TITLE:

Validation Study of the International Association for the Study of Lung Cancer Histologic Grading System of Invasive Lung Adenocarcinoma

AUTHOR(S):

Rokutan-Kurata, Mariyo; Yoshizawa, Akihiko; Ueno, Kentaro; Nakajima, Naoki; Terada, Kazuhiro; Hamaji, Masatsugu; Sonobe, Makoto; ... Date, Hiroshi; Morita, Satoshi; Haga, Hironori

IASLC

BRIEF REPORT

Validation Study of the International Association for the Study of Lung Cancer Histologic Grading System of Invasive Lung Adenocarcinoma

Check for updates

Mariyo Rokutan-Kurata, MD,[a] Akihiko Yoshizawa, MD, PhD,[a],* Kentaro Ueno, MS,[b]
Naoki Nakajima, MD, PhD,[a] Kazuhiro Terada, MD,[a] Masatsugu Hamaji, MD, PhD,[c]
Makoto Sonobe, MD, PhD,[c,d] Toshi Menju, MD, PhD,[c] Hiroshi Date, MD, PhD,[c]
Satoshi Morita, PhD,[b] Hironori Haga, MD, PhD[a]

[a]Department of Diagnostic Pathology, Kyoto University Hospital, Kyoto, Japan
[b]Department of Biomedical Statistics and Bioinformatics, Kyoto University Graduate School of Medicine, Kyoto, Japan
[c]Department of Thoracic Surgery, Kyoto University Hospital, Kyoto, Japan
[d]Department of Thoracic Surgery, Osaka Red-Cross Hospital, Osaka, Japan

ABSTRACT

Introduction: A histologic grading system for invasive lung adenocarcinoma (ADC) has been proposed by the International Association for the Study of Lung Cancer (IASLC) Pathology Committee in June 2020. This study evaluated the prognostic value of the IASLC histologic grading system (the IASLC system) in a large Japanese cohort.

Methods: We performed comprehensive histologic subtyping using the semiquantitative estimation of five major patterns and complex glandular patterns in patients with a completely resected lung ADC and determined the histologic grade using the IASLC system. Concordance index and receiver-operating characteristic curves were used to evaluate the clinical utility of the IASLC system for recurrence and death; the comparison was performed with the architectural-pattern system (the Arch system) and the grading system on the basis of the two most predominant patterns (the Sica's system).

Results: Of 1002 patients with invasive ADC, 235 had recurrent disease and 166 died of lung cancer. The concordance index and area under the curve of the IASLC system were 0.777 and 0.807 for recurrence and 0.767 and 0.776 for death, respectively. These were similar to those of the Arch system (0.763 and 0.796 for recurrence, 0.743 and 0.755 for death) and the Sica's system (0.786 and 0.814 for recurrence, 0.762 and 0.773 for death).

Conclusions: We reported that the IASLC system for invasive lung ADC has prognostic significance by evaluating a large Japanese cohort. We believe that the IASLC grading system will provide physicians with better information for postsurgery treatment.

Introduction

According to the 2015 WHO classification, lung adenocarcinoma (ADC) is classified into the following four categories: preinvasive lesion including ADC in situ (AIS), minimally invasive ADC (MIA), invasive ADC (invADC), and variants.[1] AISs and MIAs have 100% or nearly 100% 5-year overall survival (OS) estimates.[2] In contrast, there is no internationally accepted histologic predictor for invADC category, although there have been some proposals for such a grading. The architectural

grading system, which is the most frequently referenced one, is based on the predominant growth pattern of invADC. It contains the following three grades: low (lepidic ADC), intermediate (acinar and papillary ADCs), and high (solid and micropapillary ADCs) grades.[3] The drawback of the architectural grading system is that it refers to a predominant growth pattern only and it could miss a minor high-grade pattern. For instance, lepidic predominant ADC is classified into low grade even if high-grade components, including a solid or micropapillary pattern, are identified. The Sica's grading system may be superior to the architectural grading system because it considers both the predominant and second most common growth patterns.[4] Other than the five major histologic growth patterns, new patterns, such as the complex glandular pattern (CGP) and the discohesive pattern, have been confirmed as high grade.[5–8] The International Association for the Study of Lung Cancer (IASLC) Pathology Committee has recently proposed a new histologic grading system for invADC on the basis of a combination of the most predominant pattern and any high-grade histologic pattern ($\geq$20%) (the IASLC system).[9] Nevertheless, a validation study with a large cohort of patients has not been reported. In addition, most lung ADC grading systems have been studied in predominantly non-Asian cohorts[3,4,10,11]

This study aimed to evaluate the prognostic value of the IASLC system and compare it with other grading systems by reviewing lung ADCs from patients examined in a single institution.

