The Korean Journal of Pathology 2010; 44: 267-71
DOI: 10.4132/KoreanJPathol.2010.44.3.267

The Interobserver Variability for Diagnosing Pulmonary Carcinoid Tumor

Chang Hun Lee · Hee Kyung Chang[1]
Hyoun Wook Lee[2] · Dong Hoon Shin
Mee Sook Roh[3]

Department of Pathology, Pusan National University College of Medicine; [1]Department of Pathology, Kosin University College of Medicine; [2]Department of Pathology, Inje University College of Medicine; [3]Department of Pathology, Dong-A University College of Medicine, Busan, Korea

Received : July 23, 2009
Accepted : December 3, 2009

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: 051-240-2833
Fax: 051-243-7396
E-mail: msroh@dau.ac.kr

Background : Although the grade of pulmonary carcinoid tumor is routinely reported in pathology practice, there is a paucity of data on the level of agreement between pathologists. **Methods :** Data for 30 cases of surgically resected pulmonary tumors diagnosed as carcinoid tumors (19 typical carcinoids [TCs] and 11 atypical carcinoids [ACs]) were retrieved from four university hospitals. These cases were independently evaluated by five pathologists and were classified according to the 2004 World Health Organization (WHO) classification. Agreement was regarded as "unanimous" if all five pathologists agreed, and as a "majority" if four agreed. The kappa statistic was calculated to measure the degree of agreement between pathologists. **Results :** Unanimous agreement was achieved for 50.0% and a majority agreement for 83.3% of the 30 cases. The range of the kappa values extended from 0.37 to 0.89. After a consensus meeting, there was disagreement between the original diagnosis by each institute and the consensus diagnosis by the five pathologists for 40.0% of the 30 cases. Based on the consensus diagnosis, the agreement was greater for TCs than that for ACs. **Conclusions :** Discriminating carcinoid tumors is subject to interobserver variability. This study indicates that there is a need for more careful standardization and application of diagnostic criteria for making the diagnosis of pulmonary carcinoid tumor.

Key Words : Lung; Carcinoid tumor; Observer variation

Pulmonary carcinoid tumors, either typical carcinoid (TC) or atypical carcinoid (AC), are rare neuroendocrine (NE) tumors that account for 2-5% of all lung tumors.[1] The classification of NE tumors has undergone numerous changes over the years. In 1999, carcinoid tumors were classified as TC or AC, and this classification was recommended by the World Health Organization (WHO) as well as by the International Association for the Study of Lung Cancer.[2] In the 2004 WHO system, TC and AC were categorized together under the heading of carcinoid tumors, whereas large cell neuroendocrine carcinoma (LCNEC) was retained as a subtype of large cell carcinoma, and small cell lung carcinoma (SCLC) was retained as an independent category.[3] These criteria sharpened the definition of AC as a NE tumor with a prognosis that's intermediate between TC and high-grade LCNEC and SCLC.

According to the WHO classification, mitotic densities and the presence of necrosis are the most important morphologic features for distinguishing TC and AC of the lung. AC can be differentiated from TC by a mitotic count of 2 to 10 per 2 mm^2 of viable tumor or by the presence of coagulative necrosis (Fig. 1).[3] However, the diagnosis and classification of carcinoid tumors are subject to great interobserver and intraobserver variability.[4] Therefore, the spectrum of pulmonary NE tumors represents one of the most difficult areas in the surgical pathology of the lung. However, there is not much data on the level of interobserver agreement between pathologists for the classification of pulmonary carcinoid tumor. Considering the importance of pathologic diagnoses and their reproducibility, this study was conducted in a retrospective, multi-institutional setting with five pathologists reviewing the tumor histology in order to evaluate the interobserver variability in making the diagnosis of pulmonary carcinoid tumor.

