Risk model of in-hospital mortality after pulmonary resection for cancer: A national database of the French Society of Thoracic and Cardiovascular Surgery (Epithor)

Alain Bernard, MD,[a] Caroline Rivera, MD,[b] Pierre Benoit Pages, MD,[a] Pierre Emmanuel Falcoz, MD,[c] Eric Vicaut, MD,[d] and Marcel Dahan, MD[e]

Objectives: The estimation of risk-adjusted in-hospital mortality is essential to allow each thoracic surgery team to be compared with national benchmarks. The objective of this study is to develop and validate a risk model of mortality after pulmonary resection.

Methods: A total of 18,049 lung resections for non–small cell lung cancer were entered into the French national database Epithor. The primary outcome was in-hospital mortality. Two independent analyses were performed with comorbidity variables. The first analysis included variables as independent predictive binary comorbidities (model 1). The second analysis included the number of comorbidities per patient (model 2).

Results: In model 1 predictors for mortality were age, sex, American Society of Anesthesiologists score, performance status, forced expiratory volume (as a percentage), body mass index (in kilograms per meter squared), side, type of lung resection,extended resection, stage, chronic bronchitis, cardiac arrhythmia, coronary artery disease, congestive heart failure, alcoholism, history of malignant disease, and prior thoracic surgery. In model 2 predictors were age, sex, American Society of Anesthesiologists score, performance status, forced expiratory volume, body mass index, side, type of lung resection, extended resection, stage, and number of comorbidities per patient. Models 1 and 2 were well calibrated, with a slope correction factor of 0.96 and of 0.972, respectively. The area under the receiver operating characteristic curve was 0.784 (95% confidence interval, 0.76–0.8) in model 1 and 0.78 (95% confidence interval, 0.76–0.797) in model 2.

Conclusions: Our preference is for the well-calibrated model 2 because it is easier to use in practice to estimate the adjusted postoperative mortality of lung resections for cancer. (J Thorac Cardiovasc Surg 2011;141:449-58)

GTS

Lung resection is still the main curative treatment for patients with non–small cell lung cancer. The surgical lung most often concerns fragile patients with chronic obstructive pulmonary disease or other medical conditions, such as cardiovascular disease. The risk of the operation is higher in patients with concomitant respiratory or cardiac disease.[1] Several single-institution retrospective studies that examined the characteristics of patients with an increased risk of mortality after major pulmonary resection have been published.[1,2] Series that span several decades are of limited size, single-institution studies are not sufficiently reliable, and the results reported cannot be extrapolated to the practices of other surgical teams.[3,4] The development of large multicenter databases can improve the quality of estimations of postoperative death. Several scoring systems based on various comorbid conditions have previously been used to stratify patients according to the risk of complications.[1,5-7] However, these scoring systems were developed in other populations and thus might not be ideal for patients who undergo surgical intervention for non–small cell lung cancer.[1,5,6] In addition, these models do not include several other prognostic factors, such as sex, age, extent of resection, and tumor stage, which could be relevant in these patients.

Substantial variations in postoperative mortality have been reported by other countries and institutions, but direct comparison is hampered by differences in definitions and selection criteria. Case series from large clinics tend to show lower mortality than do population-based studies. This could be explained by the superior performance of specialized institutions but might also be caused by selection bias.[8] A predictive model should be developed and validated from a nationally representative thoracic surgery database. Epithor is a national database developed by the French Society of Thoracic and Cardiovascular Surgery.[9]

The objective of this study is to develop and validate a risk model for in-hospital mortality of surgical patients with non–small cell lung cancer. This model will be used to stratify the aggregate population risk for risk-adjusted

From the Department of Thoracic Surgery,[a] CHU Dijon, France; the Department of General Thoracic Surgery,[b] CHU Bordeaux, France; the Department of General Thoracic Surgery,[c] CHU Strasbourg, France; the Biostatistics Unit,[d] Hospital Fernand Widal, Paris, France; and the Department of General Thoracic Surgery,[e] CHU Toulouse, France.

Disclosures: Authors have nothing to disclose with regard to commercial support.
Received for publication Jan 27, 2010; revisions received June 10, 2010; accepted for publication June 28, 2010; available ahead of print Aug 9, 2010.

Address for reprints: Alain Bernard, MD, Department of Thoracic Surgery Hospital du Bocage, bd de Lattre de Tassigny, 21034 Dijon Cedex France (E-mail: alain.bernard@chu-dijon.fr).

0022-5223/$36.00

doi:10.1016/j.jtcvs.2010.06.044

Abbreviations and Acronyms
ASA = American Society of Anesthesiologists
AUC = area under the curve
BMI = body mass index
CI = confidence interval
DLCO = diffusion capacity for carbon monoxide
FEV = forced expiratory volume
ROC = receiver operating characteristic

comparison of performance between units. We evaluated the effect of individual comorbidities and the number of comorbidities per patient, as well as the effect of the characteristics of patients, procedures, and tumor stage on mortality in pulmonary resection.

MATERIALS AND METHODS

Epithor: The French National Database

The French Society of Thoracic and Cardiovascular Surgery database was established in 2003 as a voluntary initiative of general thoracic surgeons. Today, 70 private and public institutions record their data in the database. Details of the Epithor data collection instrument can be found on the French Society of Thoracic and Cardiovascular Surgery Web site.[9] Thoracic surgery units that applied to contribute to this database were visited and validated by the coordinator and were then sent a confidential code. Each medical record contains 50 variables, of which 14 are required to initialize and 2 to close the file. Data are sent by Internet to the national database, and patients are anonymous. Multiple coherence tests are then carried out to warn the participating site of the presence of an anomaly. Each surgeon can check the quality of the data by comparing them with national data through a quality score ranging from 0% to 100%. Variables are collected on a data form that includes information about every patient's personal characteristics, medical history, surgical procedures, cancer staging, and outcomes.