Materials and Methods
Patients
This retrospective analysis was conducted among 1241 patients with lung ADC who underwent complete resection with curative intent at Kyoto University Hospital between 2001 and 2016. Patients were excluded if they had multiple primary lung cancers, were treated with chemotherapy or radiotherapy before surgery, underwent incomplete resection, or had incomplete data to review (Supplementary Fig. 1). This study was approved by the institute's ethics committee (approval number R1814-1). Consent was waived for this retrospective study which analyzed pathologic specimens with limited clinical information.

Histologic Evaluation
All hematoxylin and eosin-stained tumor slides were reviewed and classified by two pathologists blinded to patient clinical outcomes (MRK and AY) on the basis of the slides or whole slide images. Tumors were classified on the basis of comprehensive histologic subtyping, and the percentage of each histologic component was

recorded in 5% increments according to the 2015 WHO classification.[1] Any nontraditional patterns, including a CGP referenced in the IASLC grading, were also estimated.[9]

Regarding the histologic grading of invADC, we evaluated the following approaches to achieve the best prognostic determination: (1) the architectural grading system (the Arch system); (2) the Sica's grading system (the Sica's system); and (3) the IASLC system.[9] The Arch system was based on the predominant histology as follows: low (lepidic ADC), intermediate (papillary or acinar ADC), or high (micropapillary or solid ADC or cribriform ADC which mainly consists of CGP) grade.[3,5–7] The original Sica's grading system was based on the sum of the two most predominant patterns.[4] Briefly, histologic scores were initially determined on the basis of the following three groups: grade 1 (lepidic), grade 2 (papillary or acinar), and grade 3 (solid or micropapillary). We added novel high-grade components to the grade 3 group at this time. Then, the two predominant grades were summed. Lastly, Sica's grade was calculated as follows: low grade (score $\leq$3), intermediate grade (score 4), or high grade (score 5 or 6). The IASLC system was defined as the most predominant pattern plus greater than or equal to 20% of any high-grade pattern as follows: low grade (lepidic subtype with <20% of high-grade pattern), intermediate grade (papillary or acinar subtype with <20% of high-grade pattern), and high grade (any predominant subtype with $\geq$20% of high-grade pattern).[9]

Statistical Analysis
Survival curves were developed using the Kaplan-Meier method. OS and disease-free survival (DFS) were compared using the Cox proportional hazards models and the stratified log-rank test adjusted for the stratification factor (pathologic stage). The Cox models contained patient clinical characteristics (age, sex, and pathologic stage) and the grading system as covariates. OS was defined as the interval from the date of resection to the date of death or censored at the last-known-alive date. DFS was similarly defined but included death and lung cancer recurrence, whichever occurred first, as events. To calculate median follow-up time, we used the reverse Kaplan-Meier method. To evaluate the performance of each grading system for DFS and OS, we used the area under the curve (AUC) of the time-dependent receiver-operating characteristic curve and concordance index (C-index) at 5 years.[12,13] To select patients who are most likely to benefit from postoperative chemotherapy, we also analyzed prognoses within the stage I cohort in addition to the entire cohort. We calculated point estimates and 95% confidence intervals (CIs) using the R packages, timeROC and survC1. We set the significance level at 5% (two-sided) and reported

two-sided *p* values. Data analysis and summary graphs were generated using the JMP statistical software package, version 13 (SAS Institute, Cary, NC) and R version 3.6.3 (R Foundation for Statistical Computing, Vienna, Austria). Lastly, to evaluate interobserver agreement, KT reviewed 100 cases, which were randomly selected in the same cohort. Weighted κ statistics were subsequently used to measure the reliability of the agreement between KT and the initial two authors.

