Clinical characteristics and prognosis of basaloid squamous cell carcinoma of the lung: a population-based analysis

Guangda Yuan [1] , **Cheng Zhan** [2] , **Yiwei Huang** [2] , **Donglin Zhu** [1] , **Hongya Xie** [1] , **Tengteng Wei** [1] , **Tao Lu** [2] ,
Qun Wang [2] , **Yong Yang** Corresp., [1] , **Yimeng Zhu** Corresp. [1]

[1] Department of Thoracic Surgery, The Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou,Jiangsu, China

[2] Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, Shanghai, China

Corresponding Authors: Yong Yang, Yimeng Zhu
Email address: yangy_suzhou@126.com, zhuym_suzhou@126.com

Background: This study analyzed the clinical features and prognosis of basaloid squamous cell carcinoma of the lung (BSC), and constructed a nomogram to predict the prognoses of patients. **Methods**: The information of pure BSC patients was obtained in the Surveillance, Epidemiology, and End Results database between 2004 and 2015. Then, it was evaluated, and compared with the data of lung squamous cell carcinoma (SCC), lung large cell carcinoma (LCC) and lung adenocarcinoma (LAC) patients. Subsequently, we used univariate and multivariate analyses to investigate the independent factors related to the prognoses of patients with BSC and constructed a nomogram to verify the prognoses. **Results**: A total of 425 patients diagnosed with BSC were enrolled. Compared with patients with SCC, LCC and LAC, the mean survival time of BSC patients was better than all of them. Compared with SCC, there were significant difference between the characteristics of grade (P<0.001), total stage (P<0.001), T stage (P<0.001), N stage (P<0.001), M stage (P<0.001), surgery (P<0.001), radiotherapy(P<0.001), and chemotherapy (P<0.001), while BSC also had significantly different clinical characteristics from LCC and LAC. Univariate and multivariate survival analyses showed that age (P<0.001), T stage (P<0.001), N stage (P=0.009), M stage (P<0.001), and surgery (P<0.001) were independent prognostic factors of BSC. The survival of patients undergoing lobectomy was significantly better than sublobar resection, with an OR of 0.389 (0.263–0.578). We constructed a nomogram with a C-index of 0.750 (95% confidence interval) based on the results of multivariate analysis. The calibration curves based on nomogram scores indicated that the nomogram could accurately predict the prognosis of patients. **Conclusions**: BSC had unique clinical and prognostic features. T stage, N stage, M stage, age, and surgery were independently associated with overall survival (OS). Lobectomy was a relative ideal choice for patients with BSC. The nomogram effectively predicted the OS at 1-, 3-, and 5-years.

1 **Clinical characteristics and prognosis of basaloid squamous cell carcinoma of**

2 **the lung: a population-based analysis**

3 Guangda Yuan[1*], Cheng Zhan[2*], Yiwei Huang[2*], Donglin Zhu[1], Hongya Xie[1], Tengteng Wei[1],

4 Tao Lu[2], Qun Wang[2], Yong Yang[1] & Yimeng Zhu[1]

5 1 Department of Thoracic Surgery, the Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou,

6 Jiangsu, China

7 2 Department of Thoracic Surgery, Zhongshan Hospital, Fudan University, Shanghai, China

8 [*] Guangda Yuan, Cheng Zhan and Yiwei Huang contributed equally to this work.

9 **Correspondence**

10 Yong Yang & Yimeng Zhu, Department of Thoracic Surgery, The Affiliated Suzhou Hospital of Nanjing

11 Medical University, Suzhou, Jiangsu, China.

12 Tel: 86 512 6236 4519

13 Fax: 86 512 6236 4519

14 Email: zhuym_suzhou@126.com

15   yangy_suzhou@126.com

16 **Abstract**

17 **Background**: This study analyzed the clinical features and prognosis of basaloid squamous cell carcinoma of

18 the lung (BSC), and constructed a nomogram to predict the prognoses of patients.

19 **Methods**: The information of pure BSC patients was obtained in the Surveillance, Epidemiology, and End

20 Results database between 2004 and 2015. Then, it was evaluated, and compared with the data of lung

21 squamous cell carcinoma (SCC), lung large cell carcinoma (LCC) and lung adenocarcinoma (LAC) patients.

22   Subsequently, we used univariate and multivariate analyses to investigate the independent factors related to the

23   prognoses of patients with BSC and constructed a nomogram to verify the prognoses.

24   **Results**: A total of 425 patients diagnosed with BSC were enrolled. Compared with patients with SCC, LCC

25   and LAC, the mean survival time of BSC patients was better than all of them. Compared with SCC, there were

26   significant difference between the characteristics of grade (P<0.001), total stage (P<0.001), T stage (P<0.001),

27   N stage (P<0.001), M stage (P<0.001), surgery(P<0.001), radiotherapy(P<0.001), and chemotherapy (P<0.001),

28   while BSC also had significantly different clinical characteristics from LCC and LAC. Univariate and

29   multivariate survival analyses showed that age (P<0.001), T stage (P<0.001), N stage (P=0.009), M stage

30   (P<0.001), and surgery (P<0.001) were independent prognostic factors of BSC. The survival of patients

31   undergoing lobectomy was significantly better than sublobar resection, with an OR of 0.389 (0.263–0.578).

32   We constructed a nomogram with a C-index of 0.750 (95% confidence interval) based on the results of

33   multivariate analysis. The calibration curves based on nomogram scores indicated that the nomogram could

34   accurately predict the prognosis of patients.

35   **Conclusions**: BSC had unique clinical and prognostic features. T stage, N stage, M stage, age, and surgery

36   were independently associated with overall survival (OS). Lobectomy was a relative ideal choice for patients

37   with BSC. The nomogram effectively predicted the OS at 1-, 3-, and 5-years.

38

39

40

41   **Introduction**

42   Lung cancer is the most commonly diagnosed cancer worldwide, as well as the leading cause of cancer deaths.

43   Worldwide in 2018, it accounted for 2.1 million new cases and 1.8 million deaths. Squamous cell carcinoma is

44   one of the major well-studied histological subtypes of lung cancer [1, 2]. However, basaloid squamous cell

45   carcinoma of the lung (BSC), as a rare subtype of lung squamous cell carcinoma, is less studied, and the

46   clinical features and prognostic factors remain unclear.

