ABSTRACT

Background: This study aimed to assess the clinical relevance of the parsimonious Eurolung risk scoring system for predicting postoperative morbidity, mortality, and long-term survival in Korean patients with surgically resected non-small cell lung cancer.

Methods: This retrospective analysis used the data of patients who underwent anatomical resection for non-small cell lung cancer between 2004 and 2018 at a single institution. The parsimonious aggregate Eurolung score was calculated for each patient. The Cox regression model was used to determine the ability of the Eurolung scoring system for predicting long-term outcomes.

Results: Of the 7,278 patients in the study, cardiopulmonary complications and mortality occurred in 687 (9.4%) and 53 (0.7%) patients, respectively. The rate of cardiopulmonary complications and mortality gradually increased with the increase in the Eurolung risk scores (all \( P < 0.001 \)). When risk scores were grouped into four categories, the Eurolung scoring system showed a stepwise deterioration of overall survival with the increase in risk scores, and this association was statistically significant (\( P < 0.001 \)). Multivariate Cox analysis showed that the Eurolung scoring system, classified into four categories, was a significant prognostic factor of overall survival even after adjusting for covariates such as tumor histology and pathological stage (\( P < 0.001 \)).

Conclusion: Stratification based on the parsimonious Eurolung scoring system showed good discriminatory ability for predicting postoperative morbidity, mortality, and long-term survival in South Korean patients with surgically resected non-small cell lung cancer. This might help clinicians to provide a detailed prognosis and decide the appropriate treatment option for high-risk patients with non-small cell lung cancer.

Keywords: Eurolung Risk Score; Lung Cancer; Surgery; Postoperative Complication

INTRODUCTION

Despite continuous advances in diagnosis and treatment, lung cancer remains the most frequent cause of cancer-related deaths worldwide. In Korea, the mean age of patients...
undergoing surgery for lung cancer has increased gradually, and the number of comorbidities per patient has also increased. The proportion of patients with stage I cancer is more than 50% among those undergoing surgery for non-small cell lung cancer (NSCLC), and this proportion is still increasing. Considering the alternatives to surgery such as stereotactic body radiation therapy in high-risk patients, the importance of accurate prediction of the surgery-related morbidity and mortality is increasing.

In 2016, Eurolung risk models were developed to predict the risk of cardiopulmonary morbidity and mortality after surgery for NSCLC. These were based on the European Society of Thoracic Surgeons (ESTS) database comprising 48,000 patients. However, the disadvantage of these models was the need for numerous variables to calculate the risk; hence, their clinical utility was limited. To resolve this problem, the ESTS group recently created the parsimonious Eurolung risk models using simplified variables. Furthermore, these models have been reported to have significant reliability in predicting long-term survival outcomes. However, the limitation of these models is that they were developed based on the data of the European population and were externally validated on the same population. Thus, it is essential to perform an external validation of these models among other populations of different races, such as Asians.

In this study, we sought to evaluate the clinical relevance of the parsimonious Eurolung risk scores in a Korean population with surgically resected NSCLC and to assess their utility as predictive indicators for the prognosis.

**METHODS**

**Patients**

We enrolled patients who underwent pulmonary resection for primary NSCLC at the Asan Medical Center, Seoul, Korea, between January 2004 and December 2018. Patients who underwent wedge resection or surgery for diagnostic intent were excluded from the study. In addition, patients who received neoadjuvant therapy, which is thought to have different risks related to surgery, were excluded. Patients with any other concurrent malignancy were also excluded when analyzing the long-term survival outcomes. The pathological staging of NSCLC was based on the recommendations by the 8th edition American Joint Committee on Cancer in a retrospective manner. Tumor histology was categorized according to the World Health Organization classification.

Lobectomy with systematic lymph node dissection was adopted as a standard procedure for primary lung cancer; however, segmentectomy was performed in some old patients and those with borderline lung function and early-stage adenocarcinoma. Surgical resection was performed in patients with clinical stage I–IIIA, including N2 node metastasis. All patients were followed up either until death or the last follow-up date of the study (March 1, 2021).

