

Preoperative Score to Predict Postoperative Mortality (POSPOM)

Derivation and Validation

Yannick Le Manach, M.D., Ph.D., Gary Collins, Ph.D., Reitze Rodseth, M.B.Ch.B., Ph.D., Christine Le Bihan-Benjamin, M.D., M.Sc., Bruce Biccard, M.B.Ch.B., Ph.D., Bruno Riou, M.D., Ph.D., P.J. Devereaux, M.D., Ph.D., Paul Landais, M.D., Ph.D.

ABSTRACT

Background: An accurate risk score able to predict in-hospital mortality in patients undergoing surgery may improve both risk communication and clinical decision making. The aim of the study was to develop and validate a surgical risk score based solely on preoperative information, for predicting in-hospital mortality.

Methods: From January 1, 2010, to December 31, 2010, data related to all surgeries requiring anesthesia were collected from all centers (single hospital or hospitals group) in France performing more than 500 operations in the year on patients aged 18 yr or older (n = 5,507,834). *International Statistical Classification of Diseases*, 10th revision codes were used to summarize the medical history of patients. From these data, the authors developed a risk score by examining 29 preoperative factors (age, comorbidities, and surgery type) in 2,717,902 patients, and then validated the risk score in a separate cohort of 2,789,932 patients.

Results: In the derivation cohort, there were 12,786 in-hospital deaths (0.47%; 95% CI, 0.46 to 0.48%), whereas in the validation cohort there were 14,933 in-hospital deaths (0.54%; 95% CI, 0.53 to 0.55%). Seventeen predictors were identified and included in the PreOperative Score to predict PostOperative Mortality (POSPOM). POSPOM showed good calibration and excellent discrimination for in-hospital mortality, with a *c*-statistic of 0.944 (95% CI, 0.943 to 0.945) in the development cohort and 0.929 (95% CI, 0.928 to 0.931) in the validation cohort.

Conclusion: The authors have developed and validated POSPOM, a simple risk score for the prediction of in-hospital mortality in surgical patients. (ANESTHESIOLOGY 2016; 124:570-9)

WORLDWIDE, more than 230 million major surgical procedures are carried out each year.¹ Although in-hospital death after surgery occurs infrequently in the general population,^{2,3} among particular subgroups of patients, the rate can be greater than 1 in 10.^{4,5} In the light of this substantial mortality risk, it is important that the patient, their family, and the attending physician are able to accurately and objectively predict the preoperative risk (probability) of in-hospital mortality or the risk of developing other major perioperative complications (*e.g.*, myocardial infarction, heart failure, or sudden cardiac death). Accurate risk prediction therefore becomes crucial for communicating operative risk to patients, guiding clinical decision making

What We Already Know about This Topic

- An accurate risk score for in-hospital mortality is needed to guide future research, as well as preoperative decision making.

What This Article Tells Us That Is New

- This multicenter study examining in-hospital mortality in over 5.5 million patients in France in a 1-yr period identified a 17-variable, highly sensitive, and specific risk calculator for in-hospital mortality

and management, and for forming realistic expectations of the value of undergoing surgery. Other uses include provider profiling as risk adjustment to account for differences

This article is featured in "This Month in Anesthesiology," page 1A. Corresponding article on page 523. Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are available in both the HTML and PDF versions of this article. Links to the digital files are provided in the HTML text of this article on the Journal's Web site (www.anesthesiology.org).

Submitted for publication December 15, 2014. Accepted for publication October 21, 2015. Corrected on February 16, 2017. From the Department of Anesthesia (Y.L.M.), Department of Clinical Epidemiology and Biostatistics (Y.L.M., P.J.D.), and Department of Medicine (P.J.D.), Faculty of Health Sciences, Michael DeGroote School of Medicine, McMaster University, Hamilton, Canada; Perioperative Research Group, Population Health Research Institute, Hamilton, Canada (Y.L.M., R.R.); Centre for Statistics in Medicine, Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom (G.C.); Perioperative Research Group, Department of Anaesthesia, University of KwaZulu-Natal, Pietermaritzburg, South Africa (R.R.); Department of Biostatistics and Medical Information, Necker University Hospital, Assistance Publique Hôpitaux de Paris, France (C.L.B.-B.); Perioperative Research Group, Department of Anaesthetics, Nelson R Mandela School of Medicine, University of KwaZulu-Natal, Pietermaritzburg, South Africa (B.B.); Department of Emergency Medicine and Surgery, CHU Pitié-Salpêtrière, Paris, France (B.R.); Population Health Research Institute, David Braley Cardiac, Vascular and Stroke Research Institute, Perioperative Medicine and Surgical Research Unit, Hamilton, Ontario, Canada (P.J.D.); and Montpellier 1 University, Faculty of Medicine, Department of Biostatistics, Clinical Research and Medical Informatics, Nîmes University Hospital, Nîmes, France (P.L.).

Copyright © 2015, the American Society of Anesthesiologists, Inc. Wolters Kluwer Health, Inc. All Rights Reserved. Anesthesiology 2016; 124:570-9

between centers^{6,7} or in the design and analysis of clinical trials.^{8,9}

Existing risk scores, such as the American Society of Anesthesiologists (ASA) physical status score¹⁰ and the Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity (POSSUM) scoring system, which have been externally validated for predicting the probability of in-hospital mortality after surgery have considerable limitations.¹¹ The ASA score is based on physicians subjective assessment of a patient's preoperative clinical status and categorizes patients into five broad risk groups.¹² Although widely used worldwide, the ASA score does not consider the type of surgery the patient will undergo, makes no adjustment for age, and is based on subjective criteria; all factors contributing to its very limited accuracy^{13,14} and reliability.¹⁵ The POSSUM score has been widely validated^{16–18} but includes a number of risk factors collected at discharge (*e.g.*, operative blood loss and the presence of malignancy) that precludes its preoperative use. Although other risk scores for predicting in-hospital or 30-day mortality after surgery have been developed, these have overwhelmingly been surgery-specific tools for patients undergoing cardiac surgery.^{19,20}

Therefore, to date, there are very few risk scores that use objective preoperative patient information to predict postoperative in-hospital mortality for patients scheduled for any type of surgery. The objective of this study was therefore to develop and validate a preoperative risk score that could be used to predict postoperative in-hospital mortality, based on objective and readily available preoperative clinical information.

