tools.[13–15] For rare events, even highly accurate tests will generate many false positives, with the potential consequences of excess resource utilization and complications from unnecessary treatment. Most prediction models derived from outcomes databases have positive predictive values less than 10% (low clinical precision) as the outcomes of interest are typically rare (0.1–4%).[2,16] As most patients who underwent surgery have low risk and will not develop the complication in question, high specificity is given; but, because only a small fraction of patients will screen positively as high-risk, the sensitivity typically tends to be low. A tool with sensitivity of less than 50% will, by definition, fail to identify the majority of patients who will develop the complication. This is the case with the majority of prediction models derived from outcomes databases for perioperative complications (low sensitivity and low technical precision). A consequence is that policies or care processes that preferentially allocate treatments to high-risk patients, based on these screening tools, may potentially place more patients at harm due to misdiagnoses (*i.e.,* patients who should be given treatment do not receive it). Furthermore, heterogeneity in patient populations is high in most prediction models, resulting in significant variability

in prediction accuracy.[14,17] The inherent heterogeneity and lack of technical or clinical precision are thus permanent limitations of screening tools designed to detect high risk of rare outcomes. Other challenges with rare outcomes from large databases include concerns over data reliability, missing data, unmeasured biases, unmeasured treatment effects, and lack of generalizability. These concerns apply broadly and are certainly not unique to Eikermann's work.

The impact of clinical screening tools needs to be considered in the context of the population to which it is applied and the clinicians who will use the tool.[14,15] Screening is one of several steps in the perioperative care of patients that could influence outcomes. The bevy of outcomes research-based screening tools is an emerging point to one common finding, *i.e.,* sicker patients do worse after surgery! More interestingly, several of the independent predictors are common across multiple studies of diverse outcomes. Because these diseases are coded in a binary fashion, their severity spectrum or modifiability remains untested. Should patients be investigated more intensively when there is high expected risk of complications? Should surgical techniques be determined by the estimated risk of complications? Indeed, several care processes that significantly improve outcomes have been described for patients with specific conditions.[18] An elegant way to assess prediction tools is to use a hierarchical model of diagnostic test effectiveness.[19] The hierarchical model describes the impact of a test or screening tool on patient outcomes, where screening tests that address the higher level questions have greater clinical impact. It is imperative to recognize that a great screening tool is unlikely to change outcomes unless specific interventions exist that modify the risk of the adverse outcome themselves.[20,21] The most important question among clinicians remains "What should I do differently to prevent postoperative respiratory failure?" Thus, to answer that specific question, iterative steps are needed to traverse the levels of test efficiency from accurate disease screening to identification of patients most suited for effective therapy. In the process of achieving greater screening test effectiveness, we need to be wary of these existing statistical traps.

One way to overcome the stated permanent limitations could be to develop two-level screening tools, with an initial high-sensitivity test followed by a test with high-positive predictive value. This approach has been particularly effective in screening for disease states such as obstructive sleep apnea, where determination of high risk based on a questionnaire (high sensitivity) followed by overnight home oximetry (high predictive value) provides highly accurate prediction.[22] The countermeasure for low sensitivity of prediction tools is to cast a wider net that includes all possible conditions that influence risk of respiratory failure (or other clinical outcomes of interest). This approach comes with an attendant drop in the positive predictive value. The addition of a second-level screening process, which identifies the more severe forms of specific contributory diseases, could increase the positive predictive value of the tool in a subsection of the study population. For

example, second-level testing might evaluate lung volumes, exercise tolerance, radiographic measures, or other functional measures that are shown to increase the overall predictive value.

A significant strength of Eikermann's study is that it elucidates specific risk conditions, which become targets for mechanistic research, to help us understand how these conditions influence the outcomes. With the development of the SPORC screening tool, Eikermann *et al.* have targeted an important perioperative complication that markedly increases mortality and cost of health care. In doing so, they have lit a path of discovery toward greater efficiency of screening for postoperative respiratory failure.

