Interobserver Variability in Diagnosing High-Grade Neuroendocrine Carcinoma of the Lung and Comparing It with the Morphometric Analysis

Seung Yeon Ha · Joungho Han[1]
Wan-Seop Kim[2] · Byung Seong Suh[3]
Mee Sook Roh[4]

Department of Pathology, Gachon University Gil Hospital, Gachon University of Medicine and Science, Incheon; [1]Department of Pathology, Sungkyunkwan University School of Medicine; [2]Department of Pathology, Konkuk University School of Medicine; [3]Department of Occupational Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul; [4]Department of Pathology, Dong-A University College of Medicine, Busan, Korea

Corresponding Author
Mee Sook Roh, M.D.
Department of Pathology, Dong-A University College of Medicine, Dongdaesin-dong 3-ga, Seo-gu, Busan 602-715, Korea
Tel: +82-51-240-2833
Fax: +82-51-243-7396
E-mail: msroh@dau.ac.kr

Background: Distinguishing small cell lung carcinoma (SCLC) and large cell neuroendocrine carcinoma (LCNEC) of the lung is difficult with little information about interobserver variability. **Methods:** One hundred twenty-nine cases of resected SCLC and LCNEC were independently evaluated by four pathologists and classified according to the 2004 World Health Organization criteria. Agreement was regarded as "unanimous" if all four pathologists agreed on the classification. The kappa statistic was calculated to measure the degree of agreement between pathologists. We also measured cell size using image analysis, and receiver-operating-characteristic curve analysis was performed to evaluate cell size in predicting the diagnosis of high-grade neuroendocrine (NE) carcinomas in 66 cases. **Results:** Unanimous agreement was achieved in 55.0% of 129 cases. The kappa values ranged from 0.35 to 0.81. Morphometric analysis reaffirmed that there was a continuous spectrum of cell size from SCLC to LCNEC and showed that tumors with cells falling in the middle size range were difficult to categorize and lacked unanimous agreement. **Conclusions:** Our results provide an objective explanation for considerable interobserver variability in the diagnosis of high-grade pulmonary NE carcinomas. Further studies would need to define more stringent and objective definitions of cytologic and architectural characteristics to reliably distinguish between SCLC and LCNEC.

Pulmonary neuroendocrine (NE) tumors represent four subgroups with different biologic behavior, from low-grade typical carcinoid and intermediate-grade atypical carcinoid to high-grade small cell lung carcinoma (SCLC) and large cell NE carcinomas (LCNEC). These subtypes are classified according to morphologies observed by hematoxylin and eosin (H&E) staining, as detailed in the 2004 World Health Organization (WHO) classification.[1]

LCNEC of the lung was first introduced in 1991 by Travis *et al.*,[2] who proposed it to be a distinct entity of high-grade NE carcinoma with light microscopic characteristics that differ from those of high-grade SCLC. According to the 2004 WHO classification, LCNECs are characterized by large cell size, low nuclear to cytoplasmic ratios, frequent nucleoli, vesicular, coarse or fine chromatin, organoid growth pattern with palisading or

rosettes, high mitotic rates, and necrosis. On the contrary, tumor cells of SCLC are round, oval, or spindle-shaped, usually less than the size of three small resting lymphocytes, and have scant cytoplasm, finely granular chromatin, and absent or inconspicuous nucleoli.[1,3] Although LCNEC is considered to be a disease entity that is distinct from SCLC, these two types of pulmonary NE carcinoma share common morphological characteristics, with a phenotypical spectrum that runs from SCLC to LCNEC. Therefore, distinguishing SCLC and LCNEC can occasionally be difficult in daily diagnostic practice. A recent study reported striking variability among assessors in diagnosing SCLC and LCNEC and showed fair overall agreement for all cases.[4]

However, there is limited information on the level of interobserver agreement between pathologists for the classification of

SCLC and LCNEC in Korea. This study was conducted to document the interobserver variability in diagnosing SCLC and LCNEC among a group of four pathologists with a special interest in lung cancer. Furthermore, we undertook a morphometric analysis to compare the diagnostic power of cell size in diagnosing SCLC and LCNEC, because cell size is one of the most important diagnostic criteria to distinguish LCNEC from SCLC based on the WHO classification.[1]

MATERIALS AND METHODS

Case selection

We evaluated a total of 129 cases that had the histologic diagnosis of primary pulmonary high-grade NE carcinoma. The tumors were originally diagnosed as SCLC (n = 35) or LCNEC (n = 94). Immunohistochemical staining for such general NE markers as chromogranin, synaptophysin, and CD56 was performed if necessary at the time when the original diagnosis was made at each institute. The tissues were obtained from patients who underwent surgery between 1999 and 2008 at two university hospitals, including Seoul Samsung Hospital and Dong-A University Medical Center. To ensure that there would be enough specimens for pathologic examination, only surgical cases were considered. Since we used existing data that did not identify individual subjects, informed consent by the study participants was not necessary and waived for this study.

