

# Predicting In-hospital Postoperative Mortality for the Practitioner

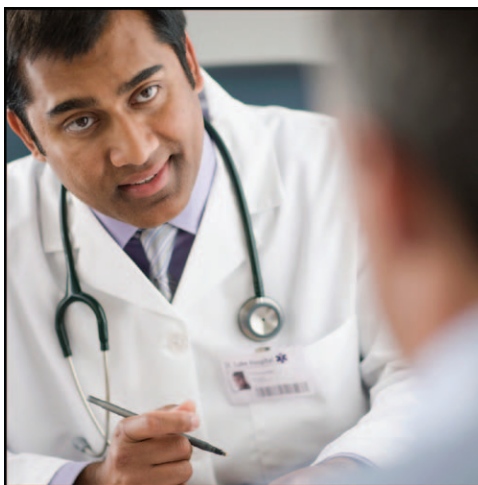
## Beyond the Numbers

Edward J. Mascha, Ph.D., Valentín Mazo, M.D.

MANAGEMENT of patients undergoing surgery has dramatically improved over the past 150 yr. Instrumental have been the advent and development of anesthesia, antisepsis, antibiotics, and progress in critical care, to name but a few. But the aging population along with the increasing complexity and higher cost of procedures coupled with the finiteness of resources represent a concern for the future. Every postoperative death, whatever the overall incidence might be, makes us wonder about complications, errors, or preventability. Some questions that never fail: Was the surgery advisable or necessary in the first place? Was the patient properly informed of the risks? Were efforts spared in his or her care? In any case, the need to properly calibrate the risk/benefit ratio before each surgery, as opposed to an average across all surgeries or those of a particular type, is undeniable.

In this issue of *ANESTHESIOLOGY*, Le Manach *et al.*<sup>1</sup> propose a predictive model of in-hospital postoperative mortality. Data for model development and validation include surgical procedures in France receiving care of an anesthesiologist (with or without anesthesia being given), with results possibly extrapolable to other countries in Western Europe. We wish to emphasize the relatively low incidence of mortality, close to 0.5%. This incidence is somewhat lower than most published reports to date,<sup>1-3</sup> possibly explained by the high proportion of less aggressive procedures and outpatient surgeries.

A highly specific definition of the outcome is essential to the development, interpretation, and comparison of predictive models. Recent disputes concerning the rates of



***“Every patient deserves and has the right to as much information as possible about their vulnerability before undergoing a surgical procedure.”***

postoperative mortality involve this definition and also the representativeness of the samples.<sup>4,5</sup> Of note, Le Manach *et al.*<sup>1</sup> have chosen mortality before hospital discharge instead of the more common 30 postoperative days. Either choice has advantages and disadvantages, but methodological consensus across the studies for the sake of better interpretation of results is desirable.

In developing their prediction model, Le Manach *et al.* appropriately include presurgical and procedure variables and not intraoperative or postsurgical variables, even though adding the latter would have improved prediction. Including only baseline variables creates a model which can be used to anticipate and plan for patient requirements during the surgery and postoperatively. Choice of specific candidate predictors (age, comorbidities, and surgical procedures)

supported by internationally agreed upon coding is methodologically strong. For supporters of parsimonious models, 17 variables may seem excessive. However, we believe that for data easily collected in the preoperative evaluation, even small improvements in model discrimination (*i.e.*, ability to separate events from nonevents) are worth the effort. Finally, clinicians often request cutpoints within a predictive scale to support decision-making, but this categorization of risk can be arbitrary; decision-making should instead be based, for example, on resources available to avoid the outcome, rather than on a uniform “probability of mortality” criterion for all patients. Therefore, we consider it appropriate that the authors have not proposed a categorization of the probability of dying to use in decision-making.

Image: ©Thinkstock.

Corresponding article on page 570.

Accepted for publication November 2, 2015. From the Departments of Quantitative Health Sciences and Outcomes Research, Cleveland Clinic, Cleveland, Ohio (E.J.M.); and Department of Anesthesiology and Postoperative Care Unit, Hospital Universitari Germans Trias i Pujol, Badalona, Barcelona, Spain (V.M.).

Copyright © 2015, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. All Rights Reserved. *Anesthesiology* 2016; 124:523-5

Instead of directly using their regression model to predict outcome, as is often done, the authors chose to create a more user-friendly points system following the methods of Sullivan *et al.*<sup>6</sup> The main advantage of the chosen approach is that points given for any of the risk factors are directly comparable because they are standardized to an easily understandable risk (*i.e.*, that associated with an increase of 5 yr of age).

Points are summed for a patient across all comorbidities, demographics, and procedure(s) that the patient might have and then a patient's total points are referred to a table to obtain the estimated probability of in-hospital mortality (see Supplemental Digital Content 6: PreOperative Score to predict PostOperative Mortality [POSPOM] scoring system in the study by Le Manach *et al.*). Inevitably, some information is lost due to rounding error when using an integer point system and also by inclusion of categories for continuous variables such as age. Therefore, we commend Le Manach *et al.* for completing their validity assessment (as done in the study by Sullivan *et al.*) by showing that predictions from the POSPOM scoring system agree well with prediction from the actual regression model upon which it was based.

In exemplary manner, Le Manach *et al.* validate their model on data from a randomly chosen set of hospitals not used in the creation of the model, as opposed to the common practice of validating on a random sample of patients from the population used in the model creation. Validation on a random sample of patients from the same population naturally leads to overly optimistic (and thus biased) predictive ability, whereas applying the model to a completely different set of patients, as done by the authors, shows more true predictive ability. Although external validation is still needed, the authors have thus made a good attempt at assessing the practical utility of their predictive scoring system. Also in exemplary manner, Le Manach *et al.* assess the two key elements of a good prediction model—discrimination (how well model separates those with and without the event) and calibration (how well the model fits the data), with the model scoring high on both.