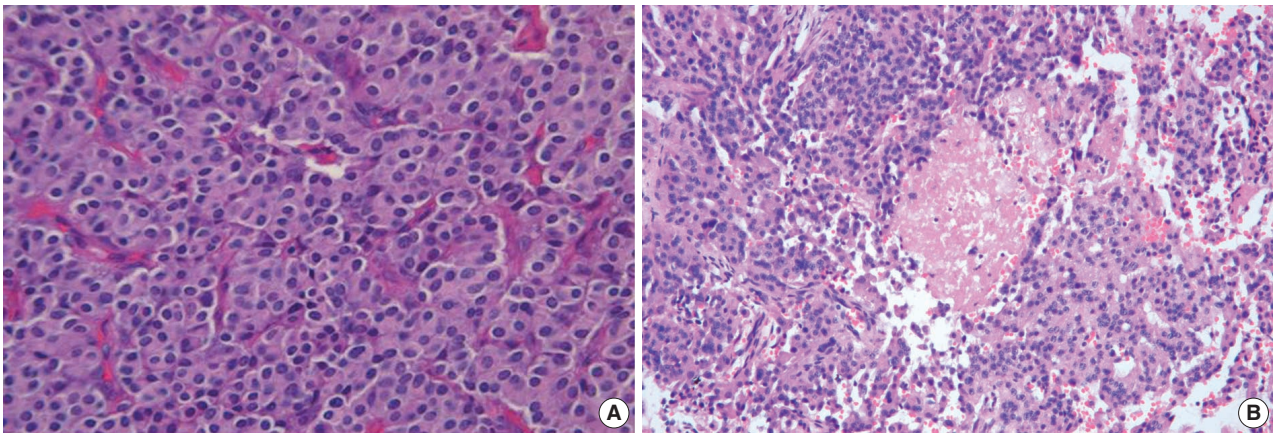

Fig. 1. Representative histologic features of pulmonary carcinoid tumors. (A) Typical carcinoid tumor consists of organoid nests of uniform cells with a moderate amount of eosinophilic cytoplasm and the nuclei with a finely granular chromatin pattern. (B) Atypical carcinoid shows the carcinoid morphology with a punctate focus of necrosis.

MATERIALS AND METHODS

Case selection

We evaluated a total of 30 cases that had the histologic diagnosis of primary pulmonary carcinoid tumor. The tumors were originally diagnosed as TC (n = 19) or AC (n = 11). The tissues were obtained from patients who were operated on between 1999 and 2008 at four university hospitals, including Pusan National University Hospital, Kosin University Hospital, Inje University Hospital and Dong-A University Medical Center. To ensure that there would be enough specimens for pathologic examination, only surgical cases were considered. Clinical records, pathological reports and clinical follow-up information were obtained when available. Postoperative pathological staging was determined according to the guidelines of the American Joint Committee on Cancer.[5]

Pathology review

All hematoxylin and eosin (HE)-stained slides were reviewed in each case based on the 2004 WHO criteria.[3] Mitoses were counted in the areas of highest mitotic activity and were counted in three sets of 10 high power fields (HPFs) and an average of the number of mitotic figures per 10 HPFs was calculated.[6] Mitotic figures were defined by the criteria described by Baak.[7] Calculations were made to estimate the number of HPFs needed to achieve the equivalent of 10 HPFs, or 2 mm^2 for each microscope commonly used by the surgical pathologists according to manufacturer and model.[6] Necrosis consisted of either punctate foci or large infarct-like zones. Superficial ulceration of endobronchial tumors and the sites of prior transbronchial or needle biopsies were not regarded as tumor necrosis.[6] The median number of HE-stained slides that were reviewed for each case was 3.4 (range, 1 to 7). Immunohistochemical staining for such general NE markers as chromogranin, synaptophysin and CD56 was performed at the time when the original diagnosis was made at each institute, if that was necessary.

Five pathologists participated in this study. Two pathologists had more than 20 years of experience in pathology, one pathologist had 15 years experience, and the remaining two pathologists had 5 years of experience in pathology. The slides were analyzed independently and in a blinded manner by the five pathologists. First, the five pathologists performed an independent pathology review, and their respective reports were sent directly to one of the authors (M.S.R.). After the individual results were compared, a review meeting was held to establish a final consensus on the histologic type using the same multi-head microscope in each case, if disagreement occurred between the five pathologists. The diagnosis that was agreed on by the majority of the pathologists was accepted as the definitive consensus diagnosis.

Statistical analysis

Agreement was regarded as "unanimous" if all five pathologists agreed on a particular diagnosis; agreement was regarded as a "majority" if four or more of the five pathologists agreed; it was regarded as a "consensus" if three or more agreed; it was considered as a "lack of consensus" if only two pathologists agreed on a particular diagnosis. To measure the interobserver agreement

between the five pathologists, the generalized kappa value was calculated with the use of Statistical Analysis Software ver. 9.1 (SAS Institute Inc., Cary, NC, 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.