Results

Clinicopathologic Characteristics

In total, 179 patients (14.4%) died during the follow-up period and 239 patients (19.3%) had relapse. In this cohort, the median follow-up time was 1836 (95% CI: 1806–1869) days. The number of patients at each pathologic stage was as follows: stage 0, 27 patients (2.2%); I, 996 patients (80.2%); II, 128 patients (10.3%); and III, 90 patients (7.3%).

We classified 1241 ADCs according to the 2015 WHO classification and the newly described CGP.[1,5] The most predominant subtype was papillary ADC (n = 539, 43.4%), solid ADC (n = 154, 12.4%), MIA (n = 153, 12.3%), lepidic ADC (n = 133, 10.7%), and acinar ADC (n = 97, 7.8%). The proportions of AIS, cribriform ADC (mainly consisting CGP), micropapillary ADC, invasive mucinous ADC (IMA), and the variants (four colloid ADCs, three fetal ADCs, and one enteric ADC) were less than 5%.

Survival Analyses and Comparison of the Grading Systems (C-index and AUC)

Because the proposal study by IASLC excluded AIS, MIA, and variants (including 51 IMAs) for evaluation of the grading system, these were also excluded in our study (Supplementary Fig. 1). Lastly, 1002 cases were enrolled. Of 1002 patients, 235 had recurrent disease and 166 died of lung cancer. Patients' demographic information of 1002 cases was illustrated in Table 1. We evaluated the prognostic significance of the following grading system in evaluating recurrence and death: (1) the Arch system, (2) the Sica's system, and (3) the IASLC system. Figure 1 presents the Kaplan-Meier curves stratified by the grading system (Fig. 1*A-F*). There were significant differences among DFS curves of low, intermediate, and high grades in every grading system both all stage cohort and only stage I cohort (all, *p* < 0.001). The C-index and AUC of the IASLC system were 0.777 (95% CI: 0.749–0.806) and 0.807 (95% CI: 0.771–0.843) for recurrence and 0.767 (95% CI: 0.724–0.810) and 0.776 (95% CI: 0.728–0.825) for death, respectively. Regarding only stage I cohort, the C-index and AUC of the IASLC system were 0.701 (95% CI: 0.651–0.750) and

Table 1. Patients' Demographic Information

Characteristic	n	%
Total	1002	100
Age (y)		
≤65	413	41.2
(mean 66.2 ± 9.97)		
≥66	589	58.8
Sex		
Male	491	49.0
Female	511	51.0
Smoking		
Smoker	541	54.0
Never smoker	461	46.0
Type of operation		
Lobectomy	773	77.1
Segmentectomy	173	17.3
Wedge	51	5.1
Pneumonectomy	5	0.5
Tumor size		
≤25 mm	635	63.4
(mean 23.9 mm ± 13.8)		
>25 mm	367	36.6
Lymph node metastasis		
Negative	849	84.7
Positive	153	15.3
Stage		
I	789	78.7
II	126	12.6
IIIA	87	8.6
2015 WHO classification		
Lepidic ADC	133	13.3
Acinar ADC	98	9.8
Papillary ADC	539	53.8
Solid ADC	154	15.4
Cribriform ADC[a]	42	4.2
MP ADC	36	3.6
Pleural invasion		
Positive	206	20.6
Negative	796	79.4
Lymphatic invasion		
Positive	115	11.5
Negative	887	88.5
Vascular invasion		
Positive	213	21.3
Negative	789	78.7
STAS		
Positive	390	38.9
Negative	612	61.1

[a]Cribriform ADC mainly revealed the complex glandular pattern.
ADC, adenocarcinoma; MP, micropapillary; STAS, spread through air space.

