Published in final edited form as:

Anesth Analg. 2014 August ; 119(2): 237–238. doi:10.1213/ANE.0000000000000324.

Making Sense of Surgical Risk When the Data Aren't Perfect

Meghan B. Lane-Fall, MD, MSc and Mark D. Neuman, MD, MSc

Department of Anesthesiology and Critical Care, University of Pennsylvania, Philadelphia, Pennsylvania

Informed consent discussions, performance measurement, and studies of patient outcomes all require accurate information regarding the likelihood of a good or bad outcome after surgery for an individual patient. Statistical risk models, typically based on multivariate analyses of retrospective data, aim to help clinicians quantify the probability of adverse outcomes for individual patients. Beginning with Lee Goldman et al.'s 1977 landmark publication of a multivariate risk index for postoperative cardiac events,¹ statistical risk models have been used for more than three decades in various surgical contexts prospectively to predict the likelihood of adverse outcomes for patients considering surgical interventions.² Predictions derived from such models can then be used to help decide whether to have surgery (e.g., medical versus surgical management of hip fracture), what kind of surgery to have (e.g., conventional versus transcatheter aortic valve replacement), or where to have surgery (e.g., outpatient surgicenter versus hospital).

In predicting risk before surgery, one size does not necessarily fit all, different risk factors might be more or less important for different groups of patients. Perhaps for this reason, different approaches to predicting risk have emerged that are specific to certain surgical specialties and specific procedures. For example, the Society of Thoracic Surgeons (STS) risk score is used to predict operative mortality for patients undergoing cardiothoracic surgery.³ In addition to being a helpful research tool for comparing patient populations, the STS score has emerged as a decision point in determining the management approach for symptomatic aortic stenosis. For instance, in the PARTNER trial, an STS score of 10% was used to define a high-risk surgical group.⁴ Only those patients at or above this cutoff were considered for transcatheter aortic valve replacement. Other surgical specialties employing procedure-specific risk models include orthopedics,⁵ and transplant surgery.⁶

Corresponding Author: Meghan B. Lane-Fall, MD, MSc, Department of Anesthesiology and Critical Care, University of Pennsylvania, 423 Guardian Drive 1121-B Blockley Hall Philadelphia, PA 19104, Phone: 215-573-7399, FAX: 215-662-7106, meghan.lane-fall@uphs.upenn.edu.

DISCLOSURES:

Name: Meghan B. Lane-Fall, MD, MSc

Contribution: This author helped write the manuscript

Attestation: Meghan B. Lane-Fall approved the final manuscript

Name: Mark D. Neuman, MD, MSc

Contribution: This author helped write the manuscript

Attestation: Mark D. Neuman approved the final manuscript

This manuscript was handled by: Gregory J. Crosby, MD

The authors declare no conflicts of interest.

In this issue of *Anesthesia & Analgesia*, Reponen and colleagues review evidence on preoperative risk stratification for patients undergoing cranial neurosurgery.⁷ The authors compared commonly used risk scores by conducting a qualitative systematic review. From a starting sample of more than 2000 papers, Reponen et al. analyzed 25 published reports describing outcomes after craniotomy as a function of five different scales: American Society of Anesthesiologists Physical Status classification (ASA; 10 studies), the Karnofsky Performance Score (KPS; 16 studies), the Charlson Comorbidity Score (CCS; 3 studies), the modified Rankin Scale (mRS; 2 studies), and the Sex, KPS, ASA classification, meningioma Location, and peritumoral Edema score (SKALE; 2 studies). The authors found evidence of some association among three of these scores and postoperative outcomes: ASA physical status was associated with mortality and with infectious complications including surgical site infection and postoperative meningitis. The KPS was associated with mortality, functional impairment and systemic complications, including pneumonia, heart failure, and deep venous thrombosis. The CCS was associated with mortality in patients who had elective treatment of intracranial aneurysms and in elderly patients with intracranial tumors.⁷ As Reponen and colleagues note, important differences in study design, patient populations, and outcome definitions across studies limit the conclusions that can be drawn from this review. The 25 studies they reviewed included retrospective and prospective designs, sample sizes ranging from 30 to almost 5000 patients, and varying definitions of operative and non-operative complications. Indeed, the authors note that they were unable to perform a quantitative meta-analysis (as opposed to a narrative review) because of marked differences in the methods and outcomes used across studies.

The work of Reponen and colleagues offers a useful starting place for further research and discussion regarding prediction of outcomes after surgery. Importantly, it highlights the importance of critically examining the individual variables contained within any risk model; ideally, a plausible physiologic mechanism should link the individual elements of a risk score to the outcome of interest. In terms of postoperative mortality, for example, increasing age and comorbidity burden might predict increased mortality. In other cases, the relationship may be less apparent. For example, the data suggest that patients with a higher ASA physical status classification may be more likely to experience postoperative infections. While this association may be supported by observations in the data, further research is likely to be required to tease out the specific mechanisms that link poor physical status to infectious outcomes and to develop effective interventions to decrease risk among future patients.

Reponen et al.'s review also highlights challenges in incorporating variables into risk scores that rely on individual clinician judgment. Indeed, the Goldman Index, and statistical risk modeling as a whole, grew out of a concern that existing tools for patient assessment, such as the ASA physical status classification, were prone to subjectivity due to their reliance on individual clinician's judgments. Yet when it has been included in statistical risk models, the ASA physical status classification turns out to be a tremendously useful predictor of postoperative outcomes specifically because it takes advantage of physicians' abilities to judge how sick an individual patient is, above and beyond the comorbidities that appear in a patient's chart. For example, a recent review of heterogeneous patient populations found

that, despite poor interrater reliability, ASA physical status was an accurate predictor of postsurgical mortality.⁸ Nonetheless, reliance on human judgment in scoring patient risk raises important concerns regarding the potential for risk assessments, such as the ASA physical status classification, to vary across raters and to be overly influenced by individual biases.