Follow-up information of all the patients was obtained from the notes of the clinical follow-up, which was conducted every 3–6 months during the first 2 years after surgery and every 6–12 months thereafter. Chest computed tomography scans were performed at the time of clinical visits or at any time when disease recurrence was suspected. Treatment modalities and chemotherapeutic regimens in cases with relapse were determined at the discretion of the attending physician.
Definitions
The residual pulmonary function was estimated as follows: it is assumed that the lungs in a normal individual have a total of 19 segments (right upper lobe: 3, right middle lobe: 2, right lower lobe: 5, left upper lobe: 3, lingula: 2, lower left lobe: 4 segments), each contributing equally to ventilation that is 1/19 (nineteenth part). Thus, the predicted postoperative forced expiratory volume in 1 second (ppoFEV1) and diffusing capacity of carbon monoxide (DLCO) were calculated as follows: preoperative FEV1 or DLCO × (19 segments – the number of segments to be removed during the surgery) ÷ 19. Cardiopulmonary complications fundamentally complied with the terminology of the ESTS and Society of Thoracic Surgeons, which included the following: prolonged air leak (lasting more than 7 days), airway stenosis, atelectasis, pneumonia, acute respiratory distress syndrome, bronchopleural fistula, pulmonary embolism, pulmonary edema, respiratory failure requiring reintubation, empyema, recurrent laryngeal nerve palsy, phrenic nerve palsy, chylothorax, postoperative bleeding, atrial arrhythmia, myocardial infarction, stroke, and acute renal failure.

Extended resection was defined as follows: 1) chest wall resection, 2) Pancoast tumors, 3) resection of the atrium, superior vena cava, aorta, diaphragm, or vertebra, 4) bronchial sleeve resection, 5) pleuropneumonectomy, 6) sleeve pneumonectomies, or 7) intrapericardial pneumonectomy. Mortality was defined as death within 30 days after surgery or death occurring at any time during the same hospital stay. The overall survival (OS) was calculated as the time interval between the date of surgery and the date of death, which was determined by reviewing the patient records from the Korean National Security Death Index Database. Aggregate Eurolung1 (morbidity) and Eurolung2 (mortality) scoring systems were developed by assigning proportional weightages of the predictors estimates, which are listed in the Eurolung model, assigned a value of 1 to the smallest coefficient. Thus, the Eurolung1 score was calculated as a sum of the points based on the following variables: 1 point for age > 70 years and ppoFEV1 < 70%; 1.5 points for male sex and extended resection; and 2 points for open surgery (as opposed to minimally invasive surgery) (Table 1). For the Eurolung2 system, age > 70 years and ppoFEV1 < 70% were assigned 1 point; male sex, open surgery (as opposed to minimally invasive surgery), and body mass index (BMI) < 18.5 kg/m² were assigned 2.5 points; and pneumonectomy was assigned 3 points (Table 1).
Statistical analyses
Continuous variables are presented as means and standard deviations, and categorical variables are presented as frequencies and percentages. The normality of individual distributions of the parameters was assessed using the Shapiro–Wilk test. Student’s t-test or the Wilcoxon rank-sum test was used for the comparison of continuous variables between the two groups, and the χ² or Fisher’s exact test was used for categorical variables between the two groups. The area under the receiver operating characteristic curve (AUC) is presented as a statistical indicator for quantifying the discrimination ability of the scoring system. The Hosmer–Lemeshow test for the goodness-of-fit was used for calibration of the logistic regression model.

To determine the long-term prognostic ability of the Eurolung2 scoring system, we conducted a survival analysis of patients who underwent complete resection for stage I, II, and III NSCLC. The Kaplan-Meier method was used to analyse OS, and their differences were assessed using the log-rank test. For multiple comparisons of the survival curves (≥ 3 curves), Bonferroni correction was applied to calculate the P values in the log-rank test. A Cox proportional-hazards model was used for the univariate and multivariate analyses to identify the prognostic ability of the Eurolung2 scoring system.