Amy Young, M.D., Satya Krishna Ramachandran, M.D., F.R.C.A., Department of Anesthesiology, University of Michigan Medical School, Ann Arbor, Michigan. rsatyak@med.umich.edu

References

1. Dimick JB, Chen SL, Taheri PA, Henderson WG, Khuri SF, Campbell DA Jr: Hospital costs associated with surgical complications: A report from the private-sector National Surgical Quality Improvement Program. J Am Coll Surg 2004; 199:531–7

2. Ramachandran SK, Nafiu OO, Ghaferi A, Tremper KK, Shanks A, Kheterpal S: Independent predictors and outcomes of unanticipated early postoperative tracheal intubation after nonemergent, noncardiac surgery. ANESTHESIOLOGY 2011; 115:44–53

3. Johnson RG, Arozullah AM, Neumayer L, Henderson WG, Hosokawa P, Khuri SF: Multivariable predictors of postoperative respiratory failure after general and vascular surgery: Results from the patient safety in surgery study. J Am Coll Surg 2007; 204:1188–98

4. Arozullah AM, Daley J, Henderson WG, Khuri SF: Multifactorial risk index for predicting postoperative respiratory failure in men after major noncardiac surgery. The National Veterans Administration Surgical Quality Improvement Program. Ann Surg 2000; 232:242–53

5. Taylor S, Kirton OC, Staff I, Kozol RA: Postoperative day one: A high risk period for respiratory events. Am J Surg 2005; 190:752–6

6. Ramachandran SK, Haider N, Saran KA, Mathis M, Kim J, Morris M, O'Reilly M: Life-threatening critical respiratory events: A retrospective study of postoperative patients found unresponsive during analgesic therapy. J Clin Anesth 2011; 23:207–13

7. Rosenberg J, Ullstad T, Larsen PN, Moesgaard F, Kehlet H: Continuous assessment of oxygen saturation and subcutaneous oxygen tension after abdominal operations. Acta Chir Scand 1990; 156:585–90

8. Gögenur I, Wildschiøtz G, Rosenberg J: Circadian distribution of sleep phases after major abdominal surgery. Br J Anaesth 2008; 100:45–9

9. Rosenberg J, Wildschiødtz G, Pedersen MH, von Jessen F, Kehlet H: Late postoperative nocturnal episodic hypoxaemia and associated sleep pattern. Br J Anaesth 1994; 72:145–50

10. Rosenberg J, Ullstad T, Rasmussen J, Hjørne FP, Poulsen NJ, Goldman MD: Time course of postoperative hypoxaemia. Eur J Surg 1994; 160:137–43

11. Brueckmann B, Uribe JV, Bateman BT, Grosse-Sandrup M, Hess DR, Schlett CL, Eikermann M: Development and validation of a score for prediction of postoperative respiratory complications. ANESTHESIOLOGY 2013; 118:1276–85

12. Squadrone V, Coha M, Cerutti E, Schellino MM, Biolino P, Occella P, Belloni G, Vilianis G, Fiore G, Cavallo F, Ranieri

VM; Piedmont Intensive Care Units Network (PICUN): Continuous positive airway pressure for treatment of postoperative hypoxemia: A randomized controlled trial. JAMA 2005; 293:589–95

13. Leeflang MM, Bossuyt PM, Irwig L: Diagnostic test accuracy may vary with prevalence: Implications for evidence-based diagnosis. J Clin Epidemiol 2009; 62:5–12

14. Moons KG, van Es GA, Deckers JW, Habbema JD, Grobbee DE: Limitations of sensitivity, specificity, likelihood ratio, and bayes' theorem in assessing diagnostic probabilities: A clinical example. Epidemiology 1997; 8:12–7

15. Mileman PA, van den Hout WB: Evidence-based diagnosis and clinical decision making. Dentomaxillofac Radiol 2009; 38:1–10

16. Ramachandran SK, Mathis MR, Tremper KK, Shanks AM, Kheterpal S: Predictors and clinical outcomes from failed Laryngeal Mask Airway Unique™: A study of 15,795 patients. ANESTHESIOLOGY 2012; 116:1217–26

17. Ramachandran SK, Josephs LA: A meta-analysis of clinical screening tests for obstructive sleep apnea. ANESTHESIOLOGY 2009; 110:928–39

18. Olsson LE, Karlsson J, Ekman I: The integrated care pathway reduced the number of hospital days by half: A prospective comparative study of patients with acute hip fracture. J Orthop Surg Res 2006; 1:3

19. Fryback DG, Thornbury JR: The efficacy of diagnostic imaging. Med Decis Making 1991; 11:88–94

20. Pinsky MR: Hemodynamic evaluation and monitoring in the ICU. Chest 2007; 132:2020–9

21. Pinsky MR, Payen D: Functional hemodynamic monitoring. Crit Care 2005; 9:566–72

22. Sharma SK, Kurian S, Malik V, Mohan A, Banga A, Pandey RM, Handa KK, Mukhopadhyay S: A stepped approach for prediction of obstructive sleep apnea in overtly asymptomatic obese subjects: A hospital based study. Sleep Med 2004; 5:351–7