Patient Population

Between January 2003 and December 2008, a total of 83,000 operations were entered into the Epithor database by 81 surgical sites, including 19,031 procedures for non–small cell lung cancer. We excluded exploratory thoracotomy (n = 982) during the study period. The final study population consisted of 18,049 pulmonary resections. Baseline demographics included age, sex, American Society of Anesthesiologists (ASA) score, performance status (Appendix 1), body mass index (BMI; in kilograms per meter squared), and forced expiratory volume (FEV; as a percentage). The diffusion capacity for carbon monoxide (DLCO) and dyspnea score were excluded because too many data were missing. In the preoperative period, the DLCO was not routinely used by thoracic surgeons. The comorbidities selected for analysis were smoking, chronic bronchitis, coronary artery disease, cardiac arrhythmia, congestive heart failure, peripheral vascular disease, pulmonary artery hypertension, alcoholism, cirrhosis, cerebral vascular event, diabetes mellitus, renal insufficiency, coagulopathy, history of malignant disease, prior thoracic surgery, preoperative chemotherapy, steroids, valvular heart disease, and pulmonary embolism. The different comorbidities are defined in Appendix 2. Very few patients were treated with preoperative radiotherapy. The type of procedure included limited resection (wedge resection or segmentectomy) lobectomy or bilobectomy and pneumonectomy. Extended resection was defined as en bloc chest wall resection, a portion of the left atrium, carinal resection, the diaphragm, and the superior vena cava. Mediastinal lymph node dissection included nodal sampling or radical lymphadenectomy. Surgical resection margins were classified as complete (R0) or microscopically (R1) or macroscopically (R2) invaded. Primary lung cancer was classified as stage I (IA or IB), II (IIA or IIB), III (IIIA or IIIB), or IV in accordance with the American Joint Committee on Cancer.[10]

Outcome Definition

In-hospital mortality included patients who died within the first 30 days after the operation and those who died later during the same hospitalization.

Missing Data

Baseline demographics (including age and sex), comorbidities, and procedure and outcome date were recorded in every case. Data concerning the ASA score, performance status, FEV, pathological stage, and surgical resection margin status were sometimes incomplete; the proportion of missing data varied between 1.8% and 23.6%. We assumed that the missing data were missing at random; that is to say, the fact that the data were missing was not related to the true (unobserved) values of the missing data.[11] We applied a multiple imputation framework to compensate for missing prognostic factor data. For categorical variables (ASA score, performance status, pathological stage, and surgical resection margin status), we applied a multinomial logistic model, and for continuous variables (FEV), we used linear regression. Missing values for ASA scores and performance status were entered by using a multinomial logistic regression model that included sex, age, FEV, procedure, extended resection, and comorbidities. Missing values for FEV were entered by using linear regression that included sex, age, ASA score, performance status, procedures, side, extended resection, and comorbidities. Missing values for surgical pathological stage and surgical resection margin were entered by using multinomial logistic regression that included sex, age, procedures, side, extended resection, ASA score, and performance status.[11]

Development of the Risk Model

To determine independent factors for in-hospital death, we first performed univariate analysis with χ^2 tests for binary and categorical variables and a *t* test for continuous variables. Variables with a level of significance of less than or equal to .1 in the univariate analysis were included in the multivariate analysis by means of logistic regression.[12] Continuous or ranked variables were tested to ensure conformity with the linear gradient by using the likelihood ratio χ^2 statistic.[12] Interaction effects were sought for all variables included in the model. All models were constructed by using backward stepwise variable selection. A step-down variable selection using Akaike's information criterion was used as a stopping rule.[12] For the purpose of the regression analysis, age and FEV were continuous variables, and the ASA score, performance status, and BMI were ranked variables. The type of pulmonary resection, pathological stage, and surgical resection margin status were transformed into dummy variables. The other variables were binary: sex, side, and extended resection. Two independent analyses were performed with comorbidity variables. The first regression analysis included variables as independent predictive binary comorbidities by using a step-down procedure (model 1). The second regression analysis included the number of comorbidities per patient as a ranked variable grouped into 4 values (0, 1, 2, and $\geq$3; model 2).

Validation

The area under the receiver operating characteristic (ROC) curve, the Somer's Dxy correlation, and the R^2 value were used to measure the discriminatory ability of the model.[12,13] The reliability of the model was estimated by the relationship between the predicted probability and the observed outcome in that sample. Calibration by plotting predicted against observed probability can estimate intercepts and slopes of curves to quantify overfitting.[13] Well-calibrated models have a slope of 1, whereas models that provide overly extreme predictions have a slope of less than 1: low predicted probabilities are too low, and high predicted probabilities are

too high. The reliability of the model was assessed with the Hosmer–Lemeshow goodness-of-fit test.[14]

The internal validation of the model was assessed with bootstrap resampling techniques.[13] Bootstrap samples consisted of n patients randomly drawn with replacement from the original dataset (training set) of size n. A model in the bootstrap sample (test set) is derived and applied to the original sample without change. The discriminatory and reliability index from the bootstrap sample minus the index computed on the original sample is an estimate of optimism. This process is repeated for 200 bootstrap replications to obtain an average optimism, which is subtracted from the final model (index-corrected) fit's apparent accuracy to obtain the overfitting-corrected estimate.[13] Calibration accuracy is estimated by using a nonparametric smoother that relates predicted probabilities to observed binary outcomes. After averaging many replications, the predicted value–specific differences are then subtracted from the apparent differences, and an adjusted calibration curve is obtained.[13] The mean absolute calibration error was also estimated.[13] The logistic regression models were compared by using the measure of discriminatory ability and the reliability index.

Discrete variables were expressed as numbers with percentages, and continuous variables were expressed as means and standard deviations. Calculations were performed with STATA 11 statistical software (StataCorp, College Station, Tex) and R statistical software, for which we used Harrell's Design library (http://www.r-project.org).