All statistical calculations were performed with R, version 4.0.2 (The R Foundation for Statistical Computing, Vienna, Austria), using the Survival, ggplot2, GGally, survminer, and rms packages. P < 0.05 was considered significant.

Ethics statement
This retrospective study was approved by the Asan Medical Center Ethics Committee/Review Board, and this center waived the need for informed consent (approval No. 2021-0544).

RESULTS

Patient characteristics
The median follow-up period was 53.6 ± 35.9 months. The clinicopathological characteristics of the patients are summarized in Table 2. A total of 7,278 patients were enrolled in this study. There were 1,531 (21.0%) patients aged > 70 years and 159 (2.2%) patients with BMI < 18.5 kg/m². In addition, 3,659 patients had ppoFEV1 < 70%. Lobectomy was performed in 5,957 patients (81.8%), and video-assisted thoracoscopic surgery was performed in 4,646 patients (63.8%). In terms of the pathological stage, there were 3,999 (54.9%), 1,554 (21.4%), 1,560 (21.4%), and 159 (2.2%) patients with stage I, II, III, and IV disease, respectively.

Short-term outcomes
Regarding cardiopulmonary complications, 687 (9.4%) patients had a cardiopulmonary event regardless of event grade, the details of which are summarized in Supplementary Table 1. Fifty-three (0.7%) patients died within 30 days after surgery or at any time during the same hospital stay. The rate of cardiopulmonary complications gradually increased with the increase in the Eurolung1 risk score (P < 0.001) (Fig. 1A). A similar trend was observed in the mortality rate (P < 0.001) (Fig. 1B). The AUC was 0.623 and 0.815 for the Eurolung1 and Eurolung2 scores, respectively (Fig. 2). The Hosmer–Lemeshow tests of the two models were not significant (P = 0.442 for the Eurolung1 and 0.390 for the Eurolung2 scoring system).
Long-term outcomes

According to the Kaplan-Meier estimates, survival curves stratified by the Eurolung2 scoring system showed a stepwise deterioration of OS with the increase in risk scores (Fig. 3A). When the risk scores were grouped into four categories, each group had a significantly different prognosis (Fig. 3B). Subgroup analyses were performed for four categories of the Eurolung2 scoring system in patients with stage I, II, and III NSCLC (Fig. 4). Although the difference in survival between the patients with scores 3–5 and 5.5–6.5 was not significant in stages II and III, the Eurolung2 risk scoring system still showed a favorable discrimination ability in all the stages.

In accordance with the Cox proportional hazard analysis, the Eurolung2 scoring system, classified into four categories, was found to be a significant prognostic factor for OS even after adjusting for confounders such as tumor histology and pathological stage in patients who underwent complete resection ($P < 0.001$) (Table 3). In addition, the higher the risk