47   BSC is an uncommon histological variant of lung cancer composed of cells exhibiting cytological and tissue

48   architectural features of both squamous cell lung carcinoma and basal cell carcinoma, while the proportion of

49   squamous cell components is less than 50% [3-5]. It is reported that BSC accounts for 3.9%–5.2% of all lung

50   squamous carcinomas and has unique clinical characteristics such as a high rate of metastasis and death,

51   according to previous research. Apart from this, BSC was once described to with overlap features of large cell

52   carcinoma (LCC) [6, 7]. Up to date, no relevant literatures have reported the difference of clinical features with

53   LCC and other non-small cell lung cancer (NSCLC). In this study, we compared clinicopathological

54   characteristics of related lung cancer subtypes in detail, then we used univariate and Cox hazards regression

55   analyses to identify risk factors affecting overall survival (OS) of BSC. We further developed a nomogram of

56   patients with BSC based on the results of survival analysis to better predict the prognoses of patients.

57   **Materials and Methods**

58   **Ethic statement**

59   We obtained permission to use data files from the public database of the Surveillance, Epidemiology, and End

60   Results (SEER) database. Thus, our research was exempted by the Ethics Committee of Suzhou Municipal

61   Hospital.

62   **Data extraction**

63   Data of all primary pure basaloid squamous cell carcinoma patients (ICD-O-3: 8083/3) between 2004 to 2015

64   were identified by the SEER*Stat software (v8.3.5, https://seer.cancer.gov/seerstat/) from the SEER database

65   (http://seer.cancer.gov/). Exclusion criteria were: 1) pathological types of non-pure-type basaloid squamous

66   cell carcinoma; 2) unknown aspects regarding differentiation, stage, and treatment methods; and 3) a history of

67   tumors in other sites (Figure 1).

68   We extracted and analyzed the data regarding patients' race, sex, age, grade, TMN stage, surgery type,

69   radiotherapy, and chemotherapy. The total stage, T stage, N stage, and M stage of all patients were manually

70   restaged according to the 8th edition of the American Joint Committee on Cancer (AJCC) lung cancer staging

71   project.

72   Chi-square tests were used for comparison of multi-class variables like race between basaloid squamous cell

73   carcinoma and other types of lung carcinomas. Rank sum tests were used for comparing two categorical

74   variables or ordered variables. The quantitative variables of age were compared by the variance analysis

75   method. In the analysis of prognostic factors for BSC, we used Kaplan-Meier analyses and log-rank tests for

76   univariate analyses, and Cox model tests for multivariate analyses. The above analyses were performed using

77   SPSS (version 25) software (SPSS, Chicago, IL, USA), all of which were two-sided tests, while $P < 0.05$ was

78   considered to be statistically significant. R language (v4.3.4) was used to generate and validate the nomogram,

79   while the main packages used were rms and Hmisc[8, 9].

80   **Results**

81   **Comparison of clinical features between BSC and squamous cell carcinoma (SCC)**

82   After screening, we enrolled 425 patients with BSC of pure type and 90006, 6997 and 160638 patients with

83   SCC, LCC and LAC, respectively.   Survival curves indicated that the survival of BSC patients was

84   significantly better than those of SCC, LCC and LACpatients (Figure 2 and Figure 5 of supplemental

85   materials). As shown in Table 1, 257 males and 168 females were included in the BSC group, with a median

86   age of 70.15 years and interquartile range (IQR) of 59.87–80.43 years. In total, 358 (84.2%) were white. The

87   grades included 0.5% grade I (well differentiated), 10.8% grade II (moderately differentiated), 64.5% grade III

88   (poorly differentiated), and 3.0% grade IV (undifferentiated). Most BSC patients (42.1%) were AJCC stage I,

89   14.8% were stage II, 23.8% were stage III, and 19.3% were stage IV. BSC patients had significantly less well

90   differentiated tumors (P<0.001), less N+ disease (P<0.001), fewer distant metastases (P<0.001), lower

91   proportion of radiotherapy (P<0.001) and chemotherapy (P<0.001), but a higher percentage of radical surgical

92   resection (P<0.001) than those of SCC and LAC patients. Conversely, only LCC patients had more

93   undifferentiated tumors (P<0.001), while much lower proportion of surgery (P<0.001), radiotherapy (P<0.001)

94   and chemotherapy (P<0.001) (Table 1).

95   **Analyses of BSC prognostic factors**

96   We used univariate analyses to investigate possible prognostic factors in patients with BSC. As shown in Table

97   1, there was a statistically significant correlation between age (P<0.001), grade (P<0.001), total stage

98   (P<0.001), T stage (P<0.001), N stage (P<0.001), M stage (P<0.001), surgery (P<0.001), radiotherapy

99   (P<0.001) and chemotherapy (P=0.013) with prognoses of BSC. In other words, elder age, lower

100   differentiation, and a higher total stage meant the worse prognosis. In addition, as shown in Figure 3A,

101   univariate analyses showed that patients who underwent surgery had better prognoses than patients who did

102   not, and were similar to patients receiving other surgical treatments (P<0.001). Particularly, the patients with

103   Stage I to Stage IV who underwent lobectomy had better benefits than those undergoing sublobar resection

104   (Figure 3B), including for total stage (P<0.0001), Stage I (P=0.00016), Stage II (P=0.0012), Stage III, and IV

105   (both P=0.03) (Figure 3C, 3D, 3E). Sex (P=0.257) and race (P=0.077) for the prognosis of BSC were not

106   statistically significant (Table 2).

107   The data revealed that the factors of age (P<0.001), T stage (P<0.001), N stage (P=0.009), M stage (P<0.001),

108   and surgery (P<0.001) with statistical significance using univariate analysis were found to be independent

109   factors according to multivariate analyses (Table 2). Multivariate analyses revealed that the older patient and

110   the higher the TMN stage, the worse the prognosis. Excluding the significant effects of T3 (P=0.003), T4

111   (P<0.001) and N2 (P=0.001), the remaining T stage and N stage of the prognoses of patients were similar.

112   Compared with T1 and N0, the odds ratios (ORs) were T2: 1.240 (0.856–1.797), T3: 1.935 (1.257–2.980), T4:

113   2.364 (1.559–3.585), and N1: 1.623(1.018–2.586), N2: 1.905(1.300–2.792), and N3: 1.826(0.951–3.505),

114   respectively. The prognoses of patients undergoing lobectomy were significantly better than other surgical

115   styles, with an OR of 0.389 (0.263–0.578) compared with nonsurgical or other surgical patients. Unlike the

116   results of univariate analysis, multivariate analysis showed that radiotherapy and chemotherapy were not

117   independent prognostic factors.

118   **Production and inspection of the nomogram**

119   We successfully constructed a nomogram based on the above independent predictors of patient outcomes

120   (Figure 4A). According to the patients' age, T stage, N stage, M stage, and surgery, we visually calculated the

121   patient's 1- (Figure 4B), 3- (Figure 4C) and 5-year (Figure 4D) survival probabilities. The C-index of this

122   nomogram was 0.750 as determined by the discriminant test. The consistency test showed that the 3-year and

123   5-year survival rates predicted by the nomogram were in good agreement with the actual 3-year and 5-year

124   survival rates, and the slope of the consistency curve was close to 1.