Materials and Methods

Study Population

Since 1996, all French hospitals, both public and private, caring for medical, surgical, and obstetric patients have been required to submit anonymous patient data to the National Hospital Discharge Data Base (NHDBB). As such, the NHDBB has complete information on all patients, in all centers. Each discharge summary submitted to the NHDBB is linked to a national grouping algorithm leading to a French Diagnosis-Related Group,²¹ thereby allowing patient comorbidities to be recorded and linked.²² As required by the French Protection Act concerning the use of anonymized hospital data, we obtained permission to access this database from the “Commission Nationale de l'Informatique et des Libertés” (Paris, France).

Selection of Patients and Procedures

We identified all surgical procedures, requiring anesthesia (*i.e.*, surgical procedures conducted in the presence of an anesthesiologist or under their supervision, whether this did or did not include an anesthetic intervention), performed in France on patients aged 18 yr or older during a 1-yr period, from January 1, 2010, to December 31, 2010. From this cohort, we selected those centers that performed more than

500 surgical procedures of any type during the 1-yr period. A center is defined as a single hospital or a group of hospitals sharing the same administration. For each patient record, the following information was extracted: sex, age (years), length of hospital stay (days), primary diagnosis (*i.e.*, the reason for the hospital stay as defined by attending physicians), and patient medical history—coded according to the *International Classification of Diseases*, 10th Revision (ICD-10).¹⁹ All medical (*e.g.*, postoperative mechanical ventilation and postoperative dialysis) or surgical (*e.g.*, reintervention related to complication and subsequent intervention(s) during the same stay) procedures performed during the hospital stay were recorded using the French classification for medical procedures “Classification Commune des Actes Médicaux” (CCAM).²⁰ The surgical procedure used for the analysis is the index operation, which reflected the planned surgical procedure at admission. For patients with multiple surgical procedures during the same stay, the index procedure was defined as the first one performed during the stay. Subsequent surgeries and/or medical procedures were recorded, but they were not used in the analysis, as they were likely to be not planned at admission.

Definition of Predictors

The ICD-10 disease categorization is complex and not commonly used in clinical practice. We therefore aggregated codes into broader disease groups to resemble the clinical observations commonly recorded during preoperative assessment (see Supplemental Digital Content 1, <http://links.lww.com/ALN/B228>). To ensure the validity of this aggregation, the process was conducted independently by three physicians with expertise in ICD-10 code management (Y.L.M., C.L.B.-B., and P.L.). Full consensus between the experts was required to include a code in a disease group. Furthermore, the experts also determined whether the disease groups, or single ICD-10 codes, were related to preoperative comorbidities or rather to postoperative adverse events. Any codes for which full agreement among experts was not obtained were excluded from the list of potential predictors.

Endpoint Definition

The primary endpoint was in-hospital mortality, defined as death after surgery and before discharge, regardless of the length of stay. Deaths occurring after hospital discharge were not considered and for the purpose of the analyses were censored, and patients were recorded as alive.

Statistical Methods

Data are expressed as mean \pm SD, number (percentage), odds ratio, and 95% CI. All *P* values were two tailed, and a *P* value less than 0.05 was considered significant.

We randomly divided the entire national cohort of French patients who underwent surgery in 2010 into two cohorts: one to develop the model and another to validate the model. Randomly splitting at the patient level is an inefficient approach to develop and validate a prediction model.^{23–25} Random splitting

a single data set at the patient level creates two cohorts that are very similar, except for random variation. Consequently, for large sample sizes, the predictive performance of the model will be very similar when evaluated in both the derivation and the validation cohorts and is thus hardly a strong test of the internal validation of the model.²⁶ We, therefore, split the cohort at the center level (*i.e.*, one hospital or several hospitals sharing the same administration) where the surgery was carried to produce derivation and validation cohorts. The centers were randomly selected to be included either in the derivation or in the validation cohort. Within a center, all patients were then either in the derivation or in the validation cohort.

We constructed a multivariable logistic model that included age, all preoperative variables (listed in Supplemental Digital Content 1, <http://links.lww.com/ALN/B228>), and predefined surgical subgroups (listed in Supplemental Digital Content 2, <http://links.lww.com/ALN/B229>) to predict an individual patient's probability of in-hospital mortality. We modeled age as both linear and nonlinear (using fractional polynomials and general additive models), but no sufficient gain in predictive performance of the model was identified and age was therefore retained as linear in the model. The large size of the derivation cohort made it likely that almost all variables examined would show a significant association with the primary outcome, resulting in an overly complex and clinically unusable final model. Therefore, we did not consider variables that occurred in the derivation cohort with a frequency less than 0.1% as potential predictors.

We developed an easy to use score (referred to as POSPOM [PreOperative Score to predict PostOperative Mortality]) from the derivation cohort using the approach described by Sullivan *et al.*²⁷ This approach assigns a positive score to less healthy risk factor states and a negative score to reflect healthier states. To develop a clinically meaningful score, a single point was made to represent a standard increase in risk (*i.e.*, a unit of risk). We defined this unit of risk as that risk associated with a 5-yr increase in age. This "amount" of risk is represented by the logistic regression coefficient associated with a 5-yr increase in age (β^{age^5}).

Based on the results of the logistic regression model, a score of 0 point was assigned to the surgical group with the lowest risk. Points were then assigned to all other groups by dividing their regression coefficients with the 5-yr age coefficient (*i.e.*, β^{age^5}). Individual patient risk was then estimated by calculating the sum of both the medical history of the patient and the surgical subgroup score.

We used both the derivation and the validation cohorts to evaluate the performance of POSPOM for predicting in-hospital mortality. Performance was evaluated by assessing the calibration and discrimination of both models. Calibration was assessed graphically by plotting the observed outcome against the predicted in-hospital probability. A smooth, nonparametric calibration line was created with the LOESS algorithm (*i.e.*, a locally weighted scatterplot smoothing) to estimate the observed probabilities of in-hospital mortality in relation to

the predicted probabilities.^{28–30} Discrimination was quantified by calculating the concordance statistic (*c*-statistic). Additional performance measures include the Yates slope (the difference in mean-predicted probabilities between those with and without the outcome), Brier score (squared difference between patient outcome and predicted risk), and scaled Brier score (scaled by the maximum Brier score and ranging from 0 to 100%). The performance of the models predicting the secondary composite endpoints was also evaluated in the validation cohort. All statistical analyses were carried out in R software (version 3.1) (<http://www.r-project.org>, last date accessed, July 21, 2015).