**Table 2. Baseline characteristics of the patients included in the study ($N = 7,278$)**

| Variable                  | Values                        |
|---------------------------|-------------------------------|
| Age, yr                   | 62.3 ± 9.5                    |
| Age > 70 yr               | 1,531 (21.0)                  |
| Sex (male)                | 4,436 (61.0)                  |
| BMI                       | 24.1 ± 3.0                    |
| BMI < 18.5 kg/m²          | 159 (2.2)                     |
| CVD                       | 481 (6.6)                     |
| CAD                       | 192 (2.6)                     |
| CKD                       | 65 (0.9)                      |
| Pulmonary function        |                               |
| FEV1 (%)                  | 89.0 ± 15.9                   |
| ppoFEV1 < 70%             | 3,659 (50.3)                  |
| DLCO (%)                  | 88.1 ± 17.7                   |
| ppoDLCO < 70%             | 3,904 (53.6)                  |
| Resection extent          |                               |
| Segmentectomy             | 605 (8.3)                     |
| Lobectomy                 | 5,957 (81.8)                  |
| Bilobectomy               | 314 (4.3)                     |
| Pneumonectomy             | 232 (3.2)                     |
| Sleeve lobectomy          | 170 (2.3)                     |
| Surgical approach         |                               |
| VATS                      | 4,646 (63.8)                  |
| Thoracotomy               | 2,329 (32.0)                  |
| Conversion to thoracotomy | 303 (4.2)                     |
| Resection status          |                               |
| Complete                  | 6,923 (95.1)                  |
| Microscopic               | 270 (3.7)                     |
| Macroscopic               | 85 (1.2)                      |
| Histology                 |                               |
| Adenocarcinoma            | 5,061 (69.6)                  |
| Squamous cell carcinoma   | 1,728 (23.7)                  |
| Others                    | 487 (6.7)                     |
| Pathological stage        |                               |
| Stage I                   | 3,999 (54.9)                  |
| Stage II                  | 1,554 (21.4)                  |
| Stage III                 | 1,560 (21.4)                  |
| Stage IV                  | 159 (2.2)                     |
| Not assessable            | 6 (0.1)                       |

Values are presented as numbers (%) or means ± standard deviations, unless otherwise indicated. BMI = body mass index, CVD = cerebrovascular disease, CAD = coronary artery disease, CKD = chronic kidney disease, FEV1 = forced expiratory volume in 1 second, ppoFEV1 = predicted postoperative forced expiratory volume in 1 second, DLCO = diffusing capacity of carbon monoxide, ppoDLCO = predicted postoperative diffusing capacity of carbon monoxide, VATS = video-assisted thoracic surgery.
score, the higher was the hazard ratio, indicating a gradual deterioration in the prognosis with an increase in the risk score (Table 3).

**DISCUSSION**

In this study, we demonstrated the clinical relevance of the parsimonious Eurolung scoring system proposed by the ESTS group. The findings from the current study indicate that the parsimonious Eurolung1 (morbidity) and Eurolung2 (mortality) scoring systems have a good discriminatory ability for cardiopulmonary postoperative morbidity and mortality in the Korean population as well as in the European population with surgically resected NSCLC. Furthermore, the Eurolung2 scoring system is an independent prognosticator regardless of tumor histology and stage and can predict the long-term prognosis.
Considering that the mean age and number of comorbidities of patients undergoing lung cancer surgery are increasing, precise risk predictions for morbidity and mortality before surgery are becoming extremely important. The Eurolung risk model is one of the latest prediction models targeted for patients who have undergone surgery for lung cancer and is based on 48,000 cases registered in the ESTS database.

In 2016, the original model was developed; however, the complexity due to the requirement of many variables has limited its access in the clinical field. In 2019, an updated model with simplified variables was published, which retained the predictive abilities of the original model. However, this model has not been externally validated in an independent sample of patients so far. In this study, we performed external validation of the parsimonious Eurolung risk score.
Validation of Parsimonious Eurolung Risk Score

Notably, we adopted the cut–off values developed in parsimonious Eurolung models, such as age > 70 years, ppoFEV1 < 70%, and BMI < 18.5 kg/m². To enhance simplicity and accessibility, continuous variables are often divided into two categories based on an optimal cut-off value in clinical practice. However, the cut-off value is a statistic estimated from the original data, which can be unstable and have a large variation. In addition, this estimate has a possibility of overfitting that is characterised by high accuracy for a classifier when evaluated on the training set but low accuracy when evaluated on a separate test set. Nonetheless, the cut-off values of the Eurolung models were still valid in our database. We believe that this is because these variables are standardized with respect to the method for measurement and can be observed consistently without the influence of other factors, thus reducing the variability.