Pathology review

A representative H&E-stained slide from each case was circulated among four pathologists and independently reviewed based on the 2004 WHO criteria.[1] Of the four pathologists that participated in this study, one pathologist had 22 years of experience in pathology, one had 20 years of experience, and the remaining two pathologists had 17 years of experience each. They were all experienced pulmonary pathologists.

Statistical analysis for interobserver agreement

Agreement was regarded as "unanimous" if all four pathologists agreed on a particular diagnosis, as a "majority" if three or more of four pathologists agreed, and as a "lack of consensus" if two pathologists had opposite diagnoses. Since only four observers participated in this study and they were all experienced pulmonary pathologists, we considered cases with no unanimous agreement to be "debated cases."

To measure the interobserver agreement between four pathologists, the generalized kappa value was calculated using SPSS ver. 18.0 (SPSS Inc., Chicago, IL, USA). We adopted the generally accepted convention for interpreting kappa values; values from 0.0-0.20 corresponded to slight agreement, 0.21-0.40 fair agreement, 0.41-0.60 moderate agreement, 0.61-0.80 substantial agreement, and 0.81-1.00 almost perfect agreement.

Morphometric analysis

We carried out morphometric analysis to measure the maximal diameter of cells that discriminate between SCLC and LCNEC using a previously described method.[5] Among 129 studied cases, we further evaluated 5 μm H&E-stained sections from 66 high-grade pulmonary NE carcinomas, including 24 unanimous LCNEC, 10 unanimous SCLC, and 32 debated cases (tumors with lack of unanimous agreement; 16 majority LCNEC, 6 majority SCLC and 10 tumors with lack of consensus) after pathology review by four observers for morphometric analysis. Images were captured using a DP70 digital camera (Olympus, Tokyo, Japan) attached to a BX51 microscope with a ×40 objective. The final image captured on the monitor had a magnification of ×400. The cell diameters of 200 randomly selected tumor cells and the cell diameters of 20 mature lymphocytes were measured for each case. The actual measurements of the morphometric parameters were done using an I-solution image analysis system ver. 8.4 (IMT i-Solution Inc., Coquitlam, BC, Canada). The system segmented the nuclei automatically from the background, and the investigator needed only to manually edit adjacent nuclei that overlapped with each other.

Statistical analysis for morphometric analysis

Receiver operating characteristic (ROC) curve analysis was performed to identify the most useful cut-off value of cell diameter that provided the greatest sum of sensitivity and specificity in predicting a unanimous diagnosis of high-grade NE carcinomas based on cell size. The area under the ROC curve (AUC) and 95% confidence interval that did not include the 0.5 value was considered to be the cell diameter that had some ability to distinguish between the groups. All calculations were performed with SPSS ver. 18.0.

RESULTS

Agreement

After the four pathologists independently performed a pathology review, unanimous agreement was achieved in 71 of 129 cases (55.0%), and a majority agreement was achieved in 102 of 129 cases (79.1%). Based on the original diagnosis,

unanimous agreement occurred for 52 (55.3%) of 94 LCNECs and for 19 (54.3%) of 35 SCLCs. A majority agreement occurred for 73 (77.7%) of 94 LCNECs and for 29 (82.9%) of 35 SCLCs.

Table 1. Agreement of the original diagnosis by each institute for all 129 cases studied

Original diagnosis	Agreement		
	Unanimous	Majority	Lack of consensus
Total cases (n=129)	71 (55.0)	102 (79.1)	27 (20.9)
LCNEC (n=94)	52 (55.3)	73 (77.7)	21 (22.3)
SCLC (n=35)	19 (54.3)	29 (82.9)	6 (17.1)

Values are presented as number (%).
LCNEC, large cell neuroendocrine carcinoma; SCLC, small cell lung carcinoma.

