

Tumor Size Evaluation according to the T Component of the Seventh Edition of the International Association for the Study of Lung Cancer's TNM Classification: Interobserver Agreement between Radiologists and Computer-Aided Diagnosis System in Patients with Lung Cancer¹

International Association for the Study of Lung Cancer의 7차 개정판 TNM 분류의 T 요소에서 종양 크기의 평가: 폐암 환자에서 영상의학과 의사와 Computer-Aided Diagnosis System의 일치도¹

Jin Kyoung Kim, MD¹, Semin Chong, MD¹, Jae Seung Seo, MD^{1,2}, Sun Jin Lee, MD¹, Heon Han, MD²

¹Department of Radiology, Chung-Ang University College of Medicine, Seoul, Korea

²Department of Radiology, Kangwon National University School of Medicine, Chuncheon, Korea

Purpose: To assess the interobserver agreement for tumor size evaluation between radiologists and the computer-aided diagnosis (CAD) system based on the 7th edition of the TNM classification by the International Association for the Study of Lung Cancer in patients with lung cancer.

Materials and Methods: We evaluated 20 patients who underwent a lobectomy or pneumonectomy for primary lung cancer. The maximum diameter of each primary tumor was measured by two radiologists and a CAD system on CT, and was staged based on the 7th edition of the TNM classification. The CT size and T-staging of the primary tumors was compared with the pathologic size and staging and the variability in the sizes and T stages of primary tumors was statistically analyzed between each radiologist's measurement or CAD estimation and the pathologic results.

Results: There was no statistically significant interobserver difference for the CT size among the two radiologists, between pathologic and CT size estimated by the radiologists, and between pathologic and CT staging by the radiologists and CAD system. However, there was a statistically significant interobserver difference between pathologic size and the CT size estimated by the CAD system ($p = 0.003$).

Conclusion: No significant differences were found in the measurement of tumor size among radiologists or in the assessment of T-staging by radiologists and the CAD system.

Index terms

Neoplasm Staging
Tomography, Spiral Computed
Carcinoma, Non-Small-Cell Lung
Observer Variation

Received June 7, 2011; Accepted August 9, 2011

Corresponding author: Semin Chong, MD
Department of Radiology, Chung-Ang University
Medical Center, Chung-Ang University College of
Medicine, 102 Heukseok-ro, Dongjak-gu,
Seoul 156-755, Korea.
Tel. 82-2-6299-1288 Fax. 82-2-6918-6229
E-mail: j3mn@chol.com

Copyrights © 2011 The Korean Society of Radiology

INTRODUCTION

According to the seventh edition TNM staging system for non-small cell lung cancer (NSCLC), the T component was divided into T1a, T1b, T2a, T2b, and T3, depending on tumor size cut-offs of 2, 3, 5 and 7 cm because of significant differ-

ences in the five-year survival rates of patients with lung cancer, which are 53%, 47%, 43%, 36% and 26%, respectively (1). Therefore, accurate assessment of tumor size and staging is essential for patient management and prognosis prediction.

Tumor size is affected by the morphology and location of the tumor, as irregular tumors and those in a juxtavascular lo-

cation are subject to the greatest degree of variability (2). In addition, consistent measurement of tumor size is necessary in the evaluation of chemotherapy response (3). Computer-aided diagnosis (CAD) can be used to complement the limitations of manual measurements when conducting serial CT examinations because measurements taken using CAD software programs are more accurate and reproducible than manual measurements (4, 5).

In our study, we evaluated tumor size and staging based on the new seventh edition T-staging criteria. To the best of our knowledge, there has been no reported study evaluating the interobserver agreement of the T component with reference to the new seventh edition TNM classification. We evaluated the clinical tumor size