Regarding the long-term survival outcomes, it is surprising that even for the same pathological stage, the 5-year survival rate varied widely according to the Eurolung score (68.7% to 93.1% in stage I, 51.4% to 78.2% in stage II, and 31.8% to 59.9% in stage III) (Fig. 4). In the Eurolung model, several variables (age, sex, ppoFEV1, and thoracotomy)

Fig. 4. Survival curves for the overall survival according to the parsimonious Eurolung risk scoring systems grouped into four categories in (A) pathological stage I, (B) stage II, and (C) stage III disease.
were commonly associated with morbidity and mortality, whereas some variables were characterized as morbidity (extended resection) or mortality (BMI and pneumonectomy). 4

We suppose that a higher aggregate score representing the extent of the physiologic condition and surgical extension is associated with the postoperative complication and affects OS. Considering that the Eurolung score is an independent prognostic factor for long-term survival, this system provides a more precise and detailed prognosis for patients with NSCLC. Furthermore, the Eurolung scoring system can help in deciding whether surgery should be performed and also the surgical extent, surgical approach, and postoperative surveillance.

The current study has some limitations: 1) This was a retrospective study at a single institution in Korea, which does not represent the general Korean population. 2) selection bias is inherent in a retrospective analysis (e.g. the indication for surgery might vary depending on the institution); 3) a relatively small number of patients were classified in the highest-risk class; therefore, all patients with a score of 7 or higher were classified into a single group, and survival analysis was performed in a total of four groups instead of the original seven groups; and 4) although a significant association was observed between the rate of mortality and risk score ($P < 0.001$), the rate of mortality was only 0.7% ($n = 53$) in the

Table 3. Univariate and multivariate analysis of the long-term survival outcomes in patients with complete resection

| Variables                  | Univariate analysis | Multivariate analysis |
|----------------------------|---------------------|-----------------------|
| Age > 70 yr                | 2.26 (2.04–2.50)    | < 0.001               |
| Sex (male)                 | 1.90 (1.71–2.10)    | < 0.001               |
| BMI < 18.5 kg/m²           | 1.61 (1.26–2.07)    | < 0.001               |
| CVD                        | 1.44 (1.22–1.70)    | < 0.001               |
| CAD                        | 1.53 (1.18–1.99)    | 0.001                 |
| CKD                        | 2.27 (1.56–3.30)    | < 0.001               |
| Pulmonary function         |                     |                       |
| ppoFEV1 < 70%              | 1.56 (1.42–1.71)    | < 0.001               |
| ppoDLCO < 70%              | 1.70 (1.55–1.87)    | < 0.001               |
| Resection extent           |                     |                       |
| Segmentectomy              | 1                   | Reference             |
| Lobectomy                  | 2.19 (1.65–2.90)    | < 0.001               |
| Bilobectomy                | 3.50 (2.51–4.87)    | < 0.001               |
| Pneumonectomy              | 4.24 (3.03–5.95)    | < 0.001               |
| Sleeve lobectomy           | 2.78 (3.3–5.95)     | < 0.001               |
| Surgical approach          |                     |                       |
| Thoracotomy                | 1                   | Reference             |
| VATS                       | 0.44 (0.40–0.48)    | < 0.001               |
| Conversion to thoracotomy  | 0.77 (0.62–0.96)    | 0.021                 |
| Histology                  |                     |                       |
| Adenocarcinoma             | 1.82 (1.65–2.01)    | < 0.001               |
| Squamous cell carcinoma    | 1.60 (1.34–1.90)    | < 0.001               |
| Others                     |                     |                       |
| Pathological stage         |                     |                       |
| Stage I                    | 1                   | Reference             |
| Stage II                   | 2.49 (2.21–2.80)    | < 0.001               |
| Stage III                  | 4.65 (4.17–5.19)    | < 0.001               |
| Mortality score            |                     |                       |
| 0–2.5                     | 1                   | Reference             |
| 3–5                       | 2.06 (1.83–2.32)    | < 0.001               |
| 5.5–6.5                   | 2.84 (2.49–3.25)    | < 0.001               |
| 7–12.5                    | 4.66 (4.01–5.42)    | < 0.001               |