Of the originally diagnosed cases of LCNEC, 21 (22.3%) lacked a consensus (two observers had opposite diagnoses), and 6 (17.1%) of 36 SCLCs lacked a consensus. Agreement was similar for both LCNEC and SCLC (Table 1, Fig. 1).

The kappa statistics for comparing the diagnosis of the four observers for the overall group of evaluated tumors are summarized in Table 2. The kappa values ranged from 0.35 to 0.81; one of the values fell into the almost perfect agreement category, one of the values fell into the substantial agreement category, three of the values fell into the moderate agreement category, and one fell into the fair agreement category.
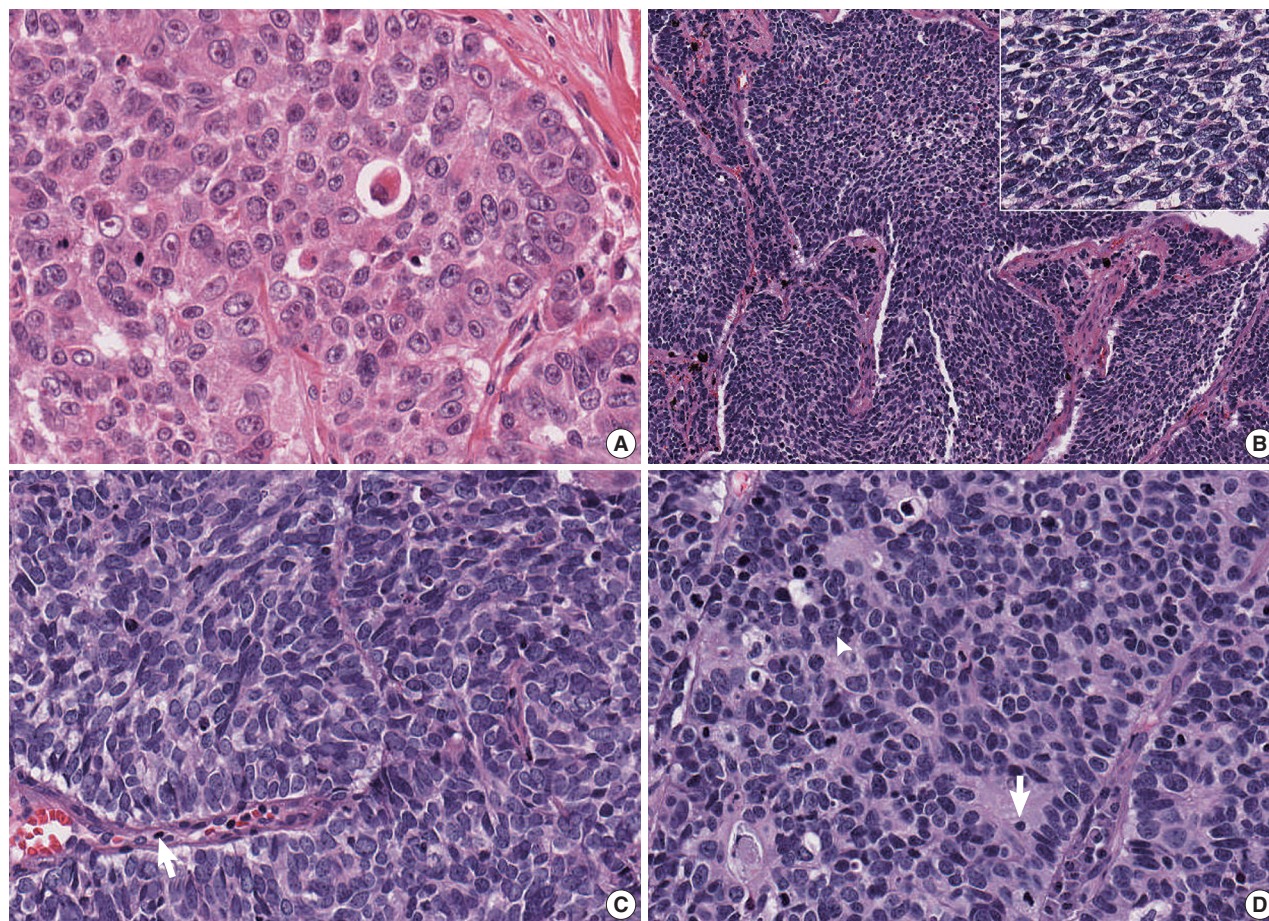