Results

Between January 1, 2010, and December 31, 2010, 7,059,447 eligible patients from 1,107 centers in France were recorded in the NHDBB. Of these, 5,507,834 patients met eligibility criteria and 2,717,902 (from 479 hospital centers) were allocated to the derivation cohort and 2,789,932 (from 479 hospital centers) were allocated to the validation cohort. The flowchart of the study is summarized in figure 1.

In-hospital mortality after surgery was 0.47% (95% CI, 0.46 to 0.48%) (12,786 deaths) in the derivation cohort and 0.54% (95% CI, 0.53 to 0.55%) (14,933 deaths) in the validation cohort. Baseline characteristics and surgical procedures distributions in both cohorts are shown in table 1.

Final Model

In the derivation cohort, 29 potential predictors (27 binary, 1 continuous, and 1 categorical variable) were evaluated for model inclusion. Of these, only 17 were retained in the final logistic model: age, ischemic heart disease, cardiac arrhythmia or heart blocks, chronic heart failure or cardiomyopathy, peripheral vascular disease, dementia, cerebrovascular disease, hemiplegia, chronic obstructive pulmonary disease, chronic respiratory failure, chronic alcohol abuse, cancer, diabetes, transplanted organ(s), chronic dialysis, chronic renal failure, and type of surgery. Final logistic model characteristics and the POSPOM are presented in tables 2 and 3 and in Supplemental Digital Content 3, <http://links.lww.com/ALN/B230>. The range of attainable POSPOM is 0 to 50; the range observed in the derivation set was 0 to 45.

Model Performance

POSPOM showed excellent discrimination (*c*-statistic: 0.944 [95% CI, 0.943 to 0.945], Brier score: 0.004 [95% CI, 0.004 to 0.004], scaled Brier score: 4.06% [95% CI, 3.76 to 5.18%], and Yates slope: 0.058 [95% CI, 0.057 to 0.059]) in the derivation cohort. In the validation cohort, similar performances (*c*-statistic: 0.929 [95% CI, 0.928 to 0.931], Brier score: 0.005 [95% CI, 0.005 to 0.005], scaled Brier score: 4.31% [95% CI, 3.25 to 5.71%], and Yates slope: 0.058 [95% CI, 0.057 to 0.060]) were observed. Inspection of calibration plot (fig. 2) demonstrates that POSPOM has good calibration with only a small underestimation of in-hospital mortality in the validation cohort for predicted probabilities ranging from 1

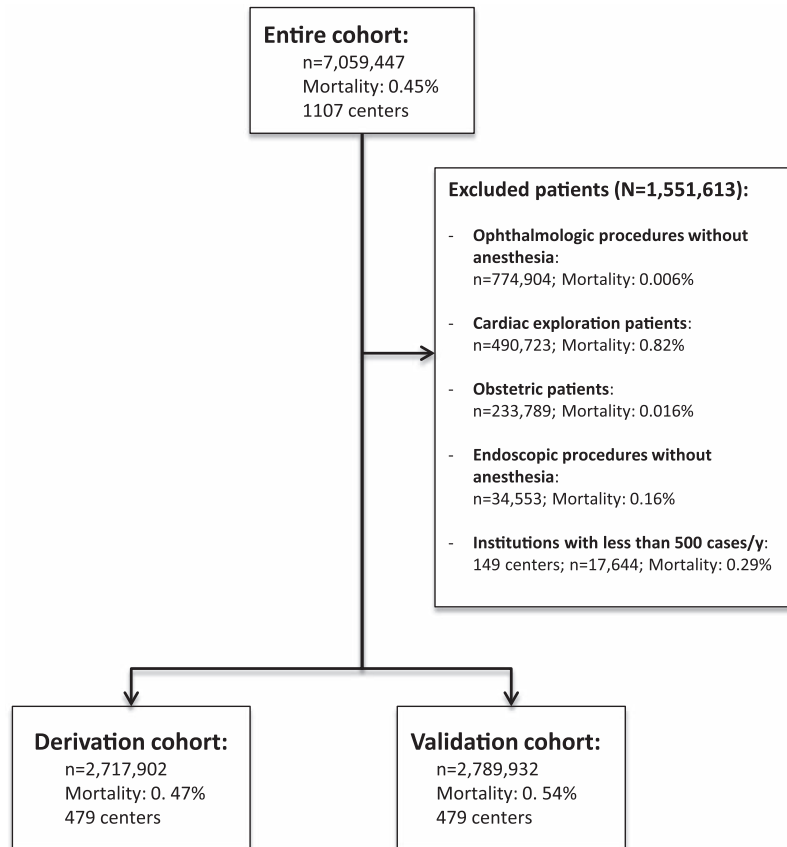


Fig. 1. Patient flowchart.

to 10%. The concordance between the predicted probabilities from the logistic regression model and the POSPOM scoring system was high (Lin concordance = 0.99), indicating little loss of predictive information when simplifying the logistic regression model to the POSPOM scoring system. In the validation cohort, POSPOM score equal to 30 (*i.e.*, predicted in-hospital mortality = 7.40%) was associated with an observed in-hospital mortality of 6.74% (95% CI, 6.40 to 7.08%). The distribution of POSPOM and the associated observed in-hospital mortality in the validation cohort are shown in figure 3. POSPOM values less than or equal to 20 were associated with a probability of in-hospital mortality less than or equal to 0.04% (*i.e.*, less than the in-hospital mortality observed in the full population—the average risk); a POSPOM value of 25 equates to a probability of in-hospital mortality of 1.73% (*i.e.*, about three times the average risk), and POSPOM values between 30 and 40 equate to a probability of in-hospital mortality of, respectively, 5.65 and 11.77% (*i.e.*, 20 times the average risk).

Discussion

Our study demonstrated that a score using only clinically relevant preoperative patient characteristics (*i.e.*, POSPOM) can accurately predict postoperative in-hospital mortality.

