Development of a patient-centered aggregate score to predict survival after lung resection for non–small cell lung cancer

Alessandro Brunelli, MD,[a,b] Michele Salati, MD,[a] Majed Refai, MD,[a] Francesco Xiumé, MD,[a] Rossana Berardi, MD,[c] Paola Mazzanti, MD,[c] and Cecilia Pompili, MD[a]

Objective: The objective of this analysis was to develop a survival aggregate score (SAS), including objective and subjective patient-based parameters, and assess its prognostic role after major anatomic resection for non–small cell lung cancer.

Methods: A total of 245 patients underwent major lung resections for non–small cell lung cancer with preoperative evaluation of quality of life (Short-Form 36v2 survey) and complete follow-up. The Cox multivariable regression and bootstrap analyses were used to identify prognostic factors of overall servival, which were weighted to construct the scoring system and summed to generate the SAS.

Results: Cox regression analysis showed that the factors negatively associated with overall survival and used to construct the score were 36-item short-form health survey physical component summary score less than 50 (hazard ratio [HR], 1.7; $P = .008$), aged older than 70 years (HR, 1.9; $P = .002$), and carbon monoxide lung diffusion capacity less than 70% (HR, 1.7; $P = .01$). Patients were grouped into 4 risk classes according to their SAS. The 5-year overall survival was 78% in class SAS0, 59% in class SAS1, 42% in class SAS2, and 14% in class SAS3 (log-rank test, $P < .0001$). SAS maintained its association with overall survival in patients with stages pT1 (log-rank test, $P = .01$), pT2 (log-rank test, $P = .02$), or pT3-4 (log-rank test, $P = .001$), and in those with stages pN0 (log-rank test, $P = .0005$) or pN1-2 (log-rank test, $P = .02$). The 5-year cancer-specific survival was 83% in class SAS0, 71% in class SAS1, 63% in class SAS2, and 17% in class SAS3 (log-rank test, $P < .0001$).

Conclusions: This system may be used to refine stratification of prognosis for clinical and research purposes. (J Thorac Cardiovasc Surg 2013;146:385-90)

GTS

Supplemental material is available online.

Earn CME credits at http://cme.ctsnetjournals.org

Objective measures of performance[1-3] and patient-perceived physical status[4-6] have been associated with survival time after surgery in patients with non–small cell lung cancer (NSCLC).

The objective of this study was to develop an aggregate scoring system, incorporating objective and subjective patient-based parameters, and to assess its association with prognosis in patients undergoing curative anatomic resection for NSCLC.

METHODS

This is a prospective longitudinal study on 407 patients who underwent lobectomy or pneumonectomy and systematic nodal dissection for NSCLC from January 2004 to December 2008. The stages of patients were determined according to the American Joint Committee on Cancer guidelines.[7]

A total of 311 patients underwent preoperative assessment of quality of life (QoL). Sixty-six of them were lost to follow-up. The remaining 245 patients (214 lobectomies and 31 pneumonectomies) were analyzed. A preliminary analysis showed that the patients with preoperative QoL assessment (included in the study) had similar overall 5-year survival (55% vs 56%; log-rank test, $P = .9$) compared with those without QoL assessment (excluded from the study).

This study was approved by the hospital Institutional Review Board, and all patients gave their informed consent to use their clinical data for scientific purposes. Operability exclusion criteria included a predicted postoperative forced expiratory volume in 1 second (FEV_1) and a predicted postoperative DLCO lower than 30%, in addition to an oxygen consumption per unit time peak lower than 10 mL/kg per minute. As a rule, operations were performed through a lateral muscle- and nerve-sparing thoracotomy by Board-certified thoracic surgeons.

Patients were extubated in the operating room and transferred to a dedicated thoracic ward. Postoperative management focused on early mobilization, antithrombotic and antibiotic prophylaxis, and physical and respiratory rehabilitation. Thoracotomy chest pain was assessed at least twice daily and controlled through a systemic, continuous infusion of nonopioid drugs. Therapy was titrated to achieve a visual analogue score lower than 5 (scale range, 0-10) during the first 48 to 72 hours. No formal

From the Divisions of Thoracic Surgery[a] and Clinical Oncology[c] and the Section of Minimally Invasive Thoracic Surgery,[b] Ospedali Riuniti Ancona, Ancona, Italy.

Disclosures: Authors have nothing to disclose with regard to commercial support.

Received for publication Jan 7, 2013; revisions received Feb 8, 2013; accepted for publication April 3, 2013; available ahead of print May 6, 2013.

Address for reprints: Alessandro Brunelli, MD, Division of Thoracic Surgery, Ospedali Riuniti Ancona, Via Conca 1, 60122 Ancona, Italy (E-mail: brunellialex@gmail.com).

0022-5223/$36.00

http://dx.doi.org/10.1016/j.jtcvs.2013.04.007

Abbreviations and Acronyms

CI = confidence interval
DLCO = carbon monoxide lung diffusion capacity
ECOG = Eastern Cooperative Oncology group
FEV_1 = forced expiratory volume in 1 second
FVC = forced vital capacity
HR = hazard ratio
LTFU = lost to follow-up
MCS = mental component summary
NSCLC = non–small cell lung cancer
PCS = physical component summary
QoL = quality of life
SAS = survival aggregate score
SF36 = 36-item short-form health survey
SF36v2 = Short-Form 36v2

preadmission or postdischarge physiotherapy or psychological supportive programs were administered.