Fig. 1. Representative histologic features of pulmonary high-grade neuroendocrine (NE) carcinomas. (A) Unanimous large cell neuroendocrine carcinoma (LCNEC) consists of organoid nesting of large, polygonal tumor cells that have abundant eosinophilic cytoplasm, coarsely granular chromatin, and prominent nucleoli. (B) Unanimous small cell lung carcinoma (SCLC) consists of a sheet-like growth pattern of tumor cells that are smaller than three lymphocytes. (B, inset) The cells have scant cytoplasm and finely granular chromatin pattern and inconspicuous nucleoli. (C) A debated case of high-grade NE carcinoma with a lack of unanimous agreement shows tumor cells with intermediate nuclear size larger than those seen in the tumor shown in (B) but smaller than those seen in the tumor shown in (A), which corresponds to three to four times the size of a lymphocyte (arrow). (D) Another case of high-grade NE carcinoma that lacks a unanimous diagnosis. Two observers diagnose it as SCLC due to the small tumor cell size (two to three times the size of a lymphocyte [arrow]), whereas the remaining two observers diagnose it as LCNEC, because the cytoplasm is not so scant, and a few nucleoli (arrowhead) and rosette-like pattern are observed.

Morphometric analysis

ROC analysis indicated that cell diameter for unanimous LCNEC yielded an AUC of 0.773 (95% confidence interval, 0.654 to 0.891). At the cut-off value of 33.8 μm, cell diameter had 70.8% sensitivity and 73.8% specificity in discriminating unanimous LCNEC from other high-grade NE carcinomas

(unanimous SCLC and debated cases). On the contrary, cell diameter for unanimous SCLC yielded an AUC of 0.905 (95% confidence interval, 0.826 to 0.985). At the cut-off value of 28.6 μm, cell diameter had 80.0% sensitivity and 92.9% specificity in discriminating unanimous SCLC from other high-grade NE carcinomas (unanimous LCNEC and debated cases) (Fig. 2). The kappa value for comparing the diagnosis based on the cut-off value of morphometric analysis and unanimous diagnosis evaluated by four observers was 0.44. Of the 32 debated cases that lacked unanimous agreement, 18 (56.2%) cases fell into cell diameter between 28.6 μm and 33.8 μm (Table 3). Moreover, seven of 10 tumors with lack of consensus (70.0%) fell into cell diameter between 28.6 μm and 33.8 μm (Table 4).

Table 2. Interobserver agreement of four pathologists according to the kappa statistics for all 129 studied cases

Observer	1	2	3	4
1	×	0.57	0.69	0.51
2		×	0.81	0.35
3			×	0.47
4				×

Table 3. Comparison of diagnosis by four pathologists with diagnosis according to the cut-off point using morphometric analysis for 66 studied cases

Diagnosis according to the cut-off point	Diagnosis by four observers			Kappa value
	Unanimous LCNEC (n=24)	Debated high-grade NE carcinoma (n=32)	Unanimous SCLC (n=10)	
LCNEC (>33.8 μm)	17 (70.8)	11 (34.4)	0 (0)	0.44
≥ 28.6 μm and ≤33.8 μm	6 (25.0)	18 (56.2)	2 (20.0)	
SCLC (<28.6 μm)	1 (4.2)	3 (9.4)	8 (80.0)	

Values are presented as number (%).
LCNEC, large cell neuroendocrine carcinoma; NE, neuroendocrine; SCLC, small cell lung carcinoma.

Table 4. Comparison of diagnosis by four pathologists with diagnosis according to the cut-off point using morphometric analysis for 32 debated cases

Diagnosis according to the cut-off point	Diagnosis by four observers		
	Majority LCNEC (n=16)	High-grade NE carcinoma with lack of consensus (n=10)	Majority SCLC (n=6)
LCNEC (>33.8 μm)	8 (50.0)	2 (20.0)	1 (16.7)
≥ 28.6 μm and ≤ 33.8 μm	7 (43.8)	7 (70.0)	4 (66.6)
SCLC (<28.6 μm)	1 (6.2)	1 (10.0)	1 (16.7)