Traditionally, the risk associated with different types of surgery has been determined by observing the rate of outcome

events after each type of surgery.³¹ However, this approach does not determine how much of the event rate is due to the surgery itself, and how much is due to the medical comorbidities of the patients undergoing the surgery. In this study, we have shown tremendous variation in patient risk within surgical categories, much of which was driven by preexisting medical conditions. For example, the unadjusted risk of in-hospital mortality after major orthopedic surgery related to trauma was 3.46% (95% CI, 3.29 to 3.63%) compared with 1.09% (95% CI, 1.02 to 1.12%) for minor vascular surgery. After adjusting for preexisting medical comorbidities, the risk of in-hospital mortality after an endovascular procedure was very similar to that of major orthopedic surgery (table 3). These findings suggest that traditional surgery-specific risk estimation based only on the observed postoperative outcome rate does not provide a true reflection of surgical risk.

During a 1-yr period in 2010, every patient in France, requiring anesthesia in a center performing more than 500 procedures per year, was included to develop the POSPOM. This combination of a sufficiently long period of inclusion and lack of patient selection resulted in a large cohort guaranteeing accurate estimates of postoperative in-hospital mortality in the overall population.

Although the inclusion of intraoperative events or early postoperative complications would have improved the performances, a preoperative risk score for providing patients and clinicians with an estimate of in-hospital mortality after surgery precludes

Table 1. Patient Characteristics in the Derivation and Validation

	Derivation Cohort N = 2,717,902		Validation Cohort N = 2,789,932	
	%	n	%	n
Demographics				
Age (yr; mean, SD)	54.6	17.9	54.9	17.9
Male sex (% , n)	53.2	1,447,022	52.8	1,473,085
Medical history				
Ischemic heart disease (% , n)	3.34	90,762	3.69	102,898
Cardiac arrhythmia (% , n)	2.19	59,614	2.39	66,615
Chronic heart failure (% , n)	0.87	23,734	0.94	26,087
Heart valve disease (% , n)	1.01	27,468	1.08	30,063
Peripheral vascular disease (% , n)	0.66	17,962	0.70	19,438
Hypertension (% , n)	13.39	363,898	13.27	370,173
Dementia (% , n)	0.61	16,507	0.68	18,819
Cerebrovascular disease (% , n)	0.27	7,391	0.28	7,828
Hemiplegia, paraplegia (% , n)	0.44	11,958	0.47	13,176
Chronic obstructive pulmonary disease (% , n)	1.12	30,380	1.24	34,714
Pulmonary circulation disorder (% , n)	0.08	2,271	0.11	3,071
Chronic respiratory failure (% , n)	0.39	10,489	0.40	11,031
Chronic alcohol abuse (% , n)	0.83	22,659	0.82	22,733
Anemia (% , n)	0.84	22,740	0.89	24,849
Cancer (% , n)	2.96	80,380	3.21	89,614
Cancer with metastasis (% , n)	1.17	31,811	1.18	32,788
Diabetes (% , n)	4.60	125,052	4.76	132,650
Drug abuse (% , n)	0.14	3,770	0.16	4,545
Organ grafted (% , n)	0.19	5,026	0.26	7,332
Obesity (% , n)	3.59	97,621	3.59	100,243
Preoperative chronic dialysis (% , n)	0.38	10,256	0.37	10,224
Chronic renal failure (% , n)	1.05	28,448	1.15	32,104
Surgical procedures				
Cardiac surgery	0.60	16,279	0.87	24,384
Major digestive surgery	1.50	40,713	1.65	45,901
Minor digestive surgery	7.15	194,260	7.56	210,962
Digestive endoscopy	24.43	664,043	24.26	676,728
Interventional neuroradiology	0.11	2,926	0.09	2,500
Interventional cardiarrhythmology	1.25	34,024	1.39	38,676
Transplant surgery	0.02	461	0.04	1,182
Renal transplantation	0.04	1,108	0.06	1,549
Gynecologic surgery	8.65	234,988	8.81	245,839
Major liver surgery	0.12	3,364	0.16	4,501
Minor liver surgery	2.10	57,049	2.20	61,251
Neurosurgery	1.46	39,709	1.48	41,356
Ophthalmologic surgery	0.69	18,872	0.47	13,031
Ear, nose, and throat	6.44	175,134	6.26	174,670
Arthroplasty and spine surgery	5.30	143,993	5.09	141,899
Trauma-related orthopedic surgery	1.76	46,004	1.69	49,165
Minor orthopedic surgery	19.69	535,139	17.87	498,677
Plastic and reconstructive surgery	5.14	139,729	5.33	148,711
Multiple trauma-related surgery	0.07	1,907	0.08	2,366
Thoracic surgery	0.52	14,126	0.64	17,768
Major urologic surgery	0.90	24,480	1.12	31,197
Minor urologic surgery	4.43	120,526	5.11	142,686
Major vascular surgery	0.54	14,673	0.58	16,303
Minor vascular surgery	3.16	85,800	3.17	88,540
Others surgery	4.00	108,595	3.95	110,090

Table 2. POSPOM Regression

	Adjusted Risk of Postoperative Death				POSPOM Points
	Regression Coefficient	OR	95% CI Lower Limit	95% CI Upper Limit	
Age (per 5-yr increase)	0.303	1.36	1.35	1.36	+1
Ischemic heart disease	0.242	1.27	1.22	1.33	+1
Cardiac arrhythmia or heart blocks	0.215	1.24	1.18	1.30	+1
Chronic heart failure or cardiomyopathy	1.124	3.08	3.02	3.14	+4
Peripheral vascular disease	0.209	1.23	1.14	1.32	+1
Dementia	0.593	1.81	1.73	1.89	+2
Cerebrovascular disease	0.207	1.23	1.10	1.36	+1
Hemiplegia	1.125	3.08	2.98	3.18	+4
Chronic obstructive pulmonary disease	0.272	1.31	1.24	1.39	+1
Chronic respiratory failure	0.788	2.20	2.11	2.29	+3
Chronic alcohol abuse	1.345	3.84	3.76	3.93	+4
Cancer	1.074	2.92	2.86	2.98	+4
Diabetes	0.189	1.21	1.16	1.26	+1
Transplanted organ(s)	0.61	1.63	1.39	1.87	+2
Preoperative chronic hemodialysis	0.375	1.84	1.77	1.92	+1
Chronic renal failure	0.775	1.45	1.32	1.57	+2

OR = odds ratio; POSPOM = PreOperative Score to predict PostOperative Mortality.