RESULTS

Agreement

After the five pathologists independently performed a pathology review, unanimous agreement was achieved in 15 of 30 cases (50.0%), a majority agreement was achieved in 25 of 30 cases (83.3%) and a consensus agreement was achieved in 30 of 30 cases (100%). For the mitotic criterion to distinguish TC and AC (< 2 per 2 mm^2 vs 2 to 10 per 2 mm^2), unanimous agreement was achieved for 16 of 30 cases (53.3%) and a majority agreement was achieved for 25 of 30 cases (83.3%). For the necrosis criterion (absent vs present), unanimous agreement was achieved for 25 of 30 cases (83.3%) and a majority agreement was achieved for 30 of 30 cases (100%). Agreement using the mitotic criterion showed substantially more variability among the pathologists than that for the necrosis criterion. Agreement results for the actual interpretation in each case are shown in Fig. 2.

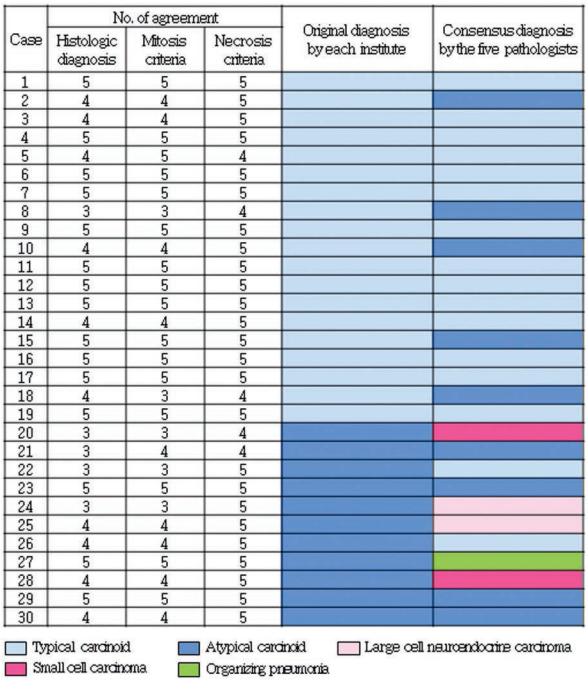

Fig. 2. Agreement results for actual interpretations for each case.

Interobserver agreement

The kappa statistic for comparing the five observers for the overall group of evaluated tumors are summarized in Table 1. The kappa values ranged from 0.37 to 0.89; one of the values fell into the almost perfect agreement category, five 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.

Consensus analysis

To establish final consensus diagnoses on the histologic type for the 15 cases with majority or consensus agreement, a review meeting was held using the same multi-head microscope in each case. An agreement was immediately reached for 11 cases. After further debate, a final consensus diagnosis was also reached for the remaining four cases. The major cause for discordance was the difficulty while evaluating mitotic figures in differentiating them from apoptosis. As the final agreement of the review meeting, the 30 tumors which were originally diagnosed as 19 TCs and 11 ACs, were reclassified as 16 TCs, 9 ACs, 2 LCNECs, 2 SCLCs and 1 organizing pneumonia (OP). There was disagreement between the original and consensus diagnoses for 40.0% (12/30) of the cases. Of the originally diagnosed cases of TC, 5 (26.3%) were changed to AC after the consensus meeting. Of the originally diagnosed cases of AC, only 4 (36.4%) were retained as AC, whereas 2 cases each (each 2 cases, 18.2%) were changed to TC, LCNEC and SCLC, and 1 (9.0%) was changed to OP (Table 2).

The clinicopathological characteristics of TC and AC based on the consensus diagnosis are summarized in Table 3. The tumor size was larger, lymph node metastasis was more frequent and the stage was higher for AC than for TC.