RESULTS

There were 690 deaths (in-hospital mortality, 3.8%; 95% confidence interval [CI], 3.5%–4.1%). The baseline patients' characteristics and in-hospital mortalities for the different variables are shown in Table 1. The type of pulmonary resection was the strongest significant predictor of in-hospital mortality: 2.4% for limited resection, 3% for lobectomy, and 7.7% for pneumonectomy ($P < .00001$, Table 1). Extended resection was associated with a 7% risk of in-hospital mortality ($P < .0001$, Table 1). Mortality was 3% in patients with left-sided pulmonary resection and 4.4% in those with right-sided pulmonary resection ($P < .0001$, Table 1). Comorbidity variables associated with increased mortality in univariate analyses included chronic bronchitis, coronary artery disease, cardiac arrhythmia, congestive heart failure, peripheral vascular disease, alcoholism, renal insufficiency, coagulopathy, history of malignant disease, and prior thoracic surgery (Table 1).

Risk Models for In-Hospital Mortality

In model 1 multivariate analysis identified age, sex, ASA score, performance status, FEV, BMI, side, lobectomy, pneumonectomy, extended resection, stage III disease, stage IV disease, chronic bronchitis, cardiac arrhythmia, coronary artery disease, congestive heart failure, alcoholism, history of malignancy disease, and prior thoracic surgery as independent predictors of in-hospital mortality (Appendix 3). In model 2 multivariate analysis identified age, sex, ASA score, performance status, FEV, BMI, side, lobectomy, pneumonectomy, extended resection, stage III disease, stage IV disease, and number of comorbidities per patient as independent predictors of in-hospital mortality (Table 2). The relationship of FEV was linear with respect to the logit of risk ($P = .23$) without a scale or spline transformation. The linear gradient of the variables age ($P = .6$), ASA score ($P = .53$), and performance status ($P = .37$) was accepted. The test for the linear gradient of the variables BMI ($P = .007$) and number of comorbidities per patient ($P = .027$) was rejected. Therefore these 2 variables were transformed into dummy variables (Table 2). Three interactions were identified between side and pneumonectomy, FEV and pneumonectomy, and FEV and extended resection (Table 2). The predicted logit at a certain value of FEV differed according to the type of pulmonary resection. In patients who had undergone pneumonectomy, the predicted logit varied little according to the value of FEV. This was not the case for patients who had undergone limited resection or lobectomy (Figure 1). The predicted logit varied little according to the value of FEV in patients who had undergone extended resection, whereas it decreased linearly among patients who had undergone simple pulmonary resection (Figure 2). Patients who had undergone right-sided pneumonectomy had an adjusted odds ratio of 2.9 (95% CI, 1.44–5.88), and patients with left-sided pneumonectomy had an adjusted odds ratio of 1.78 (95% CI, 0.87–3.645; Figure 3).

Validation

The performance of the prediction model for the training set and its ability to predict in-hospital mortality for the test set were compared (Table 3). The index-corrected Dxy correlation and R^2 values for models 1 and 2 were comparable (Table 3). The calibration plots are shown in Figure 4.The slope-correction factor for models 1 and 2 was 0.96 and 0.972, respectively (Table 3). The Hosmer–Lemeshow goodness-of-fit statistic was not statistically significant in models 1 and 2 (Table 3).

Comparison of the 2 Models

Areas under the ROC curves compared the discriminatory abilities of models 1 and 2. The area under the curve (AUC) for model 1 was 0.784 (95% CI, 0.76–0.8), whereas it was 0.78 (95% CI, 0.76–0.797) in model 2; the AUCs were not significantly different ($P = .19$). With bootstrapping, the AUC for model 1 was unchanged at 0.785 (95% CI, 0.77–0.8) and for model 2, it was unchanged at 0.78 (95% CI, 0.765–0.8). The reliability index and calibration were slightly better in model 2 than in model 1 (Table 3 and Figure 4). According to the bootstrap calibration curve using a nonparametric smoother to relate predicted probabilities to observed binary outcomes, the mean absolute error was 0.002 for model 1 and 0.0014 for model 2.

DISCUSSION

Both model 1, which included individual comorbidities, and model 2, which took into account the number of comorbidities per patient, had good discrimination, as shown by

TABLE 1. Categorical risk factors in patient survival and in-hospital mortality

Variables	Categories	Survivors (n = 17,359)	In-hospital mortality (n = 690)	*P* value
Sex	Male	13,185 (95.5%)	618 (4.5%)	.00001
	Female	4174 (98.3%)	72 (1.7%)	
Age	Years	62.5 ± 10	66.9 ± 9.3	.00001
ASA score	1	3114 (98.7%)	42 (1.3%)	.00001
	2	9351 (97%)	297 (3%)	
	3	4735 (94%)	321 (6%)	
	4	159 (84%)	30 (16%)	
Forced expiratory volume	%	72.6 ± 18.8	66.5 ± 17.8	.00001
Performance status	0	7375 (98%)	158 (2%)	.00001
	1	8233 (95.6%)	377 (4.4%)	
	2	1538 (93%)	120 (7%)	
	3	213 (86%)	35 (14%)	
Body mass index	≤17 kg/m^2	356 (90.4%)	38 (9.6%)	.0001
	18–21 kg/m^2	3206 (95.7%)	143 (4.3%)	
	22–26 kg/m^2	8126 (96%)	330 (4%)	
	>26 kg/m^2	5671 (97%)	179 (3%)	
Smoking	No	11,940 (96%)	506 (4%)	.01
	Yes	5419 (97%)	184 (3%)	
Chronic bronchitis	No	14,471 (96%)	549 (4%)	.009
	Yes	2888 (95%)	141 (5%)	
Coronary artery disease	No	16,158 (96.4%)	606 (3.6%)	.00001
	Yes	1201 (93.5%)	84 (6.5%)	
Cardiac arrhythmia	No	16,732 (96.3%)	646 (3.7%)	.00001
	yes	627 (93.4%)	44 (6.6%)	
Congestive heart failure	No	16,836 (96.3%)	640 (3.7%)	.0001
	Yes	523 (91.3%)	50 (8.7%)	
Peripheral vascular disease	No	15,583 (96.5%)	562 (3.5%)	.00001
	Yes	1776 (93.3%)	128 (6.7%)	
Pulmonary artery hypertension	No	17,314 (96%)	687 (4%)	.38
	Yes	45 (93.75%)	3 (6.25%)	
Alcoholism	No	16,416 (96.3%)	633 (3.7%)	.001
	Yes	943 (94.3%)	57 (5.7%)	
Cirrhosis	No	17,232 (96.2%)	682 (3.8%)	.2
	Yes	127 (94%)	8 (6%)	
Cerebral vascular events	No	16,814 (96%)	661 (4%)	.12
	Yes	545 (95%)	29 (5%)	
Diabetes mellitus	No	16,000 (96%)	626 (4%)	.16
	Yes	1359 (95.5%)	64 (4.5%)	
Renal insufficiency	No	17,192 (96%)	677 (4%)	.017
	Yes	167 (93%)	13 (7%)	
Coagulopathy	No	16,211 (96.3%)	627 (3.7%)	.01
	Yes	1148 (95%)	63 (5%)	
History of malignant disease	No	14,696 (96.4%)	554 (3.6%)	.002
	Yes	2663 (95%)	136 (5%)	
Prior thoracic surgery	No	16,442 (96.4%)	617 (3.6%)	.0001
	Yes	917 (92.6%)	73 (7.4%)	
Preoperative chemotherapy	No	16,095 (96.2%)	629 (3.8%)	.12
	Yes	1264 (95.4%)	61 (4.6%)	
Steroids	No	17,311 (96.2%)	688 (3.8%)	.9
	Yes	48 (96%)	2 (4%)	
Pulmonary embolism	No	17,196 (96.2%)	683 (3.8%)	.84
	Yes	163 (96%)	7 (4%)	
Valvular heart disease	No	17,299 (96.2%)	687 (3.8%)	.7
	Yes	60 (95%)	3 (4%)	