0.722 (95% CI: 0.668–0.776) for recurrence and 0.692 (95% CI: 0.620–0.765) and 0.697 (95% CI: 0.623–0.772) for death, respectively. These were similar to those of the Arch system (0.670 [95% CI: 0.612–0.728] and 0.701 [95% CI: 0.645–0.758] for recurrence, 0.646 [95% CI: 0.558–0.734] and 0.656 [95% CI: 0.582–0.730] for death) and the Sica's system (0.735 [95% CI: 0.680–0.791] and 0.750 [95% CI: 0.698–0.803] for recurrence, 0.696 [95%
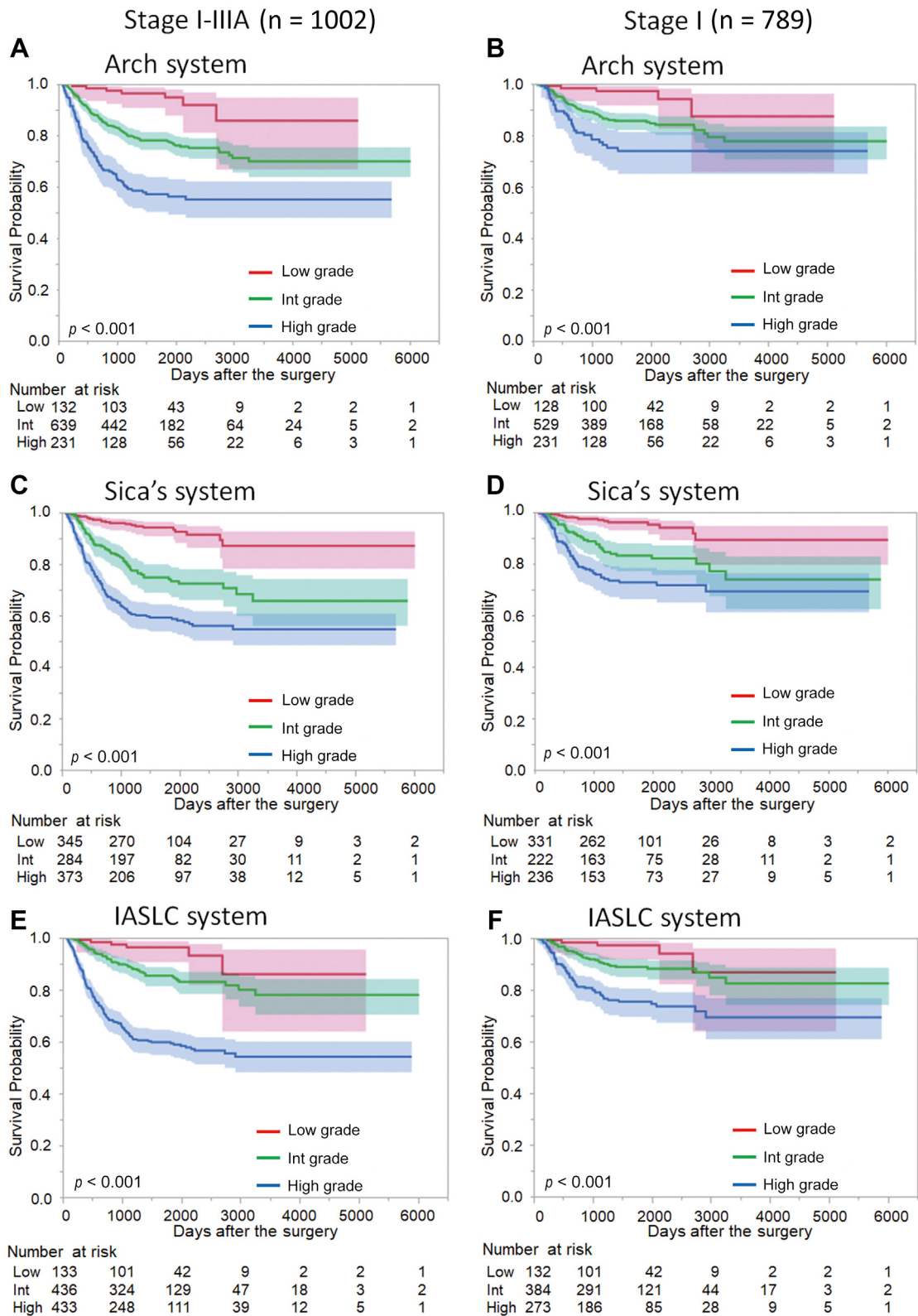
Figure 1. Disease-free survival curves for each grading system stratified by low-grade, int-grade, and high-grade groups. (*A*) The architectural grading system (Arch system) with patients with stages I to IIIA; (*B*) Arch system with patients with stage I; (*C*) Sica's grading system (Sica's system) with patients with stages I to IIIA; (*D*) Sica's system with patients with stage I; (*E*) IASLC grading system (IASLC system) with patients with stages I to IIIA; (*F*) IASLC system with patients with stage I. IASLC, International Association for the Study of Lung Cancer; Int, intermediate.

