Thoracoscore predicts midterm mortality in patients undergoing thoracic surgery

Themistocles P. Chamogeorgakis, MD,[a] Cliff P. Connery, MD,[b] Faiz Bhora, MD,[b] Andy Nabong, MA,[b] and Ioannis K. Toumpoulis, MD[a]

Drs Chamogeorgakis and Toumpoulis (*left to right*)

Objective: Thoracoscore is the first multivariate model for the prediction of in-hospital mortality after general thoracic surgery. We aimed to evaluate the performance of Thoracoscore in predicting in-hospital and midterm all-cause mortality.

Methods: We retrospectively evaluated 1675 patients who underwent thoracic surgery (lung resections [n = 626], mediastinum [n = 535], pleura and pericardium [n = 268], esophagus [n = 88], chest wall [n = 90], trachea [n = 45], and other procedures [n = 23]) from October 2002 to March 2006 at a single institution. Midterm survival data (mean follow-up 25 ± 16 months) were obtained from the National Death Index. Kaplan–Meier survival plots of the quartiles of Thoracoscore were constructed and compared with the log–rank test with adjustment for trend.

Results: Starting from the lower-risk to the higher-risk quartile, the in-hospital mortality rates were 0% (0/418), 1% (4/415), 2.5% (11/435), and 9.6% (54/407). Thoracoscore was a strong independent predictor for in-hospital mortality (odds ratio 1.20, 95% confidence intervals 1.15-.25; $P < .001$). The 2-year survivals of the Thoracoscore quartiles were 98.7% ± 0.6%, 87.0% ± 1.8%, 73.8% ± 2.3%, and 54.8% ± 2.7%, respectively ($P < .0001$). Thoracoscore was a strong independent predictor for midterm mortality (hazard ratio 1.12, 95% confidence intervals 1.11-1.14; $P < .001$).

Conclusion: Thoracoscore is a good and useful clinical tool for preoperative prediction of in-hospital and midterm mortality among patients undergoing general thoracic surgery.

From the Department of Cardiothoracic Surgery, University of Athens School of Medicine, Attikon Hospital Center, Athens, Greece[a]; and the College of Physicians and Surgeons of Columbia University, Department of Cardiothoracic Surgery, St Luke's–Roosevelt Hospital Center, New York, NY.[b]

Received for publication Feb 19, 2007; revisions received June 9, 2007; accepted for publication June 21, 2007.

Address for reprints: Themistocles P. Chamogeorgakis, MD, Sofokleous 36, Voula, 16673, Greece (E-mail: thchamogeorgakis@yahoo.com).

J Thorac Cardiovasc Surg 2007;134:883-7

0022-5223/$32.00

doi:10.1016/j.jtcvs.2007.06.020

Thoracic surgery is lacking an accepted general risk model for in-hospital mortality. Thoracoscore is the first multivariate model, and it was derived from 15,183 patients who underwent thoracic surgery in 59 French hospitals.[1] Both operative and long-term mortality may be influenced by the same set of covariates, and we have demonstrated that EuroSCORE (one of the best established and validated risk stratification models in cardiac surgery) can be used for the prediction of long-term mortality in patients undergoing cardiac surgery.[2-4] We evaluated the Thoracoscore in predicting in-hospital and midterm mortality in our thoracic surgery database.

Patients and Methods

Patient Population and Data

From October 2002 to March 2006, 1675 patients underwent thoracic surgery at the St Luke's–Roosevelt Hospital Center of Columbia University. The records of patients were retrospectively reviewed, and we were able to collect all variables of Thoracoscore[1] except for dyspnea score, which was not available in our database. Thoracic operations included lung resections (n = 626), mediastinum (n = 535),

GTS

Abbreviation and Acronym
CI = confidence interval

pleura and pericardium (n = 268), esophagus (n = 88), chest wall (n = 90), trachea (n = 45), and other procedures (n = 23).

Data Analysis

Midterm patient mortality data were obtained from the United States Social Security Death Index database (http://ssdi.genealogy.rootsweb.com). The sensitivity of the National Death Index to identify deaths is between 92% and 99% depending on which identifiers are available.[5] Social Security number alone has the best accuracy of any combination of other identifiers (first initial, last name, day of birth, month of birth, year of birth, etc) with a sensitivity of 97% and a specificity of 99%.[5] In this study we used only Social Security numbers, which were available in most patients (98.3%), and this allowed avoiding use of patients' names. Moreover, patients without a Social Security number (n = 28) were censored at the time of discharge from the hospital. The index was queried in July 2006 and patients not found in the Index were assumed to be alive at that date.