(*Continued*)

TABLE 1. Continued

Variables	Categories	Survivors (n = 17,359)	In-hospital mortality (n = 690)	*P* value
Pulmonary resection	Limited resection	1778 (97.6%)	43 (2.4%)	.00001
	Lobectomy	12,746 (97%)	409 (3%)	
	pneumonectomy	2835 (92.3%)	238 (7.7%)	
Side	Right	9753 (95.6%)	448 (4.4%)	.0001
	Left	7606 (97%)	242 (3%)	
Extended resection	No	15,805 (96.5%)	574 (3.5%)	.0001
	Yes	1554 (93%)	116 (7%)	
Mediastinal lymph node dissection	No	782 (97%)	24 (3%)	.34
	Sampling	1385 (96.5%)	50 (3.5%)	
	Lymphadenectomy	14,352 (96%)	580 (4%)	
Pathological stage	IA or IB	8613 (97.25%)	244 (2.75%)	.0001
	IIA or IIB	3051 (96%)	113 (4%)	
	IIIA or IIIB	4503 (94%)	265 (6%)	
	IV	1192 (96%)	48 (4%)	
Surgical resection margin status	R0	16,760 (96%)	652 (4%)	.006
	R1	457 (94.6%)	26 (5.4%)	
	R2	142 (92%)	12 (8%)	
No. of comorbidities per patient	None	6418 (98%)	123 (2%)	.00001
	1	5902 (96%)	231 (4%)	
	2	3884 (94%)	242 (6%)	
	3–4	1155 (92.5%)	94 (7.5%)	

the values for the area under the ROC curve, Somer's Dxy correlation, and R^2. Model 2, by taking into account the number of comorbidities per patient, considers that each of the comorbidities carries the same weight in the prediction of postoperative mortality, as demonstrated by Falcoz and colleagues.[15] However, their study was very different from ours because it involved procedures on the lung, mediastinum, and pleura for benign and malignant diseases.

TABLE 2. Logistic regression models including the number of comorbidities per patient (model 2) for prediction of in-hospital mortality

		Model 2		
Variables	Categories	Coefficients	95% CI	*P* value
Sex	Female vs male	–0.745	–1 to 0.49	.0001
Age	Increasing years	0.045	0.037 to 0.05	.0001
Side	Left vs right	–0.42	–0.62 to 0.21	.0001
ASA score	Increasing units	0.39	0.25 to 0.53	.0001
Performance status	Increasing units	0.3	0.17 to 0.41	.0001
Body mass index	$\leq$17 kg/m^2	Ref		
	18–21 kg/m^2	–0.89	–1.3 to 0.5	.0001
	22–26 kg/m^2	–1.18	–1.56 to 0.8	.0001
	>26 kg/m^2	–1.53	–1.9 to 1.13	.0001
FEV	Increasing %	–0.01	–0.016 to 0.005	.0001
Lobectomy	Yes vs no	0.56	0.23 to 0.89	.001
Pneumonectomy	Yes vs no	1.09	0.39 to 1.8	.002
Pneumonectomy · FEV	Interaction	0.01	0.0004 to 0.02	.04
Side · pneumonectomy	Interaction	–0.485	–0.83 to 0.14	.006
Extended resection	Yes vs no	–0.9	–1.74 to 0.06	.035
Extended resection · FEV	Interaction	0.018	0.006 to 0.029	.003
Stage	III vs (I or II or IV)	0.47	0.29 to 0.64	.0001
Stage	IV vs (I or II or III)	0.5	0.18 to 0.82	.002
No. of comorbidities per patient	0	Reference		
	1	0.5	0.27 to 0.73	.0001
	2	0.81	0.58 to 1	.0001
	3 or 4	0.95	0.66 to 1.25	.0001
Intercept		–6.64	–7.53 to 5.74	

CI, Confidence interval; *ASA*, American Society of Anesthesiologists; *FEV*, forced expiratory volume.