Values are presented as number (%).
LCNEC, large cell neuroendocrine carcinoma; NE, neuroendocrine; SCLC, small cell lung carcinoma.
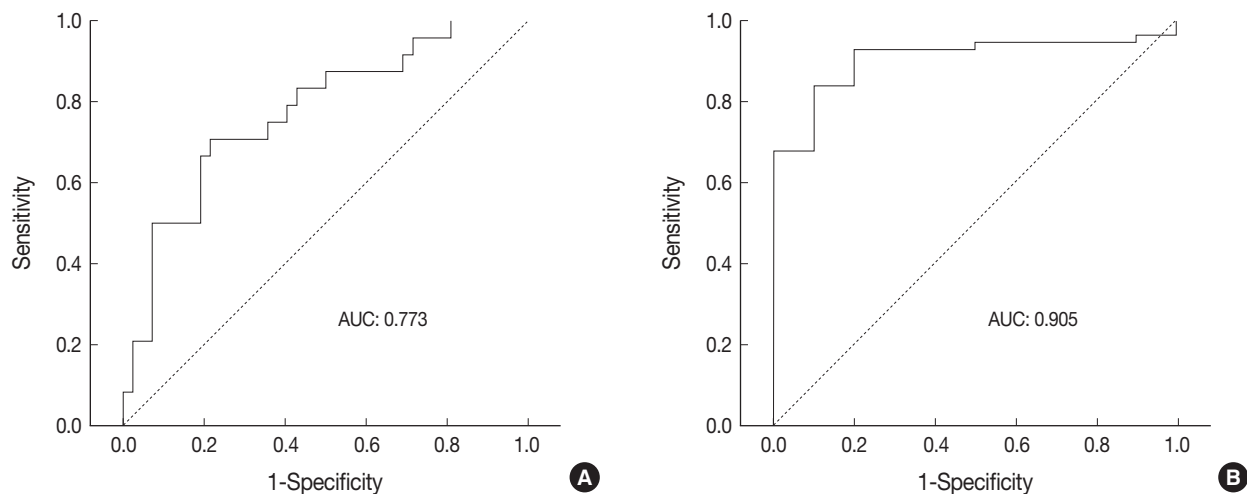

Fig. 2. Receiver operating characteristic (ROC) curve analysis using cell diameter for discriminating large cell neuroendocrine carcinoma (LCNEC) and small cell lung carcinoma (SCLC). (A) At the cut-off value of 33.8 μm, cell diameter for unanimous LCNEC yields an area under the ROC curve (AUC) of 0.773 (95% confidence interval, 0.654 to 0.891), with 70.8% sensitivity and 73.8% specificity in discriminating unanimous LCNEC from other high-grade neuroendocrine (NE) carcinomas (unanimous SCLC and debated cases). (B) At the cut-off value of 28.6 μm, the cell diameter for unanimous SCLC yields an AUC of 0.905 (95% confidence interval, 0.826 to 0.985), with 80.0% sensitivity and 92.9% specificity in discriminating unanimous SCLC from other high-grade NE carcinomas (unanimous LCNEC and debated cases).

DISCUSSION

Our results provided an objective explanation for the considerable levels of interobserver variability in the diagnosis of high-grade pulmonary NE carcinomas, with kappa values that ranged from 0.35 (fair agreement) to 0.81 (almost perfect agreement). Den Bakker *et al.*[4] reported that there was striking variability amongst observers in diagnosing SCLC and LCNEC with variable agreement, from weak agreement (kappa value = 0.19) to good agreement (kappa value = 0.54), and the overall level of agreement for all cases was fair (kappa value = 0.40). In their study, unanimity of diagnosis was achieved in 11.8% of 170 studied cases, and no consensus diagnosis was reached in 20.6%. A study by Travis *et al.*[6] revealed that unanimous diagnostic agreement occurred in 70% of SCLC and 40% of LCNECs. Similarly, our study also demonstrated that it was difficult to reach a unanimous agreement, which was achieved in only 50.0% of the 129 studied cases. These studies of interobserver variability thus raised questions about whether the distinction between SCLC and LCNEC in the lung based WHO categories is truly reproducible.

Difficulty of diagnosis in high-grade NE carcinoma is thought to occur for a variety of reasons, including the continuum of cell size and morphology from SCLC to LCNEC. Moreover, it may be true that the criteria for differentiating SCLC and LCNEC are subjectively interpreted by pathologists. In this study, the ratio of LCNEC to SCLC diagnoses varied from 0.8 to 4.8 according to observers. As shown in Fig. 1C and D, some pathologists may place emphasis on the cell size in diagnosing pulmonary NE carcinomas, whereas other pathologists may diagnose a tumor focusing on the shape of the tumor cell, organoid pattern, nuclear/cytoplasmic ratio, or nucleoli. Practically, all the features except for nuclear size are qualitative rather than quantitative.