Ethical Issues

No informed consent was obtained because the data used in this study had already been collected for clinical purposes. Furthermore, the present study did not interfere with the treatment of patients and the database was organized in a way that makes the identification of an individual patient impossible.

Statistical Analysis

Numerical variables were presented as the mean ± standard deviation, whereas discrete variables were summarized by percentages. We calculated the propensity score for in-hospital mortality according to the factors of Thoracoscore (except for dyspnea score) using its original β coefficients.[1] We also calculated the β coefficients of Thoracoscore in our database using multivariate logistic regression analysis.[6] The propensity score represents the probability that a patient would die during hospitalization. The predicted probability for each patient was calculated from the equation: Probability = odds/(1 + odds). The odds were calculated from the equation: Odds = exp(−7.3737 + [0.7679 if code of age was 1 or 1.0073 if code of age was 2] + [0.4505 × sex code] + [0.6057 × American Society of Anesthesiologists score code] + [0.6890 × Zubrod score code] + [0.8443 × priority of surgery code] + [1.2176 × procedure class code]

TABLE 1. Variables and their β coefficients of the Thoracoscore model as shown in the original model and in our thoracic surgery database (dyspnea score was not available in our database)

Variable		Value	Code	β Coefficient (original Thoracoscore, n = 10,122)	β Coefficient (our database, n = 1675)
Age (y)		<55	0		
		55-65	1	.7679	−.108
		≥65	2	1.0073	1.057
Sex	Female		0		
	Male		1	.4505	.402
ASA		≤2	0		
		≥3	1	.6057	1.909
Zubrod score		≤2	0		
		≥3	1	.6890	2.655
Dyspnea score		≤2	0		
		≥3	1	.9075	n/a
Priority of surgery	Elective		0		
	Urgent or emergency		1	.8443	.975
Procedure class	Other		0		
	Pneumonectomy		1	1.2176	3.248
Diagnosis group	Benign		0		
	Malignant		1	1.2423	.063
Comorbidity score		0	0		
		≤2	1	.7447	.093
		≥3	2	.9065	.761
Constant		n/a	n/a	−7.3737	−6.975

ASA, American Society of Anesthesiologists; *n/a,* not available.

+ [1.2423 × diagnosis group code] + [0.7447 if code of comorbidity was 1 or 0.9065 if code of comorbidity was 2]). A *C* statistic (or the area under the receiver operating characteristic curve) was used to assess the discriminatory ability of the model.[7] The area under the receiver operating characteristic curve was calculated as an index (*C* statistic) for how well the model could discriminate patients who lived and those who died during their hospitalization after thoracic surgery. The discriminative power of the model is thought excellent if the area under the receiver operating characteristic curve is greater than 0.80, very good if greater than 0.75, and good if greater than 0.70.[8] The calibration of the model was assessed by the Hosmer–Lemeshow goodness–of–fit statistic.[6] For the Hosmer–Lemeshow statistic, the predicted risks of individual patients were rank-ordered and divided into quartiles of roughly equal size, based on their predicted probability. Within each quartile of estimated risk, the number of predicted deaths was accumulated against the number of observed deaths; a $P > .05$ indicates acceptable calibration of the model. Kaplan–Meier survival plots[9] of the quartiles of modified Thoracoscore were constructed and compared with the log–rank test with adjustment for trend. Univariate logistic[6] and Cox[10] regression analysis were used to determine the odds ratio and hazard ratio of the propensity score for in-hospital and midterm mortality, respectively. All analyses were performed in SPSS 15.0 (SPSS, Inc, Chicago, Ill), and *P* values were 2-tailed. Hazard function curves of the quartiles were plotted and constructed with STATA/SE 9.1 (Stata Corporation, College Station, Tex).

Results

Table 1 shows the β coefficients of the original Thoracoscore model[1] and the β coefficients of the modified Thoracoscore (dyspnea score was not available) in our database. Age of 65 years or older, male sex, priority of surgery, and comorbidity of 3 or more showed similar β coefficients. Age between 55 and 65 years, comorbidity of 2 or less, and diagnosis group showed decreased β coefficients compared with the original Thoracoscore. Finally, American Society of Anesthesiologists score, Zubrod score, and procedure class showed increased β coefficients. There are two important reasons to explain these differences. First, the Thoracoscore model used in our study was modified by omitting one variable (dyspnea score); second, our study, which included 1675 patients, was underpowered compared with the original Thoracoscore study (10,122 patients analyzed for the development of the model).