Table 2. Summary of the Different Grade Assignments Used in the Three Grading Schemes

Variables in the Model	Disease-Free Survival		Overall Survival	
	C-Index (95% CI)	AUC (95% CI)	C-Index (95% CI)	AUC (95% CI)
Grading scheme (stages I-IIIA)				
Baseline model	0.741 (0.706-0.777)	0.773 (0.732-0.812)	0.740 (0.690-0.790)	0.751 (0.700-0.800)
Architectural system[a]	0.763 (0.729-0.797)	0.796 (0.759-0.833)	0.743 (0.690-0.796)	0.755 (0.703-0.807)
Sica's system[b]	0.786 (0.756-0.816)	0.814 (0.779-0.849)	0.762 (0.719-0.805)	0.773 (0.725-0.821)
IASLC system	0.777 (0.749-0.806)	0.807 (0.771-0.843)	0.767 (0.724-0.810)	0.776 (0.728-0.825)
Grading scheme (stage I)				
Baseline model	0.616 (0.557-0.675)	0.644 (0.583-0.705)	0.653 (0.572-0.734)	0.663 (0.590-0.736)
Architectural system[a]	0.670 (0.612-0.728)	0.701 (0.645-0.758)	0.646 (0.558-0.734)	0.656 (0.582-0.730)
Sica's system[b]	0.735 (0.680-0.791)	0.750 (0.698-0.803)	0.696 (0.617-0.774)	0.697 (0.626-0.769)
IASLC system	0.701 (0.651-0.750)	0.722 (0.668-0.776)	0.692 (0.620-0.765)	0.697 (0.623-0.772)

Note: Baseline model represents clinical characteristics only. HG patterns include solid, micropapillary, and complex glandular patterns.
[a]Architectural system was based on the predominant histology.
[b]Sica's system was based on two predominant patterns.
AUC, area under the curve; CI, confidence interval; C-index, concordance index; HG patterns, high-grade pattern; IASLC, International Association for the Study of Lung Cancer.

CI: 0.617–0.774] and 0.697 [95% CI: 0.626–0.769] for death) (Table 2).

Reproducibility Assessment

We evaluated the interobserver agreement using 100 randomly selected invADCs. When these were grouped with the IASLC grading system by the other observer (KT), we found that the κ value was 0.94. Excellent interobserver agreement was found for distinguishing tumors into three categories.

Discussion

Tumor grading systems are an important component of pathologic evaluations because they are used for selecting additional therapy in many cancer types. The IASLC Pathology Committee recently proposed a new grading system for lung ADC on the basis of multiple cohorts.[9] In this study, we reported that the proposed grading system for invasive lung ADC had prognostic significance after applying it to a large Japanese cohort.

After the study of the new histologic subtype of lung ADC,[14] some grading systems, including the predominant pattern-based grading system and the two most predominant pattern-based grading systems, were proposed as promising candidates because of their prognostic significance.[3,4] In this context, the new histologic grading system for invADC was proposed by the IASLC.[9] Nevertheless, the IASLC model has been constructed with three independent data sets that were composed of Caucasian or mixed populations. In our study, the IASLC's grading model, which was based on the combination of the most predominant pattern and the high-grade histologic pattern ($\geq$20%), had good performance. This included a C-index of 0.777 and an AUC of 0.807 for recurrence and a C-index of 0.767 and an AUC of 0.776 for death in the all-stage cohort (a C-index of 0.701 and an AUC of 0.722 for recurrence and a C-index of 0.692 and an AUC of 0.697 for death in the stage I cohort). These results had similar performance to the results of the IASLC report. In general, an AUC of 0.5 suggests no discrimination; AUC of 0.7 to 0.8, acceptable; AUC of 0.8 to 0.9, excellent; and AUC greater than 0.9, outstanding.[15] Thus, we concluded that the model is acceptable as a prognostic indicator for both recurrence and death, even in Japanese patients with resected lung ADC. This is the first validation study with a large Asian cohort. In addition, the Sica's system revealed nearly identical results (C-index and AUC) to those of the IASLC system, and further studies are needed.