OR = odds ratio, CI = confidence interval, BMI = body mass index, CVD = cerebrovascular disease, CAD = coronary artery disease, CKD = chronic kidney disease, ppoFEV1 = predicted postoperative forced expiratory volume in 1 second, ppoDLCO = predicted postoperative diffusing capacity of carbon monoxide, VATS = video–assisted thoracic surgery.
Validation of Parsimonious Eurolung Risk Score

overall cohort. Thus, the results should be interpreted with caution due to the small number of observed events. A multicentre study is necessary to overcome the limitations of this study, and a validation study is still suggested.

In conclusion, we demonstrated that stratification according to the parsimonious version of the Eurolung scoring system has good discriminatory ability in terms of postoperative morbidity and mortality in South Korean patients. This system might help clinicians in estimating a detailed prognosis and deciding the appropriate treatment in high-risk patients with NSCLC.

ACKNOWLEDGMENTS

The authors would like to thank Editage (www.editage.co.kr) for the English language editing.

SUPPLEMENTARY MATERIAL

Supplementary Table 1
Details of postoperative cardiopulmonary complications (n = 687)

Click here to view

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68(1):7-30.
   PUBMED | CROSSREF
2. Yun JK, Lee HP, Lee GD, Kim HR, Kim YH, Kim DK, et al. Recent trends in demographics, surgery, and prognosis of patients with surgically resected lung cancer in a single institution from Korea. J Korean Med Sci 2019;34(45):e291.
   PUBMED | CROSSREF
3. Brunelli A, Salati M, Rocco G, Varela G, Van Raemdonck D, Decaluwe H, et al. European risk models for morbidity (EuroLung1) and mortality (EuroLung2) to predict outcome following anatomic lung resections: an analysis from the European Society of Thoracic Surgeons database. Eur J Cardiothorac Surg 2017;51(3):490-7.
   PUBMED
4. Brunelli A, Cicconi S, Decaluwe H, Szanto Z, Falcoz PE. Parsimonious Eurolung risk models to predict cardiopulmonary morbidity and mortality following anatomic lung resections: an updated analysis from the European Society of Thoracic Surgeons database. Eur J Cardiothorac Surg 2020;57(3):455-61.
   PUBMED
5. Brunelli A, Chaudhuri N, Kefaloyannis M, Milton R, Pompili C, Tcherveniakov P, et al. Eurolung risk score is associated with long-term survival after curative resection for lung cancer. J Thorac Cardiovasc Surg 2021;161(3):776-86.
   PUBMED | CROSSREF
6. Deterbeck FC, Boffa DJ, Kim AW, Tanoue LT. The eighth edition lung cancer stage classification. Chest 2017;151(1):193-203.
   PUBMED | CROSSREF
7. Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JH, Beasley MB, et al. WHO Panel. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. J Thorac Oncol 2015;10(9):1243-60.
   PUBMED | CROSSREF
8. Cukic V. Preoperative prediction of lung function in pneumonectomy by spirometry and lung perfusion scintigraphy. Acta Inform Med 2012;20(4):221-5.
   PUBMED | CROSSREF
9. Fernandez FG, Falcoz PE, Kozower BD, Salati M, Wright CD, Brunelli A. The Society of Thoracic Surgeons and the European Society of Thoracic Surgeons general thoracic surgery databases: joint standardization of variable definitions and terminology. *Ann Thorac Surg* 2015;99(1):368-76.

10. Zhang Z. Estimating the optimal cutoff point for logistic regression. https://scholarworks.utep.edu/cgi/viewcontent.cgi?article=2564&context=open_etd. Updated 2018. Accessed January 1, 2022.

11. Altman DG. Categorising continuous variables. *Br J Cancer* 1991;64(5):975.

12. Subramanian J, Simon R. Overfitting in prediction models - is it a problem only in high dimensions? *Contemp Clin Trials* 2013;36(2):636-41.