Based on the consensus diagnosis, unanimous agreement

Table 1. Interobserver agreement of the five pathologists according to the kappa statistic for all 30 studied cases

Observer	1	2	3	4	5
1	X	0.67	0.65	0.42	0.89
2		X	0.65	0.52	0.78
3			X	0.37	0.65
4				X	0.52
5					X

Table 2. Comparison of consensus diagnoses by the five pathologists with the original diagnoses by each institute for all 30 studied cases

Consensus diagnosis	Original diagnosis	
	Typical carcinoid (n = 19)	Atypical carcinoid (n = 11)
Typical carcinoid (n = 16)	14 (73.7)	2 (18.2)
Atypical carcinoid (n = 9)	5 (26.3)	4 (36.4)
Large cell neuroendocrine carcinoma (n = 2)	0	2 (18.2)
Small cell carcinoma (n = 2)	0	2 (18.2)
Organizing pneumonia (n = 1)	0	1 (9.0)

Values are presented as number (%).

Table 4. Agreement of the consensus diagnosis for all 30 cases studied

Consensus diagnosis	Agreement		
	Unanimous	Majority	Consensus
Total cases (n = 30)	15 (50)	25 (83.3)	30 (100)
Typical carcinoid (n = 16)	11 (68.8)	15 (93.8)	16 (100)
Atypical carcinoid (n = 9)	3 (33.3)	7 (77.8)	9 (100)
Large cell neuroendocrine carcinoma (n = 2)	0 (0)	1 (50)	2 (100)
Small cell carcinoma (n = 2)	0 (0)	1 (50)	2 (100)
Organizing pneumonia (n = 1)	1 (100)		

Values are presented as number (%).

occurred for 11 (68.8%) of 16 TCs and for 3 (33.3%) of 9 ACs. A majority agreement occurred for 15 (93.8%) of 16 TCs and for 7 (77.8%) of 9 ACs. Agreement was greater for the TC than that for the AC (Table 4).

DISCUSSION

In the present study, it was difficult to reach a unanimous agreement, which was achieved in only 50% of the 30 studied cases. Considering agreement on the mitotic criterion, unanimous agreement was achieved for 16 of 30 cases (53.3%), whereas for the necrosis criterion, unanimous agreement was achieved for 25 of 30 cases (83.3%). The mitotic count varied more substantially than did that for necrosis among the pathologists, and so the main source of disagreement, in the present study, was the variability of the mitotic count. Areas of mitotic activity may be focal or they may vary from section to section. Hence, it is important to count mitoses in the most active areas. The major cause for discordance was the difficulty to differentiate mitotic figures from pyknotic apoptosis. The identification of mitotic figures

Table 3. Demographic and tumoral characteristics of the patients with typical and atypical carcinoid tumors

Factors	Consensus diagnosis	
	Typical carcinoid (n = 16)	Atypical carcinoid (n = 9)
Age (yr)		
Mean	50.6	47.8
Range	30-69	23-70
Gender		
Male	10 (62.5)	5 (55.6)
Female	6 (37.5)	4 (44.4)
Size (cm)		
Mean	2.4	3.5
Range	1.4-7.0	0.6-6.0
Lymph node metastasis	2 (12.5)	5 (55.6)
Stage		
I	14 (87.5)	4 (44.5)
II	2 (12.5)	2 (22.2)
III	0 (0)	3 (33.3)

Values are presented as number (%).

could be facilitated by the use of mitosis-specific staining or labeling techniques. Immunohistochemical labeling of the mitotic figures using mitosis-specific anti-phosphohistone H3 (PHH3) has been suggested as a promising method for identifying mitotic figures.[8,9] Further research is currently being done by our study team to determine whether immunostaining with PHH3 antibody as an adjunctive tool increases the ability of pathologists to make an accurate diagnosis of a pulmonary carcinoid tumor.

A previous study about the interobserver agreement for making the diagnosis of pulmonary NE tumors among experienced pulmonary pathologists revealed that most of the kappa values fell into the substantial agreement category.[4] However, in the present study, four out of the 10 values fell into the moderate or fair agreement category. Only two pulmonary pathologists participated in our study. As one might expect, the highest level of agreement was between the two pulmonary pathologists. However, the variability was not affected by the pathologists' experience. The level of agreement was higher between two pathologists who worked closely together than between those pathologists who did not work closely together. A parallel rating of cases by the pathologists, followed by a cooperative examination of the discordant cases at the consensus meeting, proved to be a simple and efficient means of building consensus between the observers. A centralized consensus review is therefore mandatory for adequate standardization of a pulmonary carcinoid tumor diagnosis.