A limitation of this study was that IMA was excluded. This subtype is the most common variant of lung ADC found in 51 patients in our cohort; thus, we consider that the applicability and performance of the grading system should be evaluated in the future. Furthermore, reproducibility is another important element for evaluating the performance of grading systems. In our study, we found that interobserver agreement of the IASLC grading system was excellent ($\kappa = 0.94$). This could be because this study was conducted at a single facility; thus, we consider that further study at multiple facilities should be conducted in the future.

In conclusion, we revealed the utility of the IASLC system in determining the prognosis of invADC. We also revealed that besides the IASLC grading system, the Sica's system accurately reflected the prognosis of patients with lung ADC. Though there is a need for additional investigation regarding the variants of invADC (which are excluded in our study) and the interobserver agreement, we believe that the IASLC grading system revealed acceptable performance for prognostic evaluation and will provide physicians with better information for postsurgery treatment.

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *Journal of Thoracic Oncology* at www.jto.org and at https://doi.org/10.1016/j.jtho.2021.04.008.

References

1. Travis WD, Brambilla E, Burke A, Marx AP, Nicholson AG, eds. *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart*. 4th ed. Lyon, France: International Agency for Research on Cancer (IARC); 2015.
2. Behera M, Owonikoko TK, Gal AA, et al. Lung adenocarcinoma staging using the 2011 IASLC/ATS/ERS classification: a pooled analysis of adenocarcinoma in situ and minimally invasive adenocarcinoma. *Clin Lung Cancer*. 2016;17:e57-e64.
3. Yoshizawa A, Motoi N, Riely GJ, et al. Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. *Mod Pathol*. 2011;24:653-664.
4. Sica G, Yoshizawa A, Sima CS, et al. A grading system of lung adenocarcinomas based on histologic pattern is predictive of disease recurrence in stage I tumors. *Am J Surg Pathol*. 2010;34:1155-1162.
5. Moreira AL, Joubert P, Downey RJ, Rekhtman N. Cribriform and fused glands are patterns of high-grade pulmonary adenocarcinoma. *Hum Pathol*. 2014;45:213-220.
6. Kadota K, Yeh YC, Sima CS, et al. The cribriform pattern identifies a subset of acinar predominant tumors with poor prognosis in patients with stage I lung adenocarcinoma: a conceptual proposal to classify cribriform predominant tumors as a distinct histologic subtype. *Mod Pathol*. 2014;27:690-700.
7. Warth A, Muley T, Kossakowski C, et al. Prognostic impact and clinicopathological correlations of the cribriform pattern in pulmonary adenocarcinoma. *J Thorac Oncol*. 2015;10:638-644.
8. Rokutan-Kurata M, Yoshizawa A, Nakajima N, et al. Discohesive growth pattern (Disco-p) as an unfavorable prognostic factor in lung adenocarcinoma: an analysis of 1062 Japanese patients with resected lung adenocarcinoma. *Mod Pathol*. 2020;33:1722-1731.
9. Moreira AL, Ocampo PSS, Xia Y, et al. A grading system for invasive pulmonary adenocarcinoma: a proposal from the International Association for the Study of Lung Cancer Pathology Committee. *J Thorac Oncol*. 2020;15:1599-1610.
10. Boland JM, Wampfler JA, Yang P, Yi ES. Growth pattern-based grading of pulmonary adenocarcinoma-Analysis of 534 cases with comparison between observers and survival analysis. *Lung Cancer*. 2017;109:14-20.
11. Takahashi Y, Eguchi T, Kameda K, et al. Histologic subtyping in pathologic stage I-IIA lung adenocarcinoma provides risk-based stratification for surveillance. *Oncotarget*. 2018;9:35742-35751.
12. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med*. 2011;30:1105-1117.
13. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. 2013;32:5381-5397.
14. Travis WD, Brambilla E, Noguchi M, et al. International Association for the Study of Lung Cancer/American Thoracic Society/European Respiratory Society: International Multidisciplinary Classification of Lung Adenocarcinoma: executive summary. *J Thorac Oncol*. 2011;6:244-285.
15. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5:1315-1316.