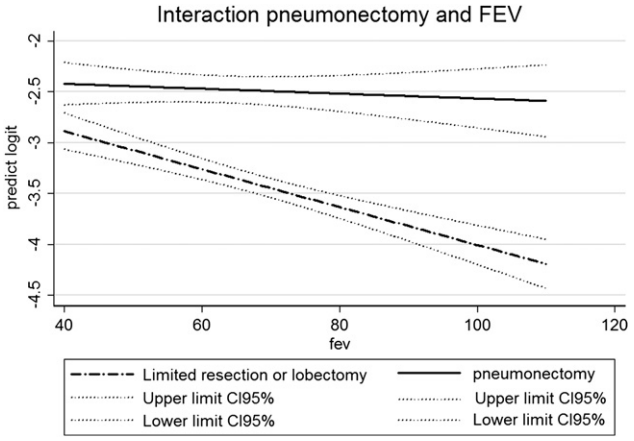

FIGURE 1. Interaction of pneumonectomy and forced expiratory volume (*FEV*; as a percentage). *CI*, Confidence interval.

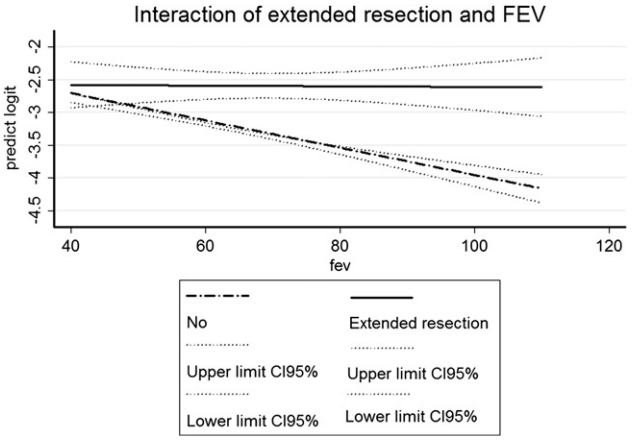

FIGURE 2. Interaction of extended resection and forced expiratory volume (*FEV*; as a percentage). *CI*, Confidence interval.

With regard to validation, model 2 yielded a slightly better calibration with a slope of 0.972 compared with model 1. The better validation of model 2 can be explained by the fact that parsimony is more important than accuracy.[13] Our preference is for model 2 because it is easier to use in practice to estimate adjusted postoperative mortality in lung resections for cancer. Among the methods for internal validation, the bootstrap method used in our study is the most efficient for estimations of internal validity in a predictive logistic regression model.[16] Ideally, this model should have been validated on an external validation dataset, but this would have been difficult because the patients are included in a national database, making it virtually impossible to find an independent validation dataset.

The values for the area under the ROC curve of the 2 models were estimated at around 0.78, indicating reasonable discriminatory ability. The DLCO, which is regarded as a prognostic variable,[17] could not be included in our model because this variable was not routinely used in French practice. In the multivariate model of laboratory values, Harpole and associates[18] included variables such as serum albumin and red blood cell count, which were not included in the Epithor database. However, the predictive values of the model of Harpole and associates are no better than the area under the ROC curve in our study. Indeed, the c-indexes for Harpole and associates' mortality models ranged from 0.749 to 0.729. Other authors[18,19] used intraoperative variables, such as operative time and blood transfusions. In these 2 studies the prediction models using these variables showed no significant superiority. Authors used individual comorbidities or the Charlson comorbidity index.[20-22] Strand and coworkers,[23] in their study based on a Swedish cancer registry, identified the Charlson comorbidity index as a prognostic factor with minimal effect on postoperative mortality. Other studies that validated the Charlson comorbidity index involved small samples.[20] Overall, our approach using the number of comorbidities per patient is easy to use and does not penalize the predictive model for in-hospital mortality.

The challenge for a national database is to find the right compromise between ease of use on a daily basis by including information with little variation and loss of precision. The possible lack of precision of these models could be explained by missing data, even though we tried to correct for this problem by using a multiple imputation method. This method assumes that missing values are independent of the occurrence of postoperative death. The estimation of missing values increases the power of the multivariate analysis, but it is possible that one cannot compensate for the lack of information. In the future, to reduce the amount of missing data, quality control audits will be implemented.

The few studies of multivariate models for the prediction of postoperative mortality in thoracic surgery did not take into account interactions with a clinical interest.[17,18,21] The adjusted level of risk of death for pneumonectomy was different depending on which lung was involved: patients with right-sided pneumonectomy were 3 times more likely to die in the hospital than were patients undergoing limited resection or lobectomy. Patients with left-sided pneumonectomy had an odds ratio of 1.78. Other interactions were taken into account in our models; these were extended resection with FEV and pneumonectomy with FEV. The risk of death in patients with extended resection or pneumonectomy was independent of FEV values. This was not the case for patients with simple resection or lobectomy or limited resection, in whom the risk of death correlated linearly with FEV values. The introduction of interaction terms in a multivariate model increases the number of parameters but probably improves its discriminative ability. Finally, because this model to estimate risk-adjusted mortality will be used in routine clinical practice, the use of variables, such as pathological stage, is justified.

The strength of this risk model is the size and quality of the dataset, which includes patients undergoing operations in France performed by private or public surgical teams. As

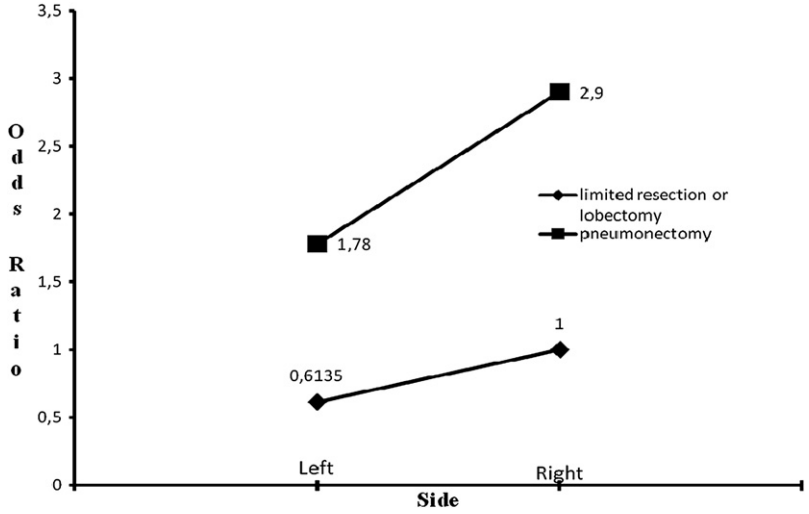

FIGURE 3. Interaction of side and type of lung resection.

shown by Falcoz and colleagues,[15] this database could be representative of practices in thoracic surgery in France. However, this report does have several limitations. As with all noncompulsory databases, there is a potential for incomplete submissions and for centers with poor outcomes to abstain from participating. The national Epithor database will develop a data verification system to reduce the amount of nonsensical and missing data by using an on-site audit procedure. Another indirect proof of the reliability of this database is that death rates in different types of lung resection are entirely consistent with the literature, especially in large databases.[22] Other predictive models developed from cancer registries or large databases[18,22–26] are comparable with our model.