Table 3. Stratification of the Risk Associated with Surgery

Surgery	N	Observed Mortality (%)	Part of Total Mortality (%)	Adjusted Risk of Postoperative Deaths			POSPOM Points
				OR	95% CI Lower Limit	95% CI Upper Limit	
Endoscopic digestive	664,043	0.01	0.46	0.01	0.01	0.01	+0
Ophthalmologic	18,872	0.02	0.03	0.01	0.00	0.01	+0
Gynecologic	234,988	0.06	1.06	0.05	0.04	0.05	+6
Other orthopedic	535,139	0.07	2.85	0.06	0.05	0.06	+6
Interventional cardiorhythmology	34,024	0.64	1.7	0.09	0.08	0.10	+8
Arthroplasty and spine	143,993	0.26	2.96	0.10	0.09	0.11	+9
Ear, nose, and throat	175,134	0.1	1.35	0.11	0.10	0.12	+9
Minor urologic	120,526	0.36	3.38	0.12	0.11	0.13	+9
Plastic	139,729	0.26	2.87	0.12	0.11	0.13	+9
Major urologic	24,480	0.83	1.59	0.26	0.23	0.28	+12
Others surgery	108,595	0.63	5.32	0.32	0.29	0.34	+12
Minor hepatic	57,049	0.64	2.86	0.33	0.29	0.35	+12
Minor gastrointestinal	194,260	0.61	9.26	0.35	0.32	0.36	+13
Renal transplant	1,108	1.08	0.09	0.37	0.20	0.49	+13
Minor vascular	85,800	1.09	7.34	0.43	0.39	0.45	+13
Orthopedic trauma	46,004	3.46	12.46	0.48	0.45	0.50	+14
Major hepatic	3,364	4.67	1.23	0.82	0.69	0.89	+15
Thoracic	14,126	3.9	4.31	0.84	0.76	0.89	+15
Neurosurgery	39,709	2.33	7.24	0.85	0.78	0.89	+15
Major vascular	14,673	4.2	4.82	0.97	0.87	1.02	+16
Major gastrointestinal	40,713	5.28	16.8	1	—	—	+16
Interventional neuroradiology	2,926	2.53	0.58	1.19	0.93	1.35	+17
Cardiac	16,279	6.11	7.78	1.38	1.26	1.44	+17
Transplant	461	15.62	0.56	5.73	4.16	6.70	+22
Multiple trauma related	1,907	7.39	1.1	5.88	4.81	6.49	+22

OR = odds ratio; POSPOM = PreOperative Score to predict PostOperative Mortality.

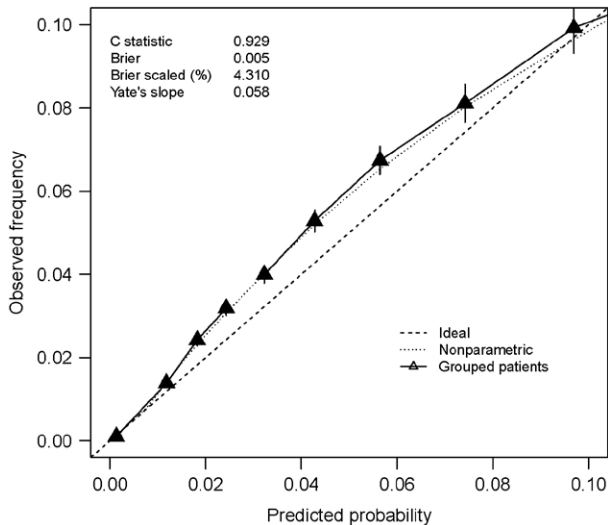


Fig. 2. Calibration plot: predicted *versus* observed in-hospital mortality for POSPOM (PreOperative Score to predict PostOperative Mortality) in the validation cohort ($n = 2,789,932$). Observed in-hospital mortality (bold line) with 95% CI for deciles of risk (triangles) and smooth. Nonparametric calibration line (dash line) was created with the LOESS algorithm (*i.e.*, a locally weighted scatterplot smoothing).

inclusion of intra- and postoperative risk factors. Accordingly, the POSPOM only included preoperative risk factors.

Comparison of Postoperative Mortality with Existing Studies

In our population of 7 million nonselected patients, we observed a postoperative in-hospital mortality of 0.5%. This contrasts with several large-scale studies that have reported postoperative mortality rates ranging from 1.3 to 4%.^{2,5,32,33} For example, the recent European Surgical Outcomes Study (EuSOS) conducted in 28 European countries reported a mortality rate as high as 4% in their sample of nonselected patients.⁵

However, as noted by the authors of the EuSOS, this value was noticeably higher than those reported in previous studies.^{2,34} Their sampling strategy, which included a 7-day observation period, is likely to have affected the case mix. Furthermore, the small proportion of European hospitals in the sample, with an overrepresentation of university hospitals in some countries, is likely to have contributed to an overestimation of in-hospital mortality. The POSPOM study, however, involved all centers, both private and public, that carried out 500 or more surgical procedures during 2010, in France, and is therefore more likely to be a true reflection of in-hospital mortality in France.

The American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP) also provides substantial information on postoperative mortality.³² Between 2005 and 2007, the 30-day mortality was 1.3% in about 300,000 patients from 200 participating hospitals.² As with the EuSOS, the sampling strategy, selection of centers and procedures, and duration of follow-up are likely to explain the difference in the observed mortality between their study and our study.³⁵

Similarly, the Veterans Administration Surgical Quality Improvement Program (VASQIP) database (January 2005 to August 2010) included 136,745 patients treated at 104 centers. Among them 1,568 patients (1.1%) sustained the primary 30-day mortality outcome. The two-fold increase in the observed mortality is likely due to the use of 30-day mortality as an outcome, as well as the type of surgical subspecialties included in the study. Only seven surgical subspecialties (vascular, general, neurosurgery, orthopedics, thoracic, urology, and otolaryngology) were recorded, and some procedures requiring anesthesia such as ophthalmologic surgery or gastrointestinal endoscopy were not considered. The ratio between postoperative myocardial infarction frequency and in-hospital mortality in our study ($0.16/0.47 = 0.36$ in the development cohort and $0.15/0.54 = 0.29$ in the validation cohort) was comparable with the Q-wave myocardial infarction frequency and 30-day mortality ratio observed in the VASQIP database ($0.3/1.1 = 0.27$).³³