125   **Discussion**

126   We conducted an in-depth analysis of BSC using the patients' data from the SEER database and we found that

127 there were significant statistical differences with SCC, LCC and LAC in terms of race, grade, total stage, T

128 stage, N stage, M stage, surgery, radiotherapy and chemotherapy. We also found that age, T stage, N stage, M

129 stage and surgery were independent influencing factors for the prognosis of patients with BSC. We then

130 plotted a nomogram. Consistency detection proved that the nomogram effectively predicted the 1-, 3- and 5-

131 year survival probabilities of patients, while the nomogram scores effectively discriminated the patients'

132 survivals.

133 BSC is an invasive subtype of squamous cell carcinoma that can be detected in the proximal bronchi [4, 10].

134 Unlike other previous studies [11], we verified that the prognosis of BSC was better than SCC, LCC and LAC.

135 In this study, we found that the prognosis of BSC in a population-based cohort was better than SCC, LCC, and

136 LAC. However, there are some previous studies reported opposite results to ours [12, 13], perhaps due to that the

137 number of cases varied, and most of other studies focused on the patients with surgery. Meanwhile, BSC has a

138 significant lower TNM stage than other lung cancer according to our results.

139 In this population-based study, BSC and other types of lung cancers had similarities in terms of age. But Moro-

140 Sibilot D et al reported that BSC patients are older than non-BSC patients[13]. Thus, Wang LC et al

141 demonstrated that there was no significant statistical difference of mean age between BSC (58.6 years) and

142 poorly differentiated squamous cell carcinoma (60.5 years)(P=0.363)[4].There were more patients with poorly

143 differentiated BSCs, while the numbers of patients of N$^+$ and M$^+$ were less than those with SCC, LCC and

144 LAC. In our study, the 5-year survival rate of the BSC patients was close to 17.6%. In other reports, the 5-year

145 survival rate for BSC of stage I and stage II was less than 15%, much lower than the 5-year survival rate of

146 47% for resectable poorly differentiated SCC [14]. However, Kim et al. reported that there was no significant

147 difference in the median survival rate between BSC and SCC in patients with stage I, without lymph node

148   metastasis [10]. Moreover, Moro et al. reported that operative modes had no difference between the prognosis of

149   BSC and poorly-differentiated SCC [13]. As shown in the Figure 6 of supplemental materials, we compared the

150   differences between the two groups by utilizing the survival curve. It clearly indicated that poorly-

151   differentiated BSCs had better 5-year prognosis than poorly-differentiated SCCs, which were similar to the

152   overall comparison results of this research. Wang LC et al also revealed that BSC and poorly differentiated

153   squamous cell carcinomas had very similar clinical features, and there were no significant differences in

154   survival rates[4], while in our results the survival of poorly differentiated BSC was superior to that of SCC with

155   the same differentiation. More research should be carried out to validate the results.

156   Currently, surgery is the best curative treatment in stage I, stage II, and some stage III non-small lung cancers

157   [15]. Thus, lobectomy is still recommended as a preferred treatment for BSC, while more patients with

158   peripheral tumors have undergone sublobar section [16]. However, both in our univariate analyses and

159   multivariate analyses, patients with lobectomy had a better prognosis than patients undergoing other therapies.

160   Our results also suggested that at any stage, even stage III and IV, the prognosis of patients with lobectomy

161   was significantly better than those with sublobar section. This may due to the radical lobectomy that reduces

162   the potential risks for relapse and distant metastases of solid tumors [17]. In addition, survival following

163   sublobar section was inferior to lobectomy for stage I non-small cell lung cancer [18]. Therefore, further studies

164   with larger cohorts, between lobectomy and sublobar section, especially when classified by histology, should

165   be performed.

166   Nomogram, as an easily available and measurable tool of statistical prediction, which provides prognostic

167   probability of specific outcomes[19, 20]. So far, multiple nomograms have been constructed for predicting

168 prognosis of different types of lung cancers[21-23]. Thus, it has even been considered more applied than the

169 traditional AJCC TNM staging system in diverse malignancies according to great quantity of previous

170 evidence[24, 25]. Furthermore, nomograms are especially advisable to deal with individual patients without

171 existing definite clinical guidelines. In general, it seems simple and convenient via utilizing nomograms to

172 predict patients' long-time survival according to their own characteristic.

173 The latest National Comprehensive Cancer Network recommends that *EGFR* mutations and other gene

174 mutations should be considered as markers for lung squamous cell carcinoma, especially for non-smokers,

175 small biopsy, or mixed squamous cell carcinoma [26, 27]. Although the gene mutation status has not been well

176 investigated in BSC, molecularly targeted treatment may still have great potential to be used in the treatment

177 for BSC.

178 The SEER database is a population-based tumor epidemiology database in the United States, covering about

179 28% of the population, including thousands of cases of lung cancers since 1973, therefore the SEER database

180 is of great help in the study of lung cancer and other tumors [28, 29]. By analyzing the cases in the entire

181 population of the SEER database, it is possible to effectively avoid the bias of the patients from the research

182 given by a single institution. Nevertheless, there is often a lack of imaging data, smoking history, gene

183 mutations, tumor markers, and data regarding other detailed treatments, especially chemotherapy regimens in

184 the SEER database. Therefore, the impact of these factors on the prognoses of patients with BSC was not

185 included in our study. These factors may significantly affect the prognoses of the patients.

186 In our study, we have selected BSC cases that met the requirements as much as possible. But there was still a

187 significant gap with the number of SCC. Though there seemed to be some controversy, it was still determined

188 by its specific characteristics. We should further pay close attention to the future prognosis of BSCs. We

189 acknowledge that the article limited the findings to epidemiological analysis and did not set more emphasis on

190 exploring the biology of rare tumors such as molecular mechanism for gene therapy strategy.

191 **Conclusion**

192 BSC has unique clinical and prognostic features that differ from SCC, LCC and LAC. Age, T stage, N stage,

193 M stage and surgery were found to be independent predictors of prognoses in patients with BSC. The

194 nomogram we constructed better predicted the patients' 1-, 3-, and 5-year survival probabilities.