The mean predicted probability of in-hospital mortality was 0.09% in the low-risk quartile, 0.35% in the mild-risk quartile, 1.60% in the medium-risk quartile, and 7.48% in the high-risk quartile. There were 54 (3.2%) in-hospital deaths. There was an increase in the presence of risk factors resulting in increased in-hospital mortality as the risk strat-

TABLE 2. Patient and disease characteristics of the quartiles according to factors used by Thoracoscore (except for dyspnea score)

Variable	Low-risk quartile (n = 418)	Mild-risk quartile (n = 415)	Medium-risk quartile (n = 435)	High-risk quartile (n = 407)	All patients (n = 1675)
Predicted probability (%), mean ± SD	0.09 ± 0.02	0.35 ± 0.13	1.60 ± 0.70	7.48 ± 4.69	2.34 ± 3.78
Predicted probability (%), range	0.06-0.15	0.18-0.66	0.71-2.90	2.97-31.38	0.06-31.38
Age, mean ± SD	32.7 ± 11.4	47.1 ± 11.2	61.7 ± 11.7	72.5 ± 8.0	53.5 ± 18.4
Male, n (%)	180 (43.1)	221 (53.3)	229 (52.6)	228 (56.0)	858 (51.2)
ASA score, mean ± SD	1.62 ± 0.58	2.51 ± 0.72	2.77 ± 0.67	3.00 ± 0.64	2.48 ± 0.84
Zubrod score, mean ± SD	0.27 ± 0.47	0.67 ± 0.75	0.84 ± 0.91	1.24 ± 1.32	0.75 ± 0.97
Priority of procedure					
Elective, n (%)	412 (98.6)	377 (90.8)	373 (85.7)	332 (81.6)	1522 (88.7)
Urgent or emergency, n (%)	6 (1.4)	38 (9.2)	62 (14.3)	75 (18.4)	183 (10.6)
Pneumonectomy, n (%)	0 (0)	0 (0)	9 (2.1)	12 (2.9)	21 (1.3)
Cancer					
Yes, n (%)	0 (0)	185 (44.6)	275 (63.2)	340 (83.5)	800 (47.8)
No, n (%)	418 (100)	230 (55.4)	160 (36.8)	67 (16.5)	875 (52.2)
Comorbidities					
0, n (%)	402 (96.2)	290 (69.9)	172 (39.5)	36 (8.8)	900 (53.7)
≤ 2, n (%)	16 (3.8)	109 (23.2)	231 (53.1)	266 (65.4)	622 (37.1)
≥ 3, n (%)	0 (0)	16 (3.9)	32 (7.4)	105 (25.8)	153 (9.2)
In-hospital mortality, n (%)	0 (0)	4 (1.0)	11 (2.5)	39 (9.6)	54 (3.2)
Midterm mortality, n (%)	5 (1.2)	59 (14.2)	110 (25.3)	185 (45.5)	359 (21.4)

SD, Standard deviation; *ASA,* American Society of Anesthesiologists.

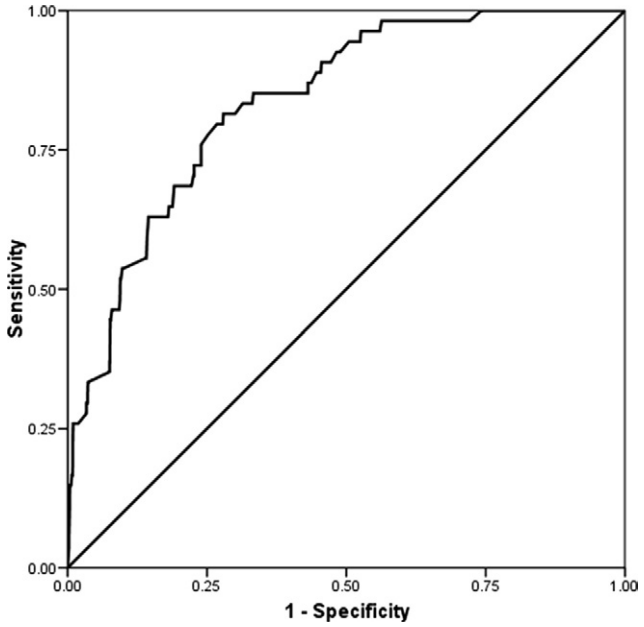

Figure 1. Receiver operating characteristic curve for in-hospital mortality of the modified Thoracoscore.