Because many fewer predictors are needed, the proposed model and resulting points system is more practical than existing models for predicting postoperative mortality.<sup>7,8</sup> An important distinction is that the proposed scoring system is intended for individual patient prediction, whereas previous models were intended more for confounding adjustment when comparing exposures or hospitals and thus did not need to be “clinician-friendly.” Still, it would be interesting to directly compare the predictive ability of the proposed method to the previously established more detailed models on their predictive ability in a common population.

At this point, someone might ask captiously: What is the usefulness of estimating a patient's probability of dying in the hospital after a surgical procedure? We think that there are several reasons to toast the birth of this promising predictive model:

1. *Every patient deserves and has the right to as much information as possible* about their vulnerability before undergoing a surgical procedure. Everyone understands what it means to have an estimated 20% chance of dying. Such estimates should ideally include CIs that give the provider and patient an expected range for the underlying probability (*e.g.*, 20% chance  $\pm$  5%). However, this information should be adequately contrasted with the chance expected without surgery, which might be higher in some cases. Moreover, in the preoperative visit, clinicians should provide information to patients of risks other than dying. Some patients with limited life expectancy may prefer not to undergo procedures involving the risk of substantial suffering afterward.
2. Knowledge of risk surely helps to safeguard the patient, allowing *maximum protection and care for the most vulnerable*. Modern computing facilitates decision-support systems, which can apply a prediction model and attempt to identify at-risk patients on the spot. However, it is difficult to establish relevant cutpoints in the risk probability curve to aid decision-making, as these benchmarks depend on the threshold probability of the disease or event at which the patient would want intervention, the resources available in each hospital setting and the measures that the scientific community have proven effective for decreasing the incidence of a deleterious event. As Vickers and Elkin<sup>9</sup> have shown, decision curve analysis can be applied to evaluate the net benefit of implementing an intervention as a function of increasing levels of risk. A model such as that developed by Le Manach *et al.* could well be the starting point for such analyses.
3. Assuming a valid model, quantifying the estimated probability of dying after surgery *versus* what is expected by the model can be a *benchmark of healthcare quality* used to compare institutions. However, it is also true that a persistent gap between predictions and observed results across many settings would suggest reconsidering the validity of the model. Historical and geographical external replication and validation is a dynamic process that never ends. For example, every good predictive model of an undesirable event is doomed to have a short life span if its use (happily) leads to implementation of preventive measures, which in turn leads to the need for more relevant models.
4. In research, this new risk score for postoperative mortality could *help in the design of clinical trials* testing the effectiveness of preventive measures in selected groups of patients according to their expected risk.

Considering the excellent performance demonstrated by the model presented by Le Manach *et al.*<sup>1</sup> after internal validation, the model clearly merits replication in other geographical and temporal environments in order to determine the transportability and generalizability of the

prediction.<sup>10</sup> That responsibility does not particularly correspond to the authors because the more independent the replication is, the more solid will be the evidence of external validation.<sup>11,12</sup> Therefore, it is time to encourage the scientific community to test whether the results recorded in France remain valid when the model is applied in other countries with similar or different resources and health strategies.

### Competing Interests

The authors are not supported by, nor maintain any financial interest in, any commercial activity that may be associated with the topic of this article.

### Correspondence

Address correspondence to Dr. Mascha: maschae@ccf.org

### References

1. Le Manach Y, Collins G, Rodseth R, Le Bihan-Benjamin C, Biccard B, Riou B, Devereaux PJ, Landais P: A PreOperative Score to predict PostOperative Mortality (POSPOM): Derivation and validation. *ANESTHESIOLOGY* 2016; 124: 570–9
2. Semel ME, Lipsitz SR, Funk LM, Bader AM, Weiser TG, Gawande AA: Rates and patterns of death after surgery in the United States, 1996 and 2006. *Surgery* 2012; 151:171–82
3. Fecho K, Lunney AT, Boysen PG, Rock P, Norfleet EA: Postoperative mortality after inpatient surgery: Incidence and risk factors. *Ther Clin Risk Manag* 2008; 4:681–8
4. Brodner G, Van Aken H: Mortality after surgery in Europe. *Lancet* 2013; 381:370
5. Pearse R, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, Vallet B, Vincent JL, Hoeft A, Rhodes A: Mortality after surgery in Europe—Authors' reply. *Lancet* 2013; 381:370–1
6. Sullivan LM, Massaro JM, D'Agostino RB Sr: Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med* 2004; 23:1631–60
7. Sessler DI, Sigl JC, Manberg PJ, Kelley SD, Schubert A, Chamoun NG: Broadly applicable risk stratification system for predicting duration of hospitalization and mortality. *ANESTHESIOLOGY* 2010; 113:1026–37
8. Dalton JE, Glance LG, Mascha EJ, Ehrlinger J, Chamoun N, Sessler DI: Impact of present-on-admission indicators on risk-adjusted hospital mortality measurement. *ANESTHESIOLOGY* 2013; 118:1298–306
9. Vickers AJ, Elkin EB: Decision curve analysis: A novel method for evaluating prediction models. *Med Decis Making* 2006; 26:565–74
10. Eisenach JC, Houle TT: Replication to advance science: Changes in *ANESTHESIOLOGY*. *ANESTHESIOLOGY* 2014; 121:209–11
11. Justice AC, Covinsky KE, Berlin JA: Assessing the generalizability of prognostic information. *Ann Intern Med* 1999; 130:515–24
12. Altman DG, Royston P: What do we mean by validating a prognostic model? *Stat Med* 2000; 19:453–73