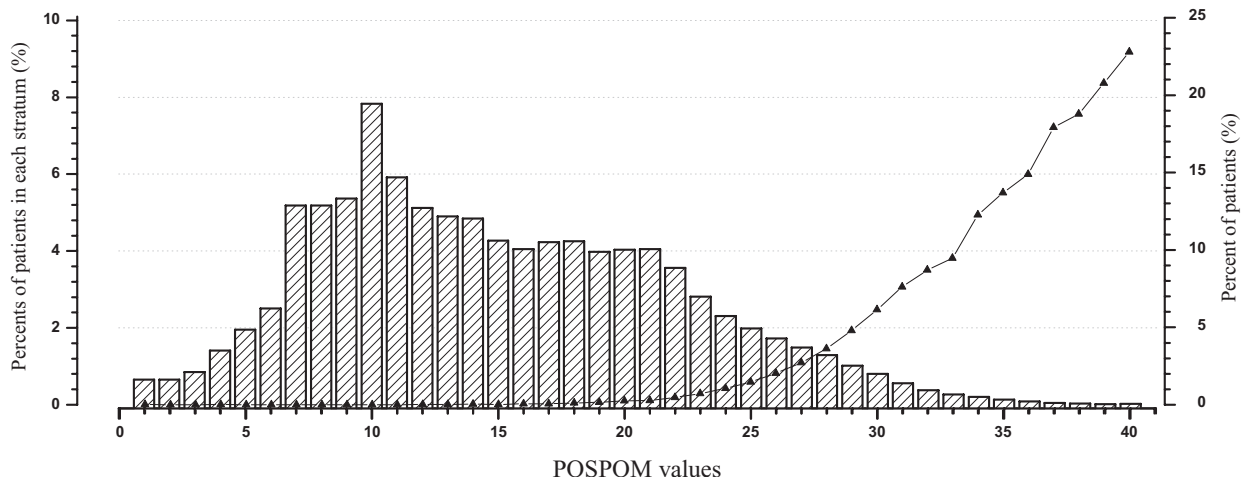


Fig. 3. Distribution of the POSPOM (PreOperative Score to predict PostOperative Mortality) values in the validation cohort ($n = 2,789,932$) in relation to the observed in-hospital mortality rate (solid line) at each POSPOM value.

In the VISION cohort study, in-hospital mortality was 1.38% (95% CI, 1.18 to 1.57%).³⁶ In the POSPOM study, when non-same-day surgery patients, older than 45 yr were selected, in-hospital mortality was 1.38% (95% CI, 1.36 to 1.40%) in the derivation cohort and 1.51% (95% CI, 1.49 to 1.53%) in the validation cohort (Supplemental Digital Content 4, <http://links.lww.com/ALN/B231>, and Supplemental Digital Content 5, <http://links.lww.com/ALN/B232>). The large increase in the postoperative cardiac outcomes is related to a major shift in the definition of the postoperative cardiac events.

Comparison of POSPOM Performances with Existing Risk Scores

ASA physical status classification¹⁰ remains widely used to stratify the preoperative risk of postoperative mortality. However, its accuracy and more importantly its reproducibility among observers³⁷ represent major limitations,³⁸ which also has an impact on risk scores that include this classification as a variable.^{2,39,40} Davenport *et al.*⁴¹ demonstrated an interdependence between the ASA physical status and the clinical risk factors in the ACS-NSQIP and that the inclusion of ASA physical status was not associated with an increase in discrimination.

Using a total of 35,179,507 patient stay records from 2001 to 2006 Medicare Provider Analysis and Review (MEDPAR) files, Sessler *et al.*⁴² developed the risk stratification indices (RSIs) for risk adjustment (including 30-day mortality and in-hospital mortality) to enable healthcare provider profiling. Although the RSIs demonstrated excellent discrimination, its calibration limits its generalizability and usefulness for providing patients and clinicians an accurate estimate of in-hospital mortality.⁴³ Furthermore, the large number of variables used in the score (*e.g.*, 184 variables for in-hospital mortality) and the infrequent occurrence of some of the preoperative variables (*e.g.*, cardiopulmonary resuscitation) limit its applicability. In a subsequent study,⁴⁴ RSIs was modified to incorporate the timing of diagnoses and procedures to develop the Present-On-Admission (POA) risk score. Observed discrimination remained high; however, almost 2,000 coefficients were used to fit the final model; this represents a clear limitation for clinical use. Both RSI and POA risk score have been constructed primarily as tools for research—for adjustment for (baseline) confounding—and not for individual patient prediction.

The collection of data from a single country is a limitation of this study. In the EuSOS,⁵ there were variations across hospitals regarding in-hospital mortality, even after adjusting for confounding. The odds ratio of in-hospital mortality varied from 0.44 (95% CI, 0.19 to 1.05; $P = 0.06$) for Finland to 6.92 (95% CI, 2.37 to 20.27; $P = 0.0004$) for Poland compared with that for United Kingdom, the country with the largest data set.⁵ We chose a single country with a pay-for-performance system. Clearly, the motives underlying the introduction and development of Diagnosis-Related Group systems vary greatly from country to country, adapted to each country, according to their individual developmental contexts or to their conception of a welfare state.²² Variations may relate to differences in the health

system model used, the relationship between providers and funders, the degree of centralization, the separation between purchasing and provision, the structure of the hospital market, the type of centers (*e.g.*, profit *vs.* nonprofit), or the level of competition between public and private structures. In this study, we considered a wide range of public and private hospitals and we included more than 5 million procedures performed in more than 1,100 hospitals. Finally, we did not consider including any interactions between any of the risk factors and age, as our intention was to keep the model simple and easy to use and that interactions rarely add to the predictive ability of the models.^{26,30} POSPOM requires validation in other countries.

Conclusions

In summary, our POSPOM risk score is a robust tool for predicting in-hospital mortality in patients undergoing surgery and has very good discriminative and calibration properties. Physicians may find it practical to use and applicable to clinical practice.