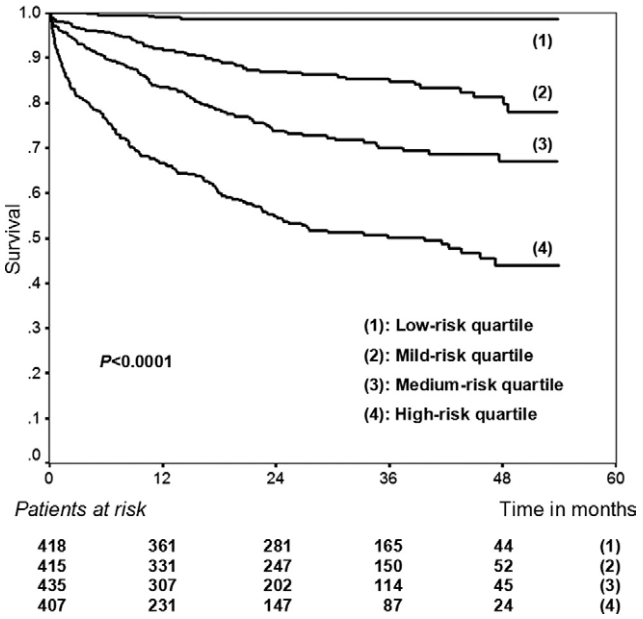

Figure 2. Kaplan–Meier survival plots of the quartiles according to the modified Thoracoscore. When the low-risk quartile was set as the reference group, the hazard ratio in the mild-risk quartile was 12.5 (95% CIs 5.0-31.2; $P < .001$), in the medium-risk quartile 24.6 (95% CIs 10.0-60.3; $P < .001$), and in the high-risk quartile 51.8 (95% CIs 21.3-125.9; $P < .001$).

ification grew (from 0% in the low-risk quartile to 9.6% in the high-risk quartile, Table 2). Modified Thoracoscore (predicted probability as calculated in our database) was a strong independent predictor for in-hospital mortality (odds ratio 1.20, 95% confidence intervals [CIs] 1.15-1.25; $P < .001$). The discriminatory ability of the modified model was excellent as measured by the C statistic (0.84, 95% CIs 0.79-0.88, Figure 1). The Hosmer–Lemeshow goodness-of-fit was not statistically significant ($P = .493$), indicating acceptable calibration of the model (Table 3).

During 43,001 person-months of follow-up, 359 (21.4%) deaths were recorded and there was an increase in midterm mortality as the risk stratification grew (Table 2). Kaplan–Meier survival plots of the modified Thoracoscore quartiles (Figure 2) diverged widely. The 2-year survivals of the quartiles were 98.7% ± 0.6%, 87.0% ± 1.8%, 73.8% ± 2.3%, and 54.8% ± 2.7% ($P < .0001$, log–rank test adjusted for trend). Similarly, higher-risk patients showed increased hazard estimate up to 36 months postoperatively compared with lower-risk patients (Figure 3). Univariate Cox regression analysis confirmed that modified Thoracoscore was a strong independent predictor for midterm mortality (hazard ratio 1.12, 95% CIs 1.11–1.14; $P < .001$).

TABLE 3. Predicted versus observed in-hospital mortality in the quartiles of the modified Thoracoscore

Quartiles of risk	No. of operations	Predicted mortality, n (%)	Observed mortality, n (%)
Low risk	418	0.71 (0.2)	0 (0)
Mild risk	415	2.77 (0.7)	4 (1.0)
Medium risk	435	11.34 (2.6)	11 (2.5)
High risk	407	39.18 (9.6)	39 (9.6)
Hosmer–Lemeshow x^2 (2 *df*)			1.265
P value			.493

df, Degrees of freedom.

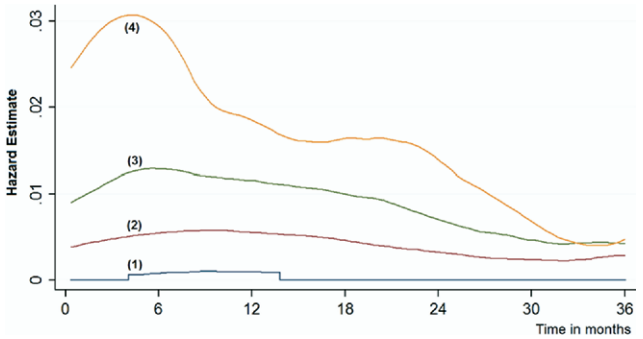

Figure 3. Hazard estimates of low-risk *(1)*, mild-risk *(2)*, medium-risk *(3)*, and high-risk *(4)* quartiles of the modified Thoracoscore.