Of the five scores that Reponen and colleagues reviewed, four had components that require some degree of judgment on the part of the assessing clinician. As such, the paper highlights opportunities for further research on opportunities to increase the reliability of such assessments across clinicians and settings, and how such clinical assessments can potentially be incorporated into large-scale prospective data collection efforts. Such efforts are ongoing outside of anesthesia through the Veterans Affairs and American College of Surgeons' National Surgical Quality Improvement Programs (NSQIP⁹ and ACS-NSQIP¹⁰ and within anesthesiology through the Anesthesia Quality Institute's National Anesthesia Clinical Outcomes Registry,¹¹ and the University of Michigan's Multicenter Perioperative Outcomes Group.¹² Ultimately, data from these and other registries may carry potential not only for the development of additional risk prediction models for neurosurgery and other specialties, but also shed light on how the collection and data on specific risk predictors can be more effectively standardized going forward.

The work of Reponen et al. make it clear that further work is needed to develop reliable and robust approaches to predicting risk for neurosurgical patients. In the meantime, however, there are patients and clinicians who need guidance as they consider therapeutic options. For patients and clinicians, the work of Reponen et al confirms that that worst functional status and greater degrees of comorbidity are likely to be associated with perioperative morbidity and mortality in craniotomy patients,⁷ and that further work is still needed to identify the specific combinations of factors that most accurately predict poor outcomes. In the meantime, then, perhaps we should consider the risk scores reviewed by Reponen and colleagues to be places to start in our conversations about operative risk among neurosurgical patients, but not the ultimate arbiters of therapeutic choices.

Acknowledgments

Funding: This work was unfunded. Mark Neuman receives salary support from the National Institutes of Health (grant number 5K08AG043548).

References

1. Goldman L, Caldera DL, Nussbaum SR, Southwick FS, Krogstad D, Murray B, Burke DS, O'Malley TA, Goroll AH, Caplan CH, Nolan J, Carabello B, Slater EE. Multifactorial index of cardiac risk in noncardiac surgical procedures. *N Engl J Med.* 1977; 297:845–50. [PubMed: 904659]
2. Neuman MD, Bosk CL. What we talk about when we talk about risk: refining surgery's hazards in medical thought. *Milbank Q.* 2012; 90:135–59. [PubMed: 22428695]
3. Shroyer ALW, Coombs LP, Peterson ED, Eiken MC, DeLong ER, Chen A, Ferguson TB Jr, Grover FL, Edwards FH, Baskett RJF, Chen JC. The society of thoracic surgeons: 30-Day operative mortality and morbidity risk models. *Ann Thorac Surg.* 2003; 75:1856–65. [PubMed: 12822628]
4. Leon MB, Smith CR, Mack M, Miller DC, Moses JW, Svensson LG, Tuzcu EM, Webb JG, Fontana GP, Makkar RR, Brown DL, Block PC, Guyton RA, Pichard AD, Bavaria JE, Herrmann HC,

- Douglas PS, Petersen JL, Akin JJ, Anderson WN, Wang D, Pocock S. Transcatheter Aortic-Valve Implantation for Aortic Stenosis in Patients Who Cannot Undergo Surgery. *N Engl J Med*. 2010; 363:1597–607. [PubMed: 20961243]
5. Maxwell MJ, Moran CG, Moppett IK. Development and validation of a preoperative scoring system to predict 30 day mortality in patients undergoing hip fracture surgery. *Br J Anaesth*. 2008; 101:511–7. [PubMed: 18723517]
 6. Onaca NN, Levy MF, Sanchez EQ, Chinnakotla S, Fasola CG, Thomas MJ, Weinstein JS, Murray NG, Goldstein RM, Klintmalm GB. A correlation between the pretransplantation MELD score and mortality in the first two years after liver transplantation. *Liver Transplantation*. 2003; 9:117–23. [PubMed: 12548503]
 7. Reponen E, Tuominen H, Korja M. Evidence for the use of preoperative risk assessment scores in elective cranial neurosurgery - a systematic review of the literature. *Anesth Analg*. 2014 IN THIS ISSUE.
 8. Moonesinghe SR, Mythen MG, Das P, Rowan KM, Grocott MPW. Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: Qualitative systematic review. *Anesthesiology*. 2013; 119:959–81. [PubMed: 24195875]
 9. Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, Chong V, Fabri PJ, Gibbs JO, Grover F, Hammermeister K, Irvin G Iii, McDonald G, Passaro E Jr, Phillips L, Scamman F, Spencer J, Stremple JF. The Department of Veterans Affairs' NSQIP: The first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. *Ann Surg*. 1998; 228:491–507. [PubMed: 9790339]
 10. Fink AS, Campbell DA Jr, Mentzer RM Jr, Henderson WG, Daley J, Bannister J, Hur K, Khuri SF. The National Surgical Quality Improvement Program in non-veterans administration hospitals: Initial demonstration of feasibility. *Ann Surg*. 2002; 236:344–54. [PubMed: 12192321]
 11. Dutton RP, DuKatz A. Quality Improvement Using Automated Data Sources: The Anesthesia Quality Institute. *Anesthesiology clinics*. 2011; 29:439–54. [PubMed: 21871404]
 12. Kheterpal S. Clinical research using an information system: The multicenter perioperative outcomes group. *Anesthesiology clinics*. 2011; 29:377–88. [PubMed: 21871400]