Databases such as Epithor are the best tools to develop a predictive model that is easy to use by thoracic surgeons to estimate adjusted in-hospital mortality. However, the development of this type of database is an additional constraint for surgeons. Ultimately, it is difficult to create national databases, such as Epithor, with all the limitations described above. The advantage of this model lies in the fact that the variables are essentially clinical and easily collected every day.

The estimation of risk-adjusted in-hospital mortality is essential in that it will allow each thoracic surgery team to be compared with national benchmarks and will foster a process of improvement in the quality of surgical practices.[22] A quality program is established from the national database Epithor to allow each thoracic surgery center to compare their results with the national average by using graphic methods, such as the funnel plot.[27] Finally, the calculation of risk-adjusted in-hospital mortality for each thoracic surgery center will be done over 3 years to obtain enough events. In-hospital mortality is only one indicator of quality among others that will be developed. The software used by the Epithor database incorporates the calculation of risk-adjusted mortality and graphics for quality control of each thoracic surgery team. Moreover, the Epithor database is a useful for the accreditation of thoracic surgeons.

In conclusion, the development of a risk model is the necessary first step in the analysis of an approach to improve clinical practice in thoracic surgery. This risk model is easy to use with the Epithor database and is robust, as demonstrated by means of validation. The use of in-hospital mortality as an indicator of quality and the development of this risk model is part of a quality assurance program piloted by the French Society of Thoracic and Cardiovascular Surgery.

We are grateful to all the French thoracic surgeons who participated in this study: Dr Michel Alauzen (Montpellier), Dr Jean-François Andro (Quimper), Dr Maxime Aubert (Grenoble),

TABLE 3. Validation of 2 logistic regression models using the bootstrap method

	Model 1				Model 2			
	Training set	Test set	Optimism	Index corrected	Training set	Test set	Optimism	Index corrected
Dxy correlation	0.576	0.56	0.0157	0.551	0.567	0.553	0.0114	0.55
R^2	0.158	0.147	0.01	0.141	0.15	0.143	0.007	0.14
Intercept	0	–0.106	0.106	–0.106	0	–0.07735	0.07735	–0.07735
Slope	1	0.9608	0.04	0.9608	1	0.972	0.028	0.972
HL (χ^2)	9.8	9			11	6.7		
HL (P value)	.27	.5			.2	.75		

HL, Hosmer–Lemeshow goodness-of-fit statistic.

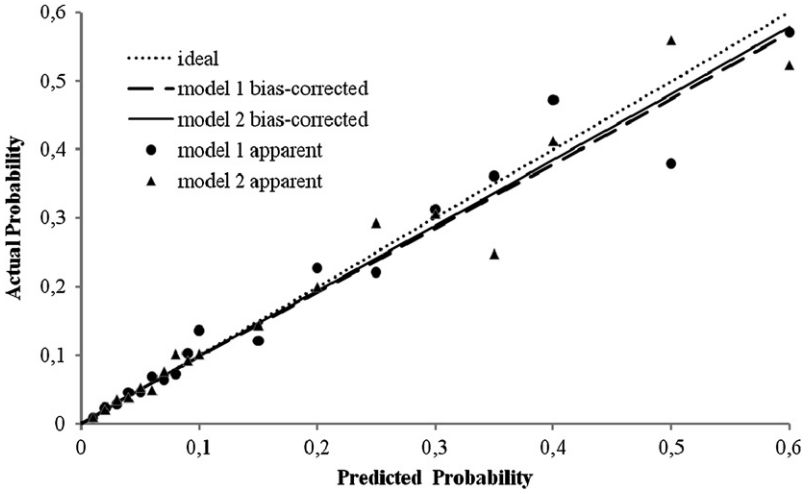

FIGURE 4. Calibration plot of observed versus predicted mortality for models 1 and 2.

Dr Jean Philippe Avaro (Marseille),Professor Jacques Azorin (Bobigny), Dr Patrick Bagan (Argenteuil), Dr François Bellenot (Cergry Pontoise), Professor Alain Bernard (Dijon), Dr Vincent Blin (Vannes), Dr Philippe Boitet (Harfleur), Dr Laurent Bordigoni (Toulon), Professor Jacques Borrelly (Nancy), Professor Pierre-Yves Brichon (Grenoble), Dr Gilles Cardot (Boulogne sur Mer), Dr Jean Michel Carrie (Saint Jean), Dr François Clement (Besançon), Professor Pierre Corbi (Poitiers),Professor Marcel Dahan (Toulouse), Dr Michel Debaert (Lille), Dr Bertrand Debrueres (Ploemeur), Dr Jean Dubrez (Bayonne), Dr Xavier Ducrocq (Strasbourg), Dr Antoine Dujon (Bois Guillaume), Professor Pascal Dumont (Tours), Dr Philippe Fernoux (Chalon sur Saône), Professor Marc Filaire (Clermont-Ferrand), Dr Eric Frassinetti (Chambéry), Dr Gil Frey (Saint Etienne), Dr Dominique Gossot (Paris), Professor Gilles Grosdidier (Nancy), Dr Benoit Guibert (Lyon), Dr Olivier Hagry (Chalon sur Saone), Dr Sophie Jaillard (Lille), Dr Jean-Marc Jarry (Aix en Provence), Dr David Kaczmarek (Saint Etienne), Dr Yves Laborde (Pau), Dr Bernard Lenot (Saint Brieuc), Dr Francis Levy (Bordeaux), Dr Laurent Lombart (Béziers), Dr Eric Marcade (Saint Grégoire), Dr Jean Paul Marcade (La Rochelle), Dr Jean Marzelle (Créteil), Professor Gilbert Massard (Strasbourg), Dr Florence Mazeres (Bayonne), Dr Eric Mensier (Lille), Dr David Metois (Orléans), Dr J. L. Michaud/E Paris (Nantes), Dr Philippe Mondine (Brest), Dr Michel Monteau (Reims), Dr Jean-Michel Moreau (Nantes), Professor Jérôme Mouroux (Nice), Dr Antoine Mugniot (Nantes), Dr Pierre Mulsant (Lyon), Dr Nidal Naffaa (Avignon), Dr Pierre Neveu (Talant), Dr Gérard Pavy (Arras), Professor Christophe Peillon (Rouen), Professor François Pons (Percy), Professor Henri Porte (Lille), Professor Jean-François Regnard (Paris), Professor Marc Riquet (Paris), Dr Babak Sadeghi Looyeh (Morlaix), Professor Pascal Thomas (Marseille), Professor Olivier Tiffet (Saint Etienne), Dr Bruno Tremblay (Meaux), Dr Jean Valla (Charenton le Pont), Professor Jean-François Velly (Pessac), Dr Bernard Wack (Metz), Dr Jean-Didier Wagner (Colmar), and Dr Didier Woelffe (Valenciennes).