Neurological or psychotropic personal medications, if present, were generally resumed the day after surgery.

QoL Assessment

QoL was assessed before the operation by the administration of the Short-Form 36v2 (SF36v2) survey. The SF36v2 questionnaire[8] is a generic instrument assessing 8 health physical and mental concepts (physical functioning, role limitation caused by physical problems, bodily pain, general health perception, vitality, social functioning, role limitation caused by emotional problems, and mental health). Scores were standardized to norms, and their weighted averages were used to create physical component summary (PCS) and mental component summary (MCS) scores on a standard scale. Norm-based scores have a mean of 50 and an SD of 10. As a consequence, for all health dimensions and component scales, any score lower than 50 is lower than the general population mean and each point represents one tenth of an SD. This allows for a direct comparison of measures among different populations and scales.

Survival

Follow-up was obtained by routine office visits, telephone contact, or data retrieved from the Regional Health Care System database. All patients were observed through April 2011. The cause of death was recorded based on physician report, ascertaining death for those cases retrieved from the Regional Health Care System database or reported by family members when telephone contact was used. The median follow-up calculated by using the reverse Kaplan-Meier method was 37 months. Perioperative mortality occurred in 3 patients in this series, and these events were counted as overall deaths and included in the survival analysis.

Statistical Analysis

The following baseline and tumor variables were tested for a possible association with survival: age, sex, body mass index, American Society of Anesthesiologist score, Eastern Cooperative Oncology group (ECOG) score, FEV_1 percentage, carbon monoxide lung diffusion capacity (DLCO percentage), FEV_1/forced vital capacity (FVC) ratio, history of coronary artery disease, renal insufficiency (creatinine level >2 mg/dL), preoperative hemoglobin level, histology (adenocarcinoma vs squamous, vs others), induction chemotherapy, and SF36v2 PCS and MCS QoL scales. For this study, PCS and MCS were categorized according to their values higher or lower than 50 (general population norms). *Survival* was defined as the interval between surgery to death or last contact. Patients who were not reported as dead at the time of the analysis were censored at the date they were last known to be alive.

Survival distribution was estimated by the Kaplan-Meier method. Significant differences in probability of surviving between the strata were evaluated by log-rank test.

Cox multivariable proportional hazards regression analysis was used to evaluate the effects of prognostic factors on survival. Predictors with $P < .1$ at univariable analysis (univariable Cox regression analysis for numeric variables and log-rank test for categorical variables) were used in the multivariable model.

Only 1 variable in a set of variables with a correlation coefficient greater than 0.5 was selected (by bootstrap procedure) and used in the regression model to avoid multicollinearity.

For the purpose of constructing the aggregate score, numeric variables were tested for a threshold effect and dichotomized by using receiver operating characteristic (curve) analysis (for identifying the best cutoff).

Bootstrap bagging with 1000 samples was used to assess the stability of the multivariable regression analysis predictors. In the bootstrap analysis, 1000 samples of the same size as the original population were drawn, with replacement from the original data set. A Cox proportional hazards regression model was repeated in each of these samples. If the final model variables occurred in most (>50%) of the bootstrap models, the original final regression model can be judged to be stable.[9] Only those variables with $P < .1$ and a bootstrap frequency of greater than 50% were retained in the final model and used to construct the score.

Construction of the Survival Aggregate Score

The scoring system was developed by proportional weighing of the significant predictor estimates, assigning a value of 1 to the smallest coefficient. An aggregate risk score was generated for each patient by summing each estimate. Finally, patients were grouped in classes of incremental risk according to their total score.

The stability of the risk score across multiple populations was further tested in 1000 bootstrapped samples drawn with replacement from the original data set.

A significance level of .05 was chosen to assess the statistical significance. All tests were performed on Stata 9.0 statistical software (Stata Corporation, College Station, Tex).

RESULTS

The characteristics of the patients included in this study are shown in Table 1.

The 5-year and median overall survival levels in the entire population were 55% and 75 months, respectively.

The following variables were associated with overall survival at univariable analysis and were used as independent predictors in the Cox proportional hazards regression analysis: age ($P = .008$), FEV_1 percentage ($P = .04$), FEV_1/FVC ratio ($P = .04$), DLCO percentage ($P = .02$), ECOG score ($P = .01$), and PCS less than 50 ($P = .006$).

To construct the aggregate score, receiver operating characteristic (curve) analysis was used to categorize the numeric variables. The following best thresholds were found: aged older than 70 years (c index, 0.58), FEV_1 lower than 80% (c index, 0.55), FEV_1/FVC less than 0.7 (c index, 0.58), DLCO less than 70% (c index, 0.58), and ECOG score greater than 2 (c index, 0.63).

Cox proportional hazards regression analysis showed that the factors negatively associated with overall survival in the