Acknowledgments

Drs. Le Manach and Landais had full access to all the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. Drs. Le Manach, Collins, and Landais helped in study concept and design; Drs. Le Manach, Le Bihan-Benjamin, and Landais in acquisition of data; Drs. Le Manach, Collins, and Landais in analysis and interpretation of data; Drs. Le Manach and Collins in drafting of the manuscript; Drs. Le Manach, Collins, Rodseth, Biccard, Devereaux, Riou, and Landais in critical revision of the manuscript for important intellectual content; and Le Manach and Collins in statistical analysis. Dr. Le Manach is supported by the Hamilton Anesthesia Associates and the Canadian Network and Centre for Trials Internationally, Hamilton, Ontario, Canada; Dr. Collins is funded by the UK Medical Research Council (grant number G1100513, London, United Kingdom); Dr. Rodseth is supported by a CIHR Scholarship (Canada-HOPE Scholarship, Toronto, Ontario, Canada), the College of Medicine of South Africa (Phyllis Kocker/Bradlow Award, Cape Town, South Africa), and the University of KwaZulu-Natal (Competitive Research Grant, Durban, South Africa); and Dr. Devereaux is supported by Career Investigator Award of Heart and Stroke Foundation of Ontario, Canada.

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Le Manach: Population Health Research Institute; David Braley Cardiac, Vascular and Stroke Research Institute; Perioperative Medicine and Surgical Research Unit, Hamilton, Ontario, Canada. yannick.lemanach@phri.ca. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

References

1. Weiser TG, Regenbogen SE, Thompson KD, Haynes AB, Lipsitz SR, Berry WR, Gawande AA: An estimation of the

- global volume of surgery: A modelling strategy based on available data. *Lancet* 2008; 372:139–44
2. Gance LG, Lustik SJ, Hannan EL, Osler TM, Mukamel DB, Qian F, Dick AW: The Surgical Mortality Probability Model: Derivation and validation of a simple risk prediction rule for noncardiac surgery. *Ann Surg* 2012; 255:696–702
 3. Jencks SF, Williams MV, Coleman EA: Rehospitalizations among patients in the Medicare fee-for-service program. *N Engl J Med* 2009; 360:1418–28
 4. Pearse RM, Harrison DA, James P, Watson D, Hinds C, Rhodes A, Grounds RM, Bennett ED: Identification and characterisation of the high-risk surgical population in the United Kingdom. *Crit Care* 2006; 10:R81
 5. Pearse RM, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, Vallet B, Vincent JL, Hoeft A, Rhodes A; European Surgical Outcomes Study (EuSOS) group for the Trials groups of the European Society of Intensive Care Medicine and the European Society of Anaesthesiology: Mortality after surgery in Europe: A 7 day cohort study. *Lancet* 2012; 380:1059–65
 6. Bottle A, Sanders RD, Mozd A, Aylin P: Provider profiling models for acute coronary syndrome mortality using administrative data. *Int J Cardiol* 2013; 168:338–43
 7. DeLong ER, Peterson ED, DeLong DM, Muhlbaier LH, Hackett S, Mark DB: Comparing risk-adjustment methods for provider profiling. *Stat Med* 1997; 16:2645–64
 8. Hayward RA, Kent DM, Vijan S, Hofer TP: Multivariable risk prediction can greatly enhance the statistical power of clinical trial subgroup analysis. *BMC Med Res Methodol* 2006; 6:18
 9. Kent DM, Hayward RA, Griffith JL, Vijan S, Beshansky JR, Califf RM, Selker HP: An independently derived and validated predictive model for selecting patients with myocardial infarction who are likely to benefit from tissue plasminogen activator compared with streptokinase. *Am J Med* 2002; 113:104–11
 10. Keats AS: The ASA classification of physical status—A recapitulation. *ANESTHESIOLOGY* 1978; 49:233–6
 11. Copeland GP, Jones D, Walters M: POSSUM: A scoring system for surgical audit. *Br J Surg* 1991; 78:355–60
 12. Wolters U, Wolf T, Stützer H, Schröder T: ASA classification and perioperative variables as predictors of postoperative outcome. *Br J Anaesth* 1996; 77:217–22
 13. Wolters U, Wolf T, Stützer H, Schröder T, Pichlmaier H: Risk factors, complications, and outcome in surgery: A multivariate analysis. *Eur J Surg* 1997; 163:563–8
 14. Peersman G, Laskin R, Davis J, Peterson MG, Richart T: ASA physical status classification is not a good predictor of infection for total knee replacement and is influenced by the presence of comorbidities. *Acta Orthop Belg* 2008; 74:360–4
 15. Mak PH, Campbell RC, Irwin MG; American Society of Anesthesiologists: The ASA Physical Status Classification: Inter-observer consistency. *American Society of Anesthesiologists. Anaesth Intensive Care* 2002; 30:633–40
 16. Lam CM, Fan ST, Yuen AW, Law WL, Poon K: Validation of POSSUM scoring systems for audit of major hepatectomy. *Br J Surg* 2004; 91:450–4
 17. Ramesh VJ, Rao GS, Guha A, Thennarasu K: Evaluation of POSSUM and P-POSSUM scoring systems for predicting the mortality in elective neurosurgical patients. *Br J Neurosurg* 2008; 22:275–8
 18. Merad F, Baron G, Pasquet B, Hennes H, Kohlmann G, Warlin F, Desrousseaux B, Fingerhut A, Ravaud P, Hay JM: Prospective evaluation of in-hospital mortality with the P-POSSUM scoring system in patients undergoing major digestive surgery. *World J Surg* 2012; 36:2320–7
 19. Nilsson J, Algotsson L, Höglund P, Lührs C, Brandt J: Comparison of 19 pre-operative risk stratification models in open-heart surgery. *Eur Heart J* 2006; 27:867–74
 20. Moonesinghe SR, Mythen MG, Das P, Rowan KM, Grocott MP: Risk stratification tools for predicting morbidity and mortality in adult patients undergoing major surgery: Qualitative systematic review. *ANESTHESIOLOGY* 2013; 119:959–81
 21. Manuel des Groupes Homogènes de Maladies; Ministère de la Santé et des Sports, France. 2010. Available at: http://www.sante.gouv.fr/fichiers/bos/2010/ste_20100004_0001_p000.pdf. Accessed July 14, 2013
 22. Landais P, Boudemaghe T, Suehs C, Dedet G, Leblanc-Benjamin C: Computerized medico-economic decision making: an international comparison, *Medical Informatics, e-Health*. Edited by Venot A, Burgun A, Quantin C. Paris, Springer, 2014, pp 191–220
 23. Altman DG, Royston P: What do we mean by validating a prognostic model? *Stat Med* 2000; 19:453–73
 24. Altman DG, Vergouwe Y, Royston P, Moons KG: Prognosis and prognostic research: Validating a prognostic model. *BMJ* 2009; 338:b605
 25. Hirsch RP: Validation samples. *Biometrics* 1991; 47:1193–4
 26. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med* 2015; 162:W1–73
 27. 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
 28. Harrell FE: *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York; London, Springer, 2001, 262–4
 29. Steyerberg EW: *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York, Springer, 2009, 270–9
 30. Collins GS, Reitsma JB, Altman DG, Moons KG: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD statement. *Ann Intern Med* 2015; 162:55–63
 31. Eagle KA, Rihal CS, Mickel MC, Holmes DR, Foster ED, Gersh BJ: Cardiac risk of noncardiac surgery: Influence of coronary disease and type of surgery in 3368 operations. CASS Investigators and University of Michigan Heart Care Program. *Circulation* 1997; 96:1882–7
 32. Khuri SF, Henderson WG, Daley J, Jonasson O, Jones RS, Campbell DA Jr, Fink AS, Mentzer RM Jr, Steeger JE; Principal Site Investigators of the Patient Safety in Surgery Study: The patient safety in surgery study: Background, study design, and patient populations. *J Am Coll Surg* 2007; 204:1089–102
 33. London MJ, Hur K, Schwartz GG, Henderson WG: Association of perioperative β -blockade with mortality and cardiovascular morbidity following major noncardiac surgery. *JAMA* 2013; 309:1704–13
 34. Noordzij PG, Poldermans D, Schouten O, Bax JJ, Schreiner FA, Boersma E: Postoperative mortality in The Netherlands: A population-based analysis of surgery-specific risk in adults. *ANESTHESIOLOGY* 2010; 112:1105–15
 35. Birkmeyer JD, Shahian DM, Dimick JB, Finlayson SR, Flum DR, Ko CY, Hall BL: Blueprint for a new American College of Surgeons: National Surgical Quality Improvement Program. *J Am Coll Surg* 2008; 207:777–82
 36. Devereaux PJ, Chan MT, Alonso-Coello P, Walsh M, Berwanger O, Villar JC, Wang CY, Garutti RI, Jacka MJ, Sigamani A, Srinathan S, Biccard BM, Chow CK, Abraham V, Tiboni M, Pettit S, Szczeklik W, Lurati Buse G, Botto F, Guyatt G, Heels-Ansdell D, Sessler DI, Thorlund K, Garg AX, Mrkobrada M, Thomas S, Rodseth RN, Pearse RM, Thabane L, McQueen MJ, VanHelder T, Bhandari M, Bosch J, Kurz A, Polanczyk C, Malaga G, Nagele P, Le Manach Y, Leuwer M, Yusuf S: Association between postoperative troponin levels and 30-day mortality among patients undergoing noncardiac surgery. *JAMA* 2012; 307: 2295–304
 37. Cuvillon P, Nouvellon E, Marret E, Albaladejo P, Fortier LP, Fabbro-Perray P, Malinovsky JM, Ripart J: American Society