References

1. Bernard A, Ferrand L, Hagry O, Benoit L, Cheynel N, Favre JP. Identification of prognostic factors determining risk groups for lung resection. *Ann Thorac Surg*. 2000;70:1161-7.
2. Ginsberg RJ, Hill LD, Eagan RT, et al. Modern Thirty-day operative mortality for surgical resections in lung cancer. *J Thorac Cardiovasc Surg*. 1983;86: 654-8.
3. Turna A, Solak O, Cetinkaya E, et al. Lactate dehydrodgenase levels predict pulmonary morbidity after lung resection for non-small cell lung cancer. *Eur J Cardiothorac Surg*. 2004;26:483-7.
4. Mishra PK, Pandey R, Shackcloth MJ, et al. Cardiac comorbidity is not a risk factor for mortality and morbidity following surgery for primary non-small cell lung cancer. *Eur J Cardiothorac Surg*. 2009;35:439-43.
5. Ferguson MK, Reeder LB, Mick R. Optimizing selection of patients for major lung resection. *J Thorac Cardiovasc Surg*. 1995;109:275-81.
6. Brunelli A, Fianchini A, Gesuita R, Carle F. POSSUM scoring system as an instrument of audit in lung resection surgery. *Ann Thorac Surg*. 1999;67: 329-31.
7. Birim O, Maat AP, Kappetein AP, van Meerbeeck JP, Damhuis RA, Bogers AJ. Validation of the Charlson comorbidity index in patients with operated primary non-small-cell lung cancer. *Eur J Cardiothorac Surg*. 2003;55: 231-40.
8. Wanatabe S, Asamura H, Suzuki K, Tsuchiya R. Recent results of postoperative mortality for surgical resections in lung cancer. *Ann Thorac Surg*. 2004;78: 999-1002.
9. French Society of Thoracic and Cardiovascular. Epithor database. Available at: http://www.epithor.net/. Accessed November 6, 2009.
10. TNM classification of malignant tumors. UICC. Edition 6. Ed Wiley; 2002. p. 99–107.
11. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*. 1999;8:3-15.
12. Hilbe JM. Logistic Regression Models. Ed A.Chapman & Hall Book; 2009. p. 1-637.
13. Harrell FE. Regression modeling strategies with applications to linear models, logistic regression and survival analysis. New-York: Springer; 2001. p. 1-330.
14. Hosmer DW, Lemeshow S. A goodness-of-fit test for multiple regression model. *Commun Stat*. 1980;A10:1043-69.
15. Falcoz PE, Conti M, Brouchet L, et al. The Thoracic Surgery Scoring System (thoracoscore): risk model for in-hospital death in 15,183 patients requiring thoracic surgery. *J Thorac Cardiothorac Surg*. 2007;133:325-32.
16. Steyerberg EW, Harrell FE, Gerard JJ, et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54:774-81.
17. Brunelli A, Rocco G. Internal validation of risk models in lung resection surgery: bootstrap versus training-and-test sampling. *J Thorac Cardiovasc Surg*. 2006; 131:1243-7.
18. Harpole DH, Decamp MM, Daley J, et al. Prognostic models of thirty-day mortality and morbidity after major pulmonary resection. *J Thorac Cardiovasc Surg*. 1999;117:969-79.
19. Yamashita S, Haga Y, Nemoto E, Nagai S, Ohta M. E-PASS (The Estimation of Physiologic Ability and Surgical Stress) scoring system helps the prediction of postoperative morbidity and mortality in thoracic surgery. *Eur Surg Res*. 2004; 36:249-55.

20. Ferguson MK, Durkin AE. A comparison of three scoring systems for predicting complications after major lung resection. *Eur J Cardiothorac Surg*. 2003;23: 35-42.
21. Birim O, Kappetein AP, Waleboer M, et al. Long-term survival after non-small cell lung cancer surgery: development and validation of a prognostic model with a preoperative and postoperative mode. *J Thorac Cardiovasc Surg*. 2006;132:491-8.
22. Boffa DJ, Allen MS, Grab JD, Gaissert HA, Harpole DH, Wright CD. Data from the Society of Thoracic Surgeons General Thoracic Surgery database: the surgical management of primary lung tumors. *J Thorac Cardiovasc Surg*. 2008;135:247-54.
23. Strand TE, Rostad H, Damhuis RA, Norstein J. Risk factors for 30-day mortality after resection of lung cancer and prediction of their magnitude. *Thorax*. 2007;62: 991-7.
24. Brunelli A, Berrisford RG, Rocco G, Varela G. European Society of Thoracic Surgeons Database Committee. The European Thoracic Database project: composite performance score to measure quality of care after major lung resection. *Eur J Cardiothorac Surg*. 2009;35:769-74.
25. Linden PA, Yeap BY, Chang MY, et al. Morbidity of lung resection after prior lobectomy: results from the Veterans Affairs National Surgical Quality Improvement Program. *Ann Thorac Surg*. 2007;83:425-31.
26. Kates M, Perez X, Gribetz J, Swanson SJ, McGinn T, Wisnivesky JP. Validation of a model to predict perioperative mortality from lung cancer resection in the elderly. *Am J Respir Crit Care Med*. 2009;179:390-5.
27. Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med*. 2005;24:1185-202.