195

References:

197 [1]. Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Niksic, M., Bonaventure, A., Valkov,
198 M., Johnson, C.J., Esteve, J., Ogunbiyi, O.J., Azevedo, E.S.G., Chen, W.Q., Eser, S., Engholm, G., Stiller,
199 C.A., Monnereau, A., Woods, R.R., Visser, O., Lim, G.H., Aitken, J., Weir, H.K., and Coleman, M.P. 2018.
200 Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37
201 513 025 patients diagnosed with one of 18 cancers from   322 population-based registries in 71 countries.
202 LANCET 391:1023-1075. 10.1016/S0140-6736(17)33326-3
203 [2]. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., and Jemal, A. 2018. Global cancer
204 statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.
205 CA Cancer J Clin 68:394-424. 10.3322/caac.21492
206 [3]. Brambilla, E., Travis, W.D., Colby, T.V., Corrin, B., and Shimosato, Y. 2001. The new World Health
207 Organization classification of lung tumours. EUROPEAN RESPIRATORY JOURNAL 18:1059-1068.
208 [4]. Wang, L.C., Wang, L., Kwauk, S., Woo, J.A., Wu, L.Q., Zhu, H., Zhan, L.Z., Sun, N.L., and Zhang, L.
209 2011. Analysis on the clinical features of 22 basaloid squamous cell carcinoma of the lung. Journal of
210 Cardiothoracic Surgery 6:10. 10.1186/1749-8090-6-10
211 [5]. Crapanzano, J.P., Loukeris, K., Borczuk, A.C., and Saqi, A. 2011. Cytological, histological, and
212 immunohistochemical findings of pulmonary carcinomas with basaloid features. DIAGNOSTIC
213 CYTOPATHOLOGY 39:92-100. 10.1002/dc.21335
214 [6]. Travis, W.D., Brambilla, E., Nicholson, A.G., Yatabe, Y., Austin, J., Beasley, M.B., Chirieac, L.R., Dacic,
215 S., Duhig, E., Flieder, D.B., Geisinger, K., Hirsch, F.R., Ishikawa, Y., Kerr, K.M., Noguchi, M., Pelosi, G.,
216 Powell, C.A., Tsao, M.S., and Wistuba, I. 2015. The 2015 World Health Organization Classification of Lung
217 Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. Journal of
218 Thoracic Oncology 10:1243-1260. 10.1097/JTO.0000000000000630
219 [7].Vignaud, J.M. 2016. [Squamous cell carcinoma, basaloid squamous cell carcinoma and adenosquamous
220 carcinoma in the lung]. ANNALES DE PATHOLOGIE 36:15-23. 10.1016/j.annpat.2015.11.011

221 [8]. Sun, F., Ma, K., Yang, X., Li, M., Shi, Y., Zhan, C., Jiang, W., and Wang, Q. 2017. A nomogram to
222 predict prognosis after surgery in early stage non-small cell lung   cancer in elderly patients. International
223 Journal of Surgery 42:11-16. 10.1016/j.ijsu.2017.04.024

224 [9]. Wang, S., Ma, K., Chen, Z., Yang, X., Sun, F., Jin, Y., Shi, Y., Jiang, W., Wang, Q., and Zhan, C. 2018. A
225 Nomogram to Predict Prognosis in Malignant Pleural Mesothelioma. WORLD JOURNAL OF SURGERY
226 42:2134-2142. 10.1007/s00268-017-4424-6

227 [10]. Kim, D.J., Kim, K.D., Shin, D.H., Ro, J.Y., and Chung, K.Y. 2003. Basaloid carcinoma of the lung: a
228 really dismal histologic variant? ANNALS OF THORACIC SURGERY 76:1833-1837.

229 [11]. Brambilla, E., Moro, D., Veale, D., Brichon, P.Y., Stoebner, P., Paramelle, B., and Brambilla, C. 1992.
230 Basal cell (basaloid) carcinoma of the lung: a new morphologic and phenotypic entity with separate prognostic
231 significance. HUMAN PATHOLOGY 23:993-1003.

232 [12]. Brambilla, C., Laffaire, J., Lantuejoul, S., Moro-Sibilot, D., Mignotte, H., Arbib, F., Toffart, A.C., Petel,
233 F., Hainaut, P., Rousseaux, S., Khochbin, S., de Reynies, A., and Brambilla, E. 2014. Lung squamous cell
234 carcinomas with basaloid histology represent a specific molecular entity. CLINICAL CANCER RESEARCH
235 20:5777-5786. 10.1158/1078-0432.CCR-14-0459

236 [13]. Moro-Sibilot, D., Lantuejoul, S., Diab, S., Moulai, N., Aubert, A., Timsit, J.F., Brambilla, C., Brichon,
237 P.Y., and Brambilla, E. 2008. Lung carcinomas with a basaloid pattern: a study of 90 cases focusing on their
238 poor prognosis. EUROPEAN RESPIRATORY JOURNAL 31:854-859. 10.1183/09031936.00058507

239 [14]. Moro, D., Brichon, P.Y., Brambilla, E., Veale, D., Labat, F., and Brambilla, C. 1994. Basaloid bronchial
240 carcinoma. A histologic group with a poor prognosis. CANCER 73:2734-2739.

241 [15]. Lang-Lazdunski, L. 2013. Surgery for nonsmall cell lung cancer. Eur Respir Rev 22:382-404.
242 10.1183/09059180.00003913

243 [16]. Zhang, H., and Shen-Tu, Y. 2015. [Clinical Advance of Sublobectomy for Early Stage Non-small Cell
244 Lung Cancer]. Zhongguo Fei Ai Za Zhi 18:565-570. 10.3779/j.issn.1009-3419.2015.09.07

245 [17]. Wang, J., and Zhao, H. 2016. [Issues Need to be Considered in Sublobectomy for Early Stage Lung
246 Cancer]. Zhongguo Fei Ai Za Zhi 19:351-354. 10.3779/j.issn.1009-3419.2016.06.10

247 [18]. Zhang, Y., Sun, Y., Wang, R., Ye, T., Zhang, Y., and Chen, H. 2015. Meta-analysis of lobectomy,
248 segmentectomy, and wedge resection for stage I non-small cell lung cancer. JOURNAL OF SURGICAL
249 ONCOLOGY 111:334-340. 10.1002/jso.23800

250 [19]. Kent, M.S., Mandrekar, S.J., Landreneau, R., Nichols, F., Foster, N.R., DiPetrillo, T.A., Meyers, B.,
251 Heron, D.E., Jones, D.R., Tan, A.D., Starnes, S., Putnam, J.J., and Fernando, H.C. 2016. A Nomogram to
252 Predict Recurrence and Survival of High-Risk Patients Undergoing Sublobar Resection for Lung Cancer: An
253 Analysis of a Multicenter Prospective Study (ACOSOG Z4032). ANNALS OF THORACIC SURGERY
254 102:239-246. 10.1016/j.athoracsur.2016.01.063

255 [20]. Zhang, J., Gold, K.A., Lin, H.Y., Swisher, S.G., Xing, Y., Lee, J.J., Kim, E.S., and William, W.J. 2015.
256 Relationship between tumor size and survival in non-small-cell lung cancer (NSCLC): an analysis of the
257 surveillance, epidemiology, and end results (SEER) registry. Journal of Thoracic Oncology 10:682-690.
258 10.1097/JTO.0000000000000456