The recognition of AC is important because of the prognostic significance when compared with that for the low-grade TC

and the high-grade LCNEC and SCLC.[10,11] The main differential diagnosis that arises with AC is the separation of this tumor from other NE tumors. In the present study, of the originally diagnosed cases of AC, only 4 (36.4%) were retained as AC, whereas 2 cases each (each 2 cases was 18.2%) were changed to TC, LCNEC and SCLC after the consensus meeting. We also found that unanimous and majority agreement occurred for 11 (68.8%) and 15 (93.8%) of the 16 TCs, respectively, whereas unanimous and majority agreement occurred for 3 (33.3%) and 7 (77.8%) of the 9 ACs, respectively. Travis *et al*.[4] reported that interobserver reproducibility is the greatest for TC and SCLC, it is less for AC, and it is lowest for LCNEC. This is not surprising because the diagnosis of pulmonary NE tumors generally reflects the historical order in which they were originally described and also the rarity of AC and LCNEC.[4]

There was disagreement between the original diagnosis by each institute and a consensus diagnosis in 40.0% of the cases. What is worse, one originally diagnosed case of AC was revised as OP after the consensus meeting. It may have been partly because the lesion was localized and composed of obliterated airspaces lined by proliferating reparative type II pneumocytes. The incidence of pulmonary carcinoid is low in general pulmonary practice. A lower prevalence of disease results in a higher number of false positive and false negative diagnoses. The substantial reproducibility observed in this study will likely not exist among general practicing pathologists, because this study was performed by pathologists who were preoccupied by the study object and were well acquainted with the diagnostic criteria. It is apparent from this study that more attention needs to be given to carefully diagnosing pulmonary carcinoid tumors.

In conclusion, discriminating carcinoid tumors is subject to interobserver variability. A centralized consensus review and education are needed to increase the diagnostic concordance rates for making the diagnosis of pulmonary carcinoid tumor among surgical pathologists. This effort may help avoid interobserver diagnostic disagreement and would have no substantial negative impact on patient care.

REFERENCES

1. Cardillo G, Sera F, Di Martino M, *et al*. Bronchial carcinoid tumors: nodal status and long-term survival after resection. Ann Thorac Surg 2004; 77: 1781-5.
2. Travis WD, Colby TV, Corrin B, *et al*. World Health Organization international histological classification of tumours: histological typing of lung and pleural tumours. 3rd ed. Berlin: Springer, 1999.
3. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. World Health Organization classification of tumours: pathology and genetics of tumours of the lung, pleura, thymus and heart. Lyon: IARC press, 2004.
4. Travis WD, Gal AA, Colby TV, Klimstra DS, Falk R, Koss MN. Reproducibility of neuroendocrine lung tumor classification. Hum Pathol 1998; 29: 272-9.
5. Mountain CF. Revisions in the international system for staging lung cancer. Chest 1997; 111: 1710-7.
6. Travis WD, Rush W, Flieder DB, *et al*. Survival analysis of 200 pulmonary neuroendocrine tumors with clarification of criteria for atypical carcinoid and its separation from typical carcinoid. Am J Surg Pathol 1998; 22: 934-44.
7. Baak JP. Mitosis counting in tumors. Hum Pathol 1990; 21: 683-5.
8. Shibata K, Inagaki M, Ajiro K. Mitosis-specific histone H3 phosphorylation in vitro in nucleosome structures. Eur J Biochem 1990; 192: 87-93.
9. Tapia C, Kutzner H, Mentzel T, Savic S, Baumhoer D, Glatz K. Two mitosis-specific antibodies, MPM-2 and phospho-histone H3 (Ser28), allow rapid and precise determination of mitotic activity. Am J Surg Pathol 2006; 30: 83-9.
10. Beasley MB, Thunnissen FB, Brambilla E, *et al*. Pulmonary atypical carcinoid: predictors of survival in 106 cases. Hum Pathol 2000; 31: 1255-65.
11. García-Yuste M, Matilla JM, Cueto A, *et al*. Typical and atypical carcinoid tumours: analysis of the experience of the Spanish Multi-centric Study of Neuroendocrine Tumours of the Lung. Eur J Cardiothorac Surg 2007; 31: 192-7.