APPENDIX 1. American Society of Anesthesiologists physical status classification system

Score 1	A normal healthy patient
Score 2	A patient with mild systemic disease
Score 3	A patient with severe systemic disease
Score 4	A patient with severe systemic disease that is a constant threat to life
Score 5	A moribund patient who is not expected to survive without operation
Score 6	A declared brain-dead patient whose organs are being removed for donor purposes
Performance status*	
Score 0	Asymptomatic
Score 1	Symptomatic but completely ambulatory
Score 2	Symptomatic, <50% in bed during the day
Score 3	Symptomatic, >50% in bed but not bedbound
Score 4	Bedbound
Score 5	Death

*Performance status is the Eastern Cooperative Oncology Group score.

APPENDIX 2. Definition of comorbidities

Comorbidity	Definition
Smoking	Chronic cigarette smokers included patients who were smoking in the 8 weeks preceding lung resection and patients not smoking for >8 weeks preceding lung resection
Chronic Bronchitis	Chronic cough present intermittently or every day or chronic sputum production
Coronary artery disease	Angina or previous myocardial infarction (>90 days) treated with percutaneous coronary intervention or surgery or medical treatment
Cardiac arrhythmia	Atrial flutter or atrial fibrillation or supraventricular tachycardia
Congestive heart failure	Previous symptoms of heart failure (acute pulmonary edema or paroxysmal nocturnal dyspnea or tachycardia or cardiomegaly on chest radiographic images) managed with medical treatment (Left ventricular function >30%)
Peripheral vascular disease	Peripheral artery occlusive disease treated with surgical intervention or medical treatment or abdominal aortic aneurysm
Pulmonary artery hypertension	A mean pulmonary artery pressure >25 mm Hg without right ventricular failure
Alcoholism	More than 2 alcoholic beverages per day for men and >1 alcoholic beverage per day for women
Cirrhosis	Replacement of liver tissue by fibrosis, scar tissue, and regenerative nodules with a Child–Pugh score of A
Cerebral vascular events	Stroke or transient ischemic attack caused by thrombosis or embolism or caused by hemorrhage
Diabetes mellitus	Type 1 or type 2 diabetes
Renal insufficiency	Chronic renal insufficiency treated with dialysis
Coagulopathy	Acquired or autoimmune or genetic clotting disorder or bleeding disorder
History of malignant disease	Malignancies treated within the previous 5 years other than the lung cancer being resected
Prior thoracic surgery	Previous ipsilateral or contralateral thoracotomy
Steroids	Long-term use of corticosteroids
Pulmonary embolism	History of pulmonary embolism
Valvular heart disease	Aortic valve stenosis or aortic insufficiency or mitral valve stenosis or mitral insufficiency treated by medication or surgery*

*Vahanian A, Baumgartner H, Bax J, et al. Guidelines on the management of valvular heart disease: The Task Force on the Management of Valvular Heart Disease of the European Society of Cardiology. *Eur Heart J*. 2007;28:230-68.

APPENDIX 3. Logistic regression model with independent predictive binary comorbidities (model 1) for prediction of in-hospital mortality

		Model 1		
Variables	**Categories**	**Coefficients**	**95% CI**	***P* value**
Sex	Female vs male	–0.74	–1 to 0.48	.0001
Age	Increasing years	0.048	0.039 to 0.057	.0001
Side	Left vs right	–0.43	–0.635 to 0.22	.0001
ASA score	Increasing units	0.39	0.25 to 0.53	.0001
Performance status	Increasing units	0.31	0.19 to 0.43	.0001
Body mass index	$\leq$17 kg/m^2	Reference		
	18–21 kg/m^2	–0.87	–1.27 to 0.48	.0001
	22–26 kg/m^2	–1.15	–1.54 to 0.77	.0001
	>26 kg/m^2	–1.46	–1.86 to 1.07	.001
FEV	Increasing %	–0.011	–0.017 to 0.005	.0001
Lobectomy	Yes vs no	0.63	0.29 to 0.97	.0001
Pneumonectomy	Yes vs no	1.07	0.36 to 1.78	.004
Pneumonectomy · FEV	Interaction	0.01	0.0012 to 0.02	.02
Side · pneumonectomy	Interaction	–0.47	–0.82 to 0.13	.008
Extended resection	Yes vs no	–0.92	–1.76 to 0.08	.03
Extended resection · FEV	Interaction	0.018	0.006 to 0.03	.002
Stage	III vs I or II or IV	0.49	0.31 to 0.66	.0001
Stage	IV vs I or II or III	0.53	0.2 to 0.85	.001
Chronic bronchitis	Yes vs no	0.2	0.002 to 0.4	.05
Cardiac arrhythmia	Yes vs no	0.38	0.05 to 0.72	.02
Coronary artery disease	Yes vs no	0.44	0.2 to 0.7	.0001
Congestive heart failure	Yes vs no	0.52	0.2 to 0.84	.001
Peripheral vascular disease	Yes vs no	0.39	0.18 to 0.6	.0001
Alcoholism	Yes vs no	0.645	0.34 to 0.94	.0001
History of malignant disease	Yes vs no	0.36	0.15 to 0.56	.001
Prior thoracic surgery	Yes vs no	0.82	0.54 to 1.09	.0001
Intercept		–6.7	–7.6 to 5.8	

CI, Confidence interval; *ASA*, American Society of Anesthesiologists; *FEV*, forced expiratory volume.