259 [21]. Zhang, B., Yuan, Z., Zhao, L., Pang, Q., and Wang, P. 2017. Nomograms for predicting progression and
260 efficacy of post-operation radiotherapy   in IIIA-pN2 non-small cell lung cancer patients. Oncotarget 8:37208-
261 37216. 10.18632/oncotarget.16564

262 [22]. Young, K.A., Efiong, E., Dove, J.T., Blansfield, J.A., Hunsinger, M.A., Wild, J.L., Shabahang, M.M.,
263 and Facktor, M.A. 2017. External Validation of a Survival Nomogram for Non-Small Cell Lung Cancer Using
264 the National Cancer Database. ANNALS OF SURGICAL ONCOLOGY 24:1459-1464. 10.1245/s10434-017-
265 5795-5
266 [23]. Ye, L., Shi, S., Zeng, Z., Huang, Y., Hu, Y., and He, J. 2018. Nomograms for predicting disease
267 progression in patients of Stage I non-small cell lung cancer treated with stereotactic body radiotherapy.
268 JAPANESE JOURNAL OF CLINICAL ONCOLOGY 48:160-166. 10.1093/jjco/hyx179
269 [24]. Liang, W., Zhang, L., Jiang, G., Wang, Q., Liu, L., Liu, D., Wang, Z., Zhu, Z., Deng, Q., Xiong, X., Shao,
270 W., Shi, X., and He, J. 2015. Development and validation of a nomogram for predicting survival in patients
271 with resected non-small-cell lung cancer. JOURNAL OF CLINICAL ONCOLOGY 33:861-869.
272 10.1200/JCO.2014.56.6661
273 [25]. Xie, D., Marks, R., Zhang, M., Jiang, G., Jatoi, A., Garces, Y.I., Mansfield, A., Molina, J., and Yang, P.
274 2015. Nomograms Predict Overall Survival for Patients with Small-Cell Lung Cancer Incorporating
275 Pretreatment Peripheral Blood Markers. Journal of Thoracic Oncology 10:1213-1220.
276 10.1097/JTO.0000000000000585
277 [26]. Keedy, V.L., Temin, S., Somerfield, M.R., Beasley, M.B., Johnson, D.H., McShane, L.M., Milton, D.T.,
278 Strawn, J.R., Wakelee, H.A., and Giaccone, G. 2011. American Society of Clinical Oncology provisional
279 clinical opinion: epidermal growth factor receptor (EGFR) Mutation testing for patients with advanced non-
280 small-cell lung cancer considering first-line EGFR tyrosine kinase inhibitor   therapy. JOURNAL OF
281 CLINICAL ONCOLOGY 29:2121-2127. 10.1200/JCO.2010.31.8923
282 [27]. Felip, E., Gridelli, C., Baas, P., Rosell, R., and Stahel, R. 2011. Metastatic non-small-cell lung cancer:
283 consensus on pathology and molecular tests, first-line, second-line, and third-line therapy: 1st ESMO
284 Consensus Conference in Lung Cancer; Lugano 2010. ANNALS OF ONCOLOGY 22:1507-1519.
285 10.1093/annonc/mdr150
286 [28]. Yang, X., Sun, F., Chen, L., Shi, M., Shi, Y., Lin, Z., Feng, M., Zhan, C., Jiang, W., and Wang, Q. 2017.
287 Prognostic value of visceral pleural invasion in non-small cell lung cancer: A propensity score matching study
288 based on the SEER registry. JOURNAL OF SURGICAL ONCOLOGY 116:398-406. 10.1002/jso.24677
289 [29]. Yang, X., Zhan, C., Li, M., Huang, Y., Zhao, M., Yang, X., Lin, Z., Shi, Y., Jiang, W., and Wang, Q.
290 2018. Lobectomy Versus Sublobectomy in Metachronous Second Primary Lung Cancer: A Propensity Score
291 Study. ANNALS OF THORACIC SURGERY 106:880-887. 10.1016/j.athoracsur.2018.04.071

292

293

294

295

296

297

298

299

300

301

302

303

304

305

Figure legends

Figure 1 The flow diagram of the selection process for the study cohort.

Figure 2 The survival curve of basaloid squamous cell carcinoma of the lung (BSC), lung squamous cell carcinoma (SCC), large cell carcinoma (LCC) and lung adenocarcinoma (LAC).

Figure 3 (A) Survival analyses for patients with different surgeries and without surgery. (B) Survival analyses for patients with lobectomy and sublobar resection classified by tumor histology and tumor stage. (C) Stage I (1-, 3-, 5-year survival rate of lobectomy vs. sublobar section: 82.0%, 53.2%, 31.5% vs. 70.7%, 46.3%, 17.1%). (D) Stage II (1-, 3-, 5-year survival rate of lobectomy vs. sublobar section : 71.1%, 40.0%, 31.1% vs. 75.0%, 12.5%, 12.5%). (E) Stage III and IV (1-, 3-, 5-year survival rate of lobectomy vs. sublobar section: 67.4%, 32.6%, 18.6% vs. 46.2%, 7.7%, 0.0%).

Figure 4 (A) Nomogram for prediction of 1-, 3- and 5- year overall survival of patients with basaloid squamous cell carcinoma of the lung (BSC). Calibration plots of the nomogram prediction of (B) 1-year, (C) 3-year and (D) 5-year overall survival nomogram calibration curves. The red line represents equality of the observed and predicted probability.

Supplementary material

Figure 5 (A) The survival curve of basaloid squamous cell carcinoma of the lung (BSC) and lung squamous

322   cell carcinoma (SCC). (B) The survival curve of basaloid squamous cell carcinoma of the lung (BSC) and

323   large cell carcinoma (LCC). (C) The survival curve of basaloid squamous cell carcinoma of the lung (BSC)

324   and lung adenocarcinoma (LAC).

325   **Figure 6** The survival curve of poor differentiated basaloid squamous cell carcinoma of the lung (PBSC) and

326   poor differentiated lung squamous cell carcinoma (PSCC).

327

328

Comparison of the clinipathological characteristics of basaloid squamous cell carcinoma (BSC), squamous cell carcinoma (SCC), large cell carcinoma (LCC) and lung adenocarcinoma (LAC).

*P value for chi-square test. BSC, Basaloid squamous cell carcinoma; SCC, squamous cell carcinoma; LCC, large cell carcinoma ; LAC, lung adenocarcinoma; IQR, interquartile range.

1   **Table 1**. Comparison of the clinipathological characteristics of basaloid squamous cell carcinoma (BSC), squamous cell carcinoma of the lung (SCC),
2   large cell carcinoma of the lung (LCC) and lung adenocarcinoma (LAC).