- of Anesthesiologists' physical status system: A multicentre Francophone study to analyse reasons for classification disagreement. *Eur J Anaesthesiol* 2011; 28:742–7
38. Cohen ME, Bilimoria KY, Ko CY, Richards K, Hall BL: Effect of subjective preoperative variables on risk-adjusted assessment of hospital morbidity and mortality. *Ann Surg* 2009; 249:682–9
 39. Ragg JL, Watters DA, Guest GD: Preoperative risk stratification for mortality and major morbidity in major colorectal surgery. *Dis Colon Rectum* 2009; 52:1296–303
 40. Dalton JE, Kurz A, Turan A, Mascha EJ, Sessler DI, Saager L: Development and validation of a risk quantification index for 30-day postoperative mortality and morbidity in noncardiac surgical patients. *ANESTHESIOLOGY* 2011; 114:1336–44
 41. Davenport DL, Bowe EA, Henderson WG, Khuri SF, Mentzer RM Jr: National Surgical Quality Improvement Program (NSQIP) risk factors can be used to validate American Society of Anesthesiologists Physical Status Classification (ASA PS) levels. *Ann Surg* 2006; 243:636–41; discussion 641–4
 42. 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
 43. Sigakis MJ, Bittner EA, Wanderer JP: Validation of a risk stratification index and risk quantification index for predicting patient outcomes: in-hospital mortality, 30-day mortality, 1-year mortality, and length-of-stay. *ANESTHESIOLOGY* 2013; 119:525–40
 44. 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

ANESTHESIOLOGY REFLECTIONS FROM THE WOOD LIBRARY-MUSEUM

A “Nitrous Oxide Gas” Advertising Bookmark from Dr. A. K. Harroun



According to the July 1882 issue of *The Odontographic Journal*, “the bogus dental college of Morrison & Co., at Delevan, Wis[consin] ... issued an ‘announcement’ ... with ... a list of its graduates ... [whose] sheepskins are going ‘dog cheap’ at twelve dollars.” One of the 1881 graduates listed is “A. K. Harroun, ... Pennsylvania.” On his advertising bookmark above, that graduate lists himself as “A. K. Harroun, D.D.S. Dental Rooms ... Honesdale, Pa.” Because his office shifted from 125 to 255 Main Street, Dr. Harroun had to strike the old address and stamp in the new one on his old stock of bookmarks. He is careful to note that, “Teeth extracted without pain with Nitrous Oxide Gas.”

From the mid-1860s through the 1880s, most dentists who administered laughing gas to patients were titled “Dr.” as an honorific rather than by an earned degree. Among dentists who learned their profession from a preceptor rather than a professor, most resisted ordering mail order degrees from “Morrison & Co.” This advertising bookmark is part of the Wood Library-Museum’s Ben Z. Swanson Collection. (Copyright © the American Society of Anesthesiologists, Inc.)

George S. Bause, M.D., M.P.H., Honorary Curator, ASA’s Wood Library-Museum of Anesthesiology, Schaumburg, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYC@aol.com.