Discussion

The Thoracoscore model was constructed to predict mortality during hospital stay among patients undergoing the whole range of thoracic surgery.[1] Such models may be used to assess the clinical outcomes of thoracic surgery in an objective risk-adjusted manner and allow useful comparisons to be made between countries, hospitals, and even individual surgeons. We confirmed the performance and calibration of Thoracoscore in our North American thoracic surgery database and we found a similar *C* index of 0.84. The risk for in-hospital mortality was increased by 20% for every 1% increase in the calculated modified Thoracoscore in our database (range 0.06%–31.38%). Thoracoscore works very well for in-hospital mortality and, in addition, we demonstrated that it also works very well for all-cause midterm mortality (mean follow-up 25 months). Groups at higher risk for in-hospital mortality continue to be at higher risk for midterm mortality. The risk for midterm mortality was increased by 12% for every 1% increase in the modified Thoracoscore.

Midterm patient follow-up represents another aspect of monitoring and prediction of patient outcomes, quality of care, and quality improvement in thoracic surgery. There are also additional reasons for estimating the risk for midterm mortality. These include determination of indications for surgery, proper informed consent, and identification of patients at high-risk for midterm mortality to have more careful follow-up and appropriate conservative therapy. Both early and late outcomes are important considerations, and optimization of prognosis may require separate models, although simple models covering early and late outcomes would be attractive. We showed clearly that modified Thoracoscore can also be used to forecast midterm mortality and can be used to inform the decision about whether to operate, taking into consideration both early and midterm mortality.

Our study has several limitations. First, this is a retrospective study. Nevertheless, the data on the risk factors analyzed have been collected with highly standardized methods for The Society of Thoracic Surgeons database. Second, we examined all-cause mortality and we were unable to determine the cause of death (thoracic or nonthoracic). However, for practical purposes, prediction of overall mortality is probably more important in the whole context of thoracic surgery after a midterm follow-up period. Third, this study refers to a single-center database, and it is likely that selection of patients for thoracic surgery, as well as race variation, which differ widely among thoracic surgery units, may be important determinants of early and midterm outcome. Fourth, dyspnea score was not available in our database, and this changed the β coefficients in the remaining variables. However, the modified Thoracoscore showed very good discriminative power in both in-hospital and midterm all-cause mortality. Finally, the inclusion in the final model of major postoperative complications may further improve its accuracy in predicting midterm mortality.[11] Modified Thoracoscore is a good clinical tool for preoperative prediction of in-hospital and midterm mortality among patients undergoing general thoracic surgery. This score needs further validation and refinements to adopt the changes in thoracic surgery, including minimally invasive and robotically assisted procedures.

References

1. Falcoz PE, Conti M, Brouchet L, Chocron S, Puyraveau M, Mercier M, et al. The Thoracic Surgery Scoring System (Thoracoscore): risk model for in-hospital death in 15,183 patients requiring thoracic surgery. *J Thorac Cardiovasc Surg.* 2007;133:325-32.
2. Toumpoulis IK, Anagnostopoulos CE, DeRose JJ, Swistel DG. European system for cardiac operative risk evaluation predicts long-term survival in patients with coronary artery bypass grafting. *Eur J Cardiothorac Surg.* 2004;25:51-8.
3. Toumpoulis IK, Anagnostopoulos CE, Toumpoulis SK, DeRose JJ, Swistel DG. EuroSCORE predicts long-term mortality after heart valve surgery. *Ann Thorac Surg.* 2005;79:1902-8.
4. Toumpoulis IK, Anagnostopoulos CE, Ioannidis JP, Toumpoulis SK, Chamogeorgakis T, Swistel DG, et al. The importance of independent risk-factors for long-term mortality prediction after cardiac surgery. *Eur J Clin Invest.* 2006;36:599-607.
5. Williams BC, Demitrack LB, Fries BE. The accuracy of the National Death Index when personal identifiers other than Social Security number are used. *Am J Public Health.* 1992;82:1145-7.
6. Hosmer DW, Taber S, Lemeshow S. The importance of assess fit of logistic regression models: a case study. *Am J Public Health.* 1991; 81:1630-5.
7. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143: 29-36.
8. Sweats JA. Measuring the accuracy of diagnostic systems. *Science.* 1988;240:1285-93.
9. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958;53:547-81.
10. Cox DR. Regression models and life-tables. *J R Stat Soc.* 1972;34: 187-220.
11. Chamogeorgakis T, Anagnostopoulos CE, Connery CP, Ashton RC, Dosios T, Kostopanagiotou G, et al. Independent predictors for early and midterm mortality after thoracic surgery. *Thorac Cardiovasc Surg.* 2007. In press.