Characteristics	BSC	SCC	p value	LCC	p value	LAC	p value
Race, n (%)			0.351		0.039		0.049
White	358(84.2)	75272(83.6)		5682(81.2)		128685(80.1)	
Black	43(10.1)	10617(11.8)		999(14.3)		18644(11.6)	
Other	24(5.7)	4117(4.6)		316(4.5)		13309(8.3)	
Age, median [IQR]	70.15 (59.87-80.43)	70.41 (60.66-80.16)	0.139	67.92 (56.87-78.97)	0.303	68.37 (57.27-79.47)	0.376
Sex, n (%)			0.503		0.443		<0.001
Male	257(60.5)	55849(62.1)		4099(58.6)		78527(48.9)	
Female	168 (39.5)	34157(37.9)		2898(41.4)		82111(51.1)	
Grade, n (%)			<0.001		<0.001		<0.001
Well differetiated	2(0.5)	1974(2.2)		15(0.2)		11955(7.4)	
Moderately differetiated	46(10.8)	26606(29.6)		80(1.1)		36005(22.4)	
Poorly differetiated	274(64.5)	32307(35.9)		2115(30.2)		43291(27.0)	
Undifferetiated	13(3.0)	648(0.7)		2299(32.9)		968(0.6)	
Unknown	90(21.2)	28471(31.6)		2488(35.6)		68419(42.6)	
Total stage,n (%)			<0.001		<0.001		<0.001
I	179(42.1)	27683(30.8)		1460(20.9)		41408(25.8)	
II	63(14.8)	10946(12.2)		534(7.6)		10560(6.6)	
III	101(23.8)	26297(29.2)		1858(26.6)		32528(20.2)	
IV	82(19.3)	25080(27.8)		3145(44.9)		76142(47.4)	
T stage, n (%)			<0.001		<0.001		0.001
T1	133(31.3)	9151(10.2)		1292(10.2)		43558(27.1)	
T2	139(32.7)	66275(73.6)		2346(73.6)		47776(29.8)	
T3	68(16.0)	3596(4.0)		880(4.0)		22072(13.7)	
T4	85(20.0)	10984(12.2)		2479(12.2)		47232(29.4)	

N stage, n (%)			<0.001		<0.001		<0.001
N0	269(63.3)	9151(10.2)		2847(40.7)		72104(44.9)	
N1	50(11.8)	66275(73.6)		663(9.5)		14110(8.8)	
N2	84(19.8)	3596(4.0)		2628(37.5)		54487(33.9)	
N3	22(5.1)	10984(12.2)		859(12.3)		19937(12.4)	
M stage,n (%)			<0.001		<0.001		<0.001
M0	343(80.7)	58972(65.5)		3854(55.1)		84496(52.6)	
M1	82(19.3)	31034(34.5)		3143(44.9)		76142(47.4)	
Surgery, n (%)			<0.001		<0.001		<0.001
Not performed	144(33.9)	60336(67.0)		5041(72.0)		112857(70.3)	
Lobectomy	199(46.8)	20978(23.3)		1377(19.7)		35329(22.0)	
Sublobar resection	62(14.6)	6153(6.8)		438(6.3)		10982(6.8)	
Pneumonectomy	20(4.7)	2539(2.8)		141(2.0)		1470(0.9)	
Radiotherapy , n (%)			<0.001		<0.001		<0.001
No	104(24.5)	4(0.0)		3980(56.9)		98823(61.5)	
Yes	321(75.5)	90002 (100.0)		3017(43.1)		61815(38.5)	
Chemotherapy, n (%)			<0.001		<0.001		<0.001
No	134(31.5)	53403(59.3)		3944(56.4)		88220(54.9)	
Yes	291(68.5)	36603(40.7)		3053(43.6)		72418(45.1)	

5 BSC, Basaloid squamous cell carcinoma; SCC, squamous cell carcinoma; LCC, large cell carcinoma of the lung; LAC, lung adenocarcinoma; IQR,

6 interquartile range.

Univariate and multivariate Cox proportional hazards analysis of BSC.

*P value for chi-square test. C-index=0.750.

1   **Table 2**. Univariate and multivariate Cox proportional hazards analysis

Patient characteristics	Univariate analysis		Multivariate analysis	
	HR (95%CI)	p value	HR (95%CI)	p value
Race		0.077		-
White	Reference		-	-
Black	1.251(0.811-1.931)	0.311	-	-
Other	0.502(0.248-1.019)	0.056	-	-
Age		<0.001	1.032(1.017-1.048)	<0.001
Sex		0.257		-
Male	Reference		-	-
Female	0.855(0.652-1.121)		-	-
Grade		<0.001		0.248
Well differetiated	Reference		Reference	
Moderately differetiated	0.776(0.181-3.322)	0.733	2.179(0.404-11.756)	0.365
Poorly differetiated	0.994(0.245-4.033)	0.993	1.894(0.369-9.712)	0.444
Undifferetiated	1.748(0.375-8.139)	0.477	4.198(0.689-25.591)	0.120
Unknown	2.291(0.537-9.421)	0.251	2.101(0.401-10.988)	0.379
AJCC 8th T stage		<0.001		<0.001
T1	Reference		Reference	
T2	1.052(0.735-1.507)	0.780	1.240(0.856-1.797)	0.256
T3	1.909(1.274-2.862)	0.002	1.935(1.257-2.980)	0.003
T4	2.689(1.879-3.849)	<0.001	2.364(1.559-3.585)	<0.001
AJCC 8th N stage		<0.001		0.009
N0	Reference		Reference	
N1	1.678(1.088-2.587)	0.019	1.623(1.018-2.586)	0.042
N2	3.218(2.360-4.390)	<0.001	1.905(1.300-2.792)	0.001
N3	5.013(2.855-8.804)	<0.001	1.826(0.951-3.505)	0.071
AJCC 8th M stage		<0.001		<0.001

M0	Reference		Reference	
M1	4.253(3.170-5.704)	<0.001	2.399(1.695-3.393)	<0.001
Surgery		<0.001		<0.001
Not performed	Reference		Reference	
Lobectomy	0.187(0.136-0.258)	<0.001	0.389(0.263-0.578)	<0.001
Sublobar resection	0.442(0.304-0.643)	<0.001	0.889(0.567-1.394)	0.608
Pneumonectomy	0.384(0.209-0.706)	<0.001	0.614(0.317-1.189)	0.148
Radiotherapy		<0.001		
No	Reference		Reference	
Yes	2.173(1.631-2.895)	<0.001	1.076(0.768-1.508)	0.669
Chemotherapy				
No	Reference		Reference	
Yes	1.423(1.076-1.882)	0.013	0.725(0.506-1.039)	0.080

3   The total stage was not an independent variable related to T, N, and M stage, therefore it was excluded in the multivariate analysis.

