Interobserver Agreement in the Nuclear Grading of Primary Pulmonary Adenocarcinoma

Yoshimasa Nakazato, MD,† Akiko Miyagi Maeshima, MD, ‡ Yuichi Ishikawa, MD,§ Yasushi Yatabe, MD,‖ Junya Fukuoka, MD,¶ Tomoyuki Yokose, MD,# Yasuhiko Tomita, MD,** Yuko Minami, MD,†† Hisao Asamura, MD, ‡‡ Keisei Tachibana, MD, §§ Tomoyuki Goya, MD, §§ and Masayuki Noguchi, MD††*

Introduction: Nuclear grading involves an evaluation of the size and shape of nuclei and the percentage of tumor cells that are in the mitotic phase. To estimate the degree of aggressiveness, this approach has been applied to various types of carcinomas, such as breast carcinoma and pulmonary adenocarcinoma (Nakazato et al.). In the present study, we estimated and evaluated the interobserver variability of nuclear grading in primary pulmonary adenocarcinomas.

Methods: We selected 122 primary pulmonary adenocarcinomas measuring 2 cm or less in diameter. Eight pathologists independently evaluated the nuclear factors, using the nuclear grading system reported previously by Nakazato et al. The same pathologists also used both the international multidisciplinary classification of pulmonary adenocarcinoma (2011 International Association for the Study of Lung Cancer classification) and Noguchi's classification, and assessed the extent of the lepidic pattern in the largest cut surface of the tumor. Interobserver agreement was evaluated using the κ statistic. The disease-free survival curves of the patients were obtained using the Kaplan–Meier method and analyzed with the log-rank test.

Results: The mean (±SD) κ values for the two histological classifications, the extent of the lepidic pattern, and nuclear grading were 0.46 ± 0.09, 0.48 ± 0.09, 0.45 ± 0.16, and 0.58 ± 0.09, respectively. The cases judged as negative on the basis of nuclear grading showed a significantly better prognosis (5-year disease-free survival rate; 91.8% ± 2.7) than the positive cases did (68.6% ± 3.1).

Conclusion: Nuclear grading is practical for prognostic evaluation of pulmonary adenocarcinoma. The interobserver agreement for nuclear grading is significantly higher than for histological classifications and the extent of the lepidic pattern. Nuclear grading is a reliable prognostic indicator for small adenocarcinomas.

Key Words: Nuclear grading, Pulmonary adenocarcinoma, Interobserver agreement, Histological classification, Extent of lepidic area.

(*J Thorac Oncol.* 2013;8: 736-743)

The histological grading of lung cancer is a significant indicator of prognosis.[1] Many prognostic factors of pulmonary carcinoma have been shown to correlate with patient outcome.[1] Among them, histopathological subtype, epidermal growth factor receptor expression, and degree of tumor differentiation have been reported to be useful for estimation of cancer malignancy.[2] Although most studies have shown that tumor grading is associated with patient survival, it has been questioned whether it can be applied practically or has clinicopathological significance.[3] Furthermore, methods for assessing tumor grade vary and are usually a composite of different parameters including histopathological subtype, the proportion of the lepidic lesion, and nuclear size.[3,4] Although these parameters yield prognostic information, their relative importance remains uncertain. As all these parameters involve subjective assessment, interobserver variability should be sufficiently tested before they can be used on a practical basis.

Davis et al.[5] found a significant association between the tumor grade of breast cancer and prognosis, as determined by a central review panel and pathologists at a local clinic, despite the fact that precise agreement between the panel and the pathologists was achieved in only 54% of cases. Henson[6] noted that the most appropriate endpoint for testing the value and validity of tumor grading is patient outcome. He considers that variation in grading among individual pathologists *may not significantly interfere with the estimation of prognosis in cancer patients*. However, agreement of interobserver diagnosis or judgment should be suppressed as far as possible.

*Department of Anatomic and Diagnostic Pathology, Dokkyo University School of Medicine, Mibu, Japan; †Division of Diagnostic Pathology, Gunma Prefectural Cancer Center, Gunma, Japan; ‡Clinical Laboratory, National Cancer Center Hospital, Tokyo, Japan; §Division of Pathology, The Cancer Institute Hospital of the Japanese Foundation For Cancer Research, Tokyo, Japan; ‖Departments of Pathology and Molecular Diagnostics, Aichi Cancer Center Hospital, Nagoya, Japan; ¶Laboratory of Pathology, Toyama University Hospital, Toyama, Japan; #Department of Pathology, Kanagawa Cancer Center, Kanagawa, Japan; **Department of Pathology, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka, Japan; ††Department of Pathology, Faculty of Medicine, University of Tsukuba, Ibaraki, Japan; ‡‡Thoracic Surgery Divisions, National Cancer Center Hospital, Tokyo, Japan; and §§Department of Surgery, Institute of Medical Sciences, Kyorin University, Tokyo, Japan.

Disclosure: The authors declare no conflicts of interest.