For this reason, our attention shifted to morphometric analysis of cell size that is objective and a graphical method to compare discriminating power. In this study, the kappa value for comparing the diagnosis according to the cut-off point of morphometric analysis and unanimous diagnosis by the four observers was 0.44. Of the 32 debated tumors that lacked unanimous agreement, 18 (56.2%) fell into nuclear diameter between 28.6 μm and 33.8 μm, which corresponds to three to four times the size of a lymphocyte. Moreover, seven of 10 tumors with lack of consensus (70.0%) fell into having a cell diameter between 28.6 μm and 33.8 μm. The morphological WHO criteria for diagnosing SCLC and LCNEC have proposed the use of an arbitrary cut-off that is three times the size of a lymphocyte to distinguish 'small' from 'large' cells.[1] However, Marchevsky *et al.*[7] reported that 5 of 16 SCLCs exhibited a predominant number of neoplastic cells that were larger than three normal lymphocytes, while 4 of 12 LCNECs had a predominant number of small cells classified by cell size alone, and these results were similar to those of our study. Therefore, we demonstrated that there is a continuum of cell sizes from SCLC to LCNEC with substantial cell size overlap and indicated that cell size alone is insufficient for reliable distinction between SCLC and LCNEC, and those cases falling in the middle cell size are difficult to categorize.

Meanwhile, to determine the biological identity, similarity, and difference among high-grade pulmonary NE carcinomas, there is an obvious need for well-defined criteria. It is well-known that misclassification of a tumor could potentially mask differences in biological characteristics and clinical outcome. Due to the lack of a clear understanding of these high-grade NE carcinomas, studies that have reported inconsistency in clinical management of high-grade NE carcinomas suggest that LCNEC should be treated as a non-small cell lung cancer or SCLC.[8,9] This is a critical point, especially in clarifying the histology-specific sensitivity to treatment.

Although a constellation of morphological features distinguishing SCLC and LCNEC clearly exists, the dilemma of how to classify tumors that have features intermediate between SCLC and LCNEC must be resolved. It has been suggested that an alternative approach to classifying SCLC and LCNEC may be adopted by combining them into a single group of high-grade NE carcinoma.[7] However, there is insufficient data to justify combining all the variants as a single entity. Therefore, further studies of high-grade pulmonary NE carcinomas would need to define more stringent and objective definitions of cytologic and architectural characteristics that would enable a pathologist to reliably distinguish between SCLC and LCNEC.

REFERENCES

1. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. World Health Organization classification of tumour: pathology and genetics of tumours of the lung, pleura, thymus and heart. Lyon: IARC Press, 2004.

2. Travis WD, Linnoila RI, Tsokos MG, *et al*. Neuroendocrine tumors of the lung with proposed criteria for large-cell neuroendocrine carcinoma: an ultrastructural, immunohistochemical, and flow cytometric study of 35 cases. Am J Surg Pathol 1991; 15: 529-53.

3. Travis WD. Advances in neuroendocrine lung tumors. Ann Oncol 2010; 21 Suppl 7: vii65-71.

4. den Bakker MA, Willemsen S, Grünberg K, *et al*. Small cell carcinoma of the lung and large cell neuroendocrine carcinoma interobserver variability. Histopathology 2010; 56: 356-63.

5. Jeong SM, Ha SY, An J, *et al*. Morphometric analysis for pulmonary small cell carcinoma using image analysis. Korean J Pathol 2011; 45: 87-91.

6. Travis WD, Gal AA, Colby TV, Klimstra DS, Falk R, Koss MN. Reproducibility of neuroendocrine lung tumor classification. Hum Pathol 1998; 29: 272-9.

7. Marchevsky AM, Gal AA, Shah S, Koss MN. Morphometry confirms the presence of considerable nuclear size overlap between "small cells" and "large cells" in high-grade pulmonary neuroendocrine neoplasms. Am J Clin Pathol 2001; 116: 466-72.

8. Varlotto JM, Medford-Davis LN, Recht A, *et al*. Should large cell neuroendocrine lung carcinoma be classified and treated as a small cell lung cancer or with other large cell carcinomas? J Thorac Oncol 2011; 6: 1050-8.

9. Asamura H, Kameya T, Matsuno Y, *et al*. Neuroendocrine neoplasms of the lung: a prognostic spectrum. J Clin Oncol 2006; 24: 70-6.