Figure 1

The flow diagram of the selection process for the study cohort.

The flow diagram of the selection process for the study cohort.

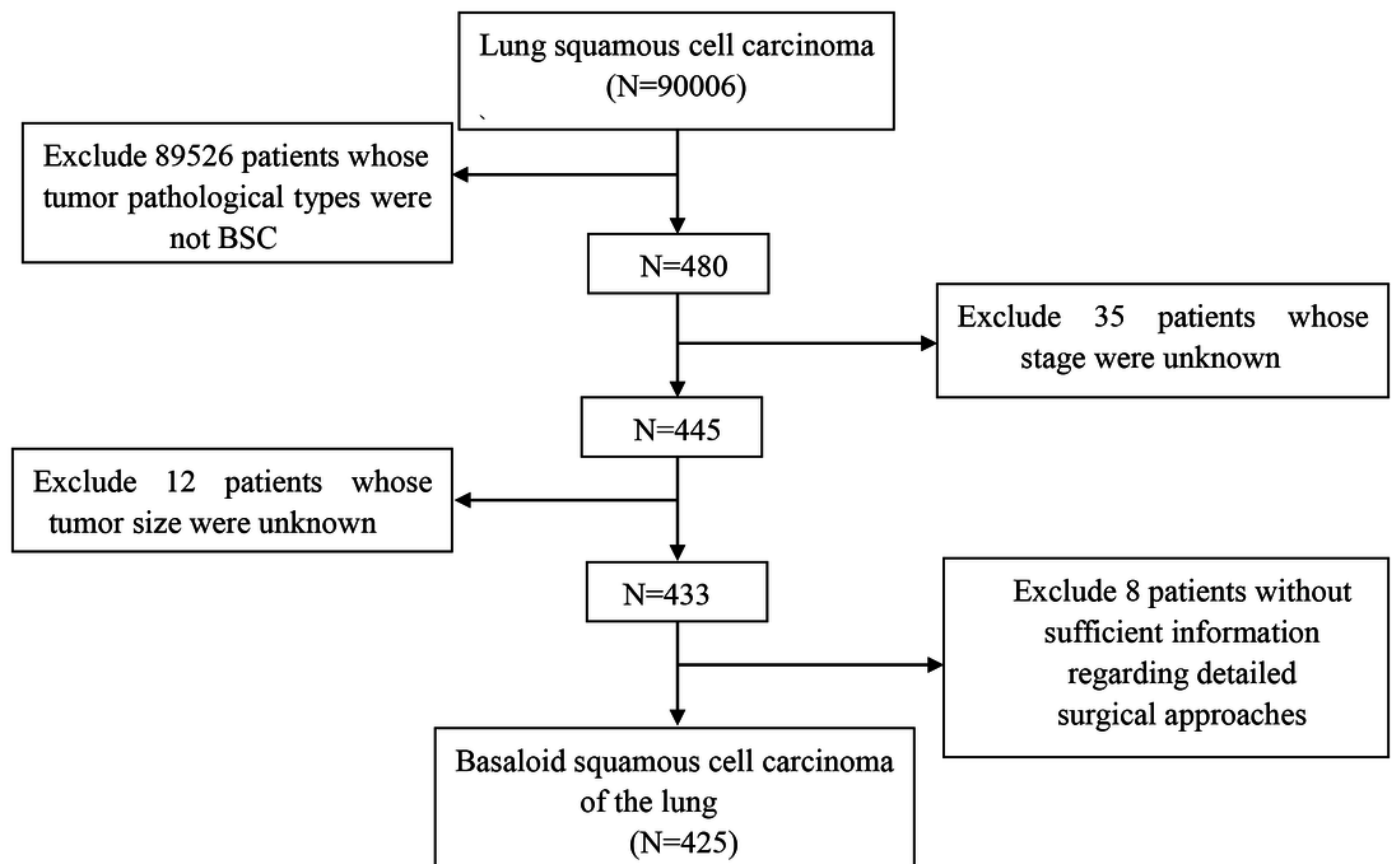



Figure 2

Survival for BSC、 SCC、 LCC and LAC

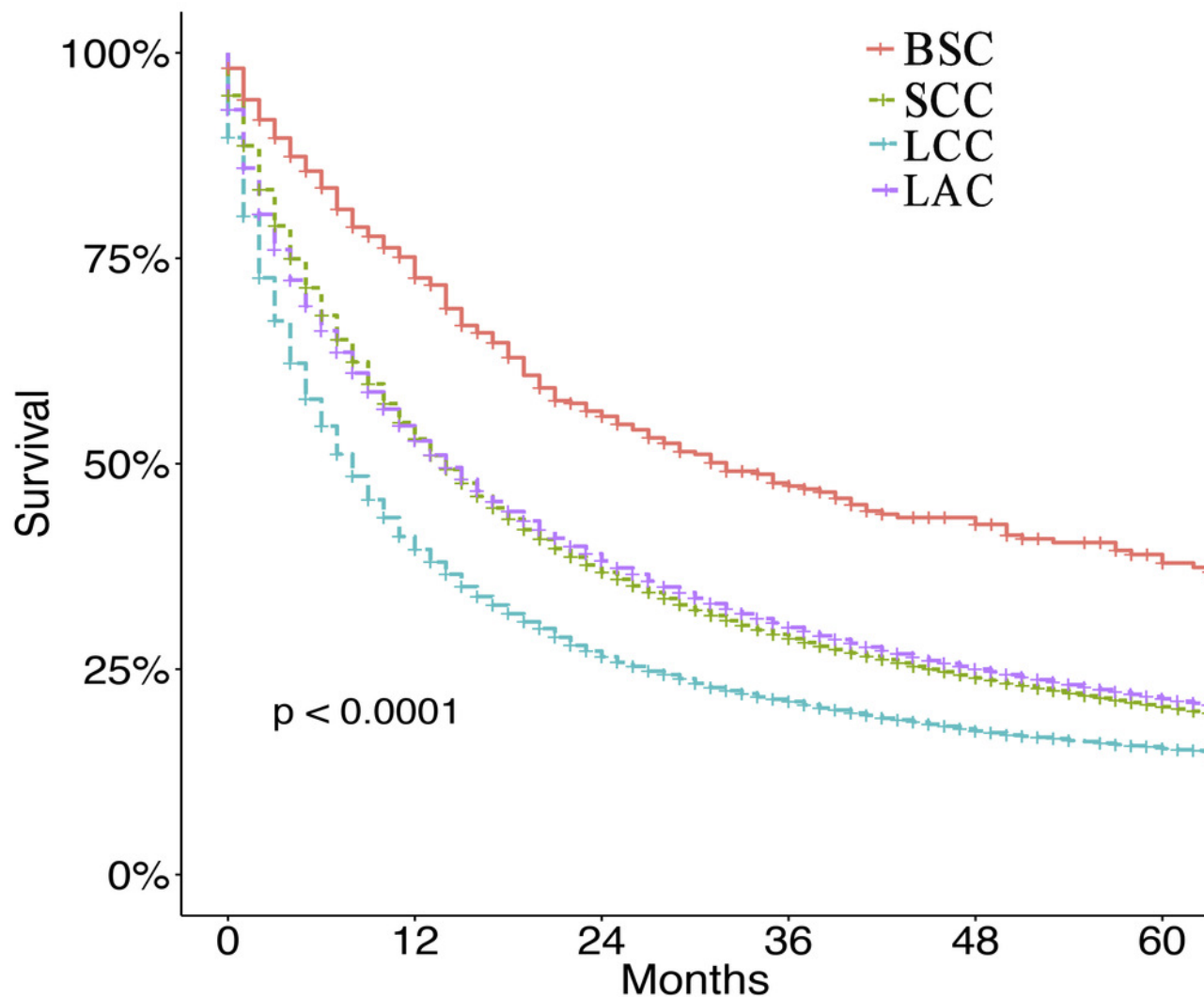
The survival curves of basaloid squamous cell carcinoma of the lung (BSC) 、 squamous cell carcinoma (SCC) 、 large cell carcinoma (LCC) and lung adenocarcinoma (LAC) .

Patients at risk						
BSC	425	266	175	130	104	75
SCC	90006	44210	26900	18399	13298	9658
LCC	6997	2796	1789	1351	1071	885
LAC	60482	77177	48437	32991	23365	16747

Figure 3

survival for surgical modalities

(A) Survival analyses for patients with different surgeries and without surgery. (B) Survival analyses for patients with lobectomy and sublobar resection classified by tumor histology and tumor stage. (C) Stage I (1-, 3-, 5-year survival rate of lobectomy vs. sublobar section: 82.0%, 53.2%, 31.5% vs. 70.7%, 46.3%, 17.1%). (D) Stage II (1-, 3-, 5-year survival rate of lobectomy vs. sublobar section : 71.1%, 40.0%, 31.1% vs. 75.0%, 12.5%, 12.5%). (E) Stage III and IV (1-, 3-, 5-year survival rate of lobectomy vs. sublobar section: 67.4%, 32.6%, 18.6% vs. 46.2%, 7.7%, 0.0%).

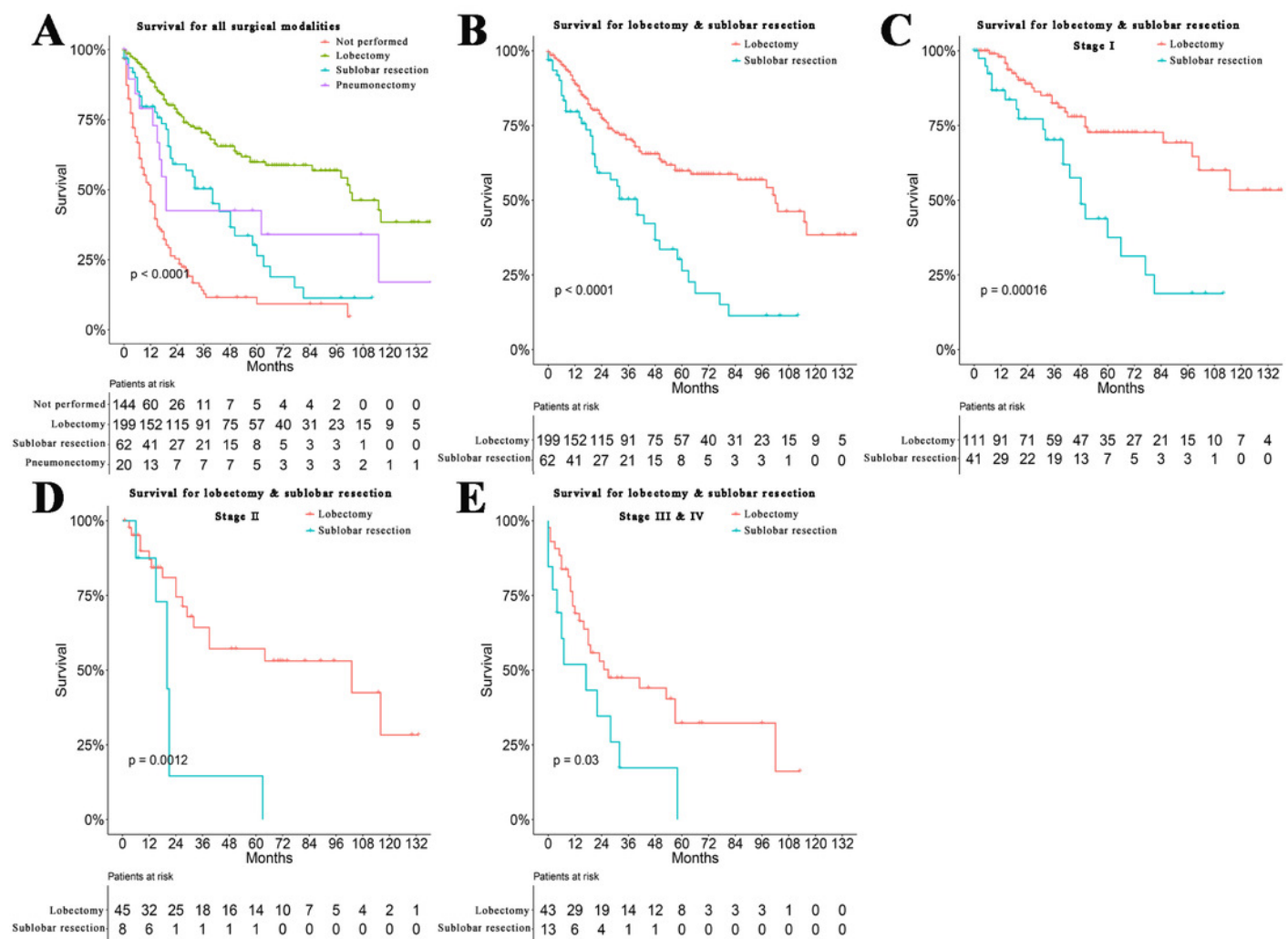



Figure 4

Nomogram and overall survival nomogram calibration curves.

(A) Nomogram for prediction of 1-, 3- and 5- year overall survival of patients with basaloid squamous cell carcinoma of the lung (BSC). Calibration plots of the nomogram prediction of (B) 1-year,(C) 3-year and (D) 5-year overall survival nomogram calibration curves. The red line represents equality of the observed and predicted probability.