Address for correspondence: Masayuki Noguchi, MD, Department of Pathology, Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, Ibaraki 305–8575, Japan. E-mail: nmasayuk@md.tsukuba.ac.jp

ISSN: 1556-0864/13/0806-0736

Our previous study of the relationship between nuclear grade and prognosis revealed a strong correlation, even when data were collected from more than one pathologist.[4] Taking interobserver agreement into consideration, we investigated interobserver variability of the nuclear grading system reported previously. The present study examined the usefulness of nuclear grading in comparison with the international multidisciplinary classification of pulmonary adenocarcinoma,[7] Noguchi's classification,[8] and the extent of the lepidic pattern in the largest cut surface of the tumor.[9,10]

PATIENTS AND METHODS

Patients

Primary tumors were obtained from 275 patients with pulmonary adenocarcinomas 2 cm or lesser in maximum diameter, who had been treated surgically between January 2001 and December 2003 at the National Cancer Center Hospital, Tokyo. From this series of 275 patients, 95 who were alive without tumor were randomly selected. In addition, 27 patients who had a relapse were also selected. A total of 122 (95 + 27) patients were selected for this study. A scheme of the study population is shown in Supplemental Figure 1 (Supplemental Digital Content 1, http://links.lww.com/JTO/A405). One hundred three patients underwent standard surgical resection of their tumors along with mediastinal and pulmonary hilar lymph node dissection. Five patients underwent segmentectomy of their tumors along with pulmonary hilar lymph node dissection. Fourteen patients underwent partial resection of their tumors along with a sampling lymph node dissection. Informed consent for specimen collection was obtained from all patients. None of the patients selected had received neoadjuvant chemotherapy or radiotherapy before surgery.

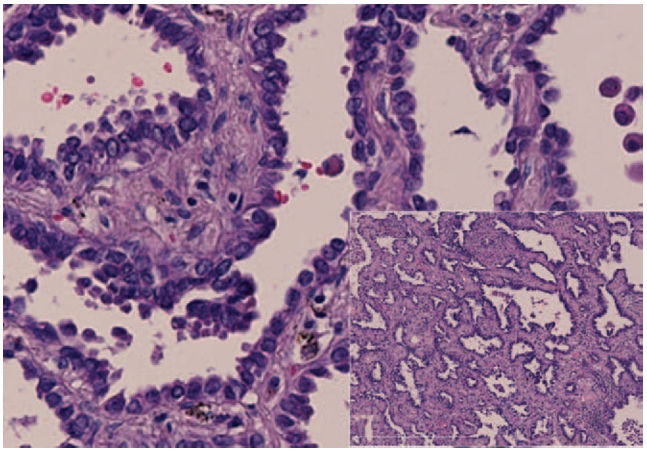

FIGURE 1. Histology of a small pulmonary adenocarcinoma judged to be *lepidic-pattern predominant* on the basis of the international multidisciplinary classification (8 of 8 pathologists), *type B* by Noguchi's classification (4 of 8 pathologists), and *negative* by the nuclear grading system (8 of 8 pathologists). Hematoxylin and eosin stain, original magnifications ×400 and ×40.

Tissue Specimens and Pathologic Information

The resected specimens were fixed with 10% to 15% neutral buffered formalin at room temperature and then embedded in paraffin for histologic examination. As the tumors were 2 cm or less in maximum diameter, we were able to make one paraffin block per tumor. The cut surface of each tumor was stained with hematoxylin and eosin and examined by light microscopy. Tumors were classified according to the criteria of the international multidisciplinary classification[4] of lung adenocarcinoma (2011 International Association for the Study of Lung Cancer [IASLC] classification) and the histological criteria proposed by Noguchi et al.[5] (Noguchi's classification). In terms of the extent of the lepidic pattern, we used 75% as cutoff rates. In the previous studies by Yokose et al.,[11] Anami et al.,[10] and Yoshizawa et al.[2] 75%, 50%, and predominancy of lepidic growth area (2011 IASLC classification) were used as cutoff rates to compare the prognosis. In the present study, we chose the strictest criteria to represent the extent of the lepidic growth pattern. Diagnoses were performed by eight pathologists (MN, YI, AM, JF, YT, TY, YY, and YN) using a BX50 or BX51 microscope (Olympus, Tokyo, Japan). To avoid any errors through use of different models of microscope, two participants (YN and MN) using BX50 and BX51 microscopes confirmed that their results were similar. However, it is always necessary to be mindful of potential errors when applying such criteria practically in routine examinations. To decrease such measurement bias, we used a blind test method. Cancer staging of specimens was determined in accordance with the International Union against Cancer tumor, node, metastasis classification of malignant tumors, 7th edition.

Nuclear Grading Criteria

We formerly reported that the nuclear grading system to be a useful parameter for prognostication of small pulmonary adenocarcinomas.[3] In the present trial, we attempted to apply nuclear grading to routine histological diagnosis. As a standard for nuclear size, we used small lymphocytes. The mean (±SD) values of the nuclear area (NA) and nuclear dimension (ND) of small lymphocytes were NA 14 μm^2 (± 4 SD) and ND 3.9 μm (± 0.03 SD), respectively. Therefore, the NA of 67 μm^2, which was indicated to be the critical size in the former study was approximately five times larger than that of lymphocytes, and the ND level of 10.7 μm, which was also indicated previously to be the critical size was approximately three times larger. We chose fields where the large nuclei were frequently observed from the largest cut surface of each tumor. Necrotic and inflammatory areas were avoided, and overlapping nuclei were omitted. Any tumor cell was judged to be positive, if its NA and ND were five and three times larger, respectively, than the corresponding values of small lymphocytes.

The participants reviewed each slide in detail, using low magnification (×40) and selected a few areas that included tumor cells with large nuclei. The number of tumor cells judged to be positive was then counted, and a field was judged as *positive field* if it contained more than five positive cells. The case was judged as a *positive case* if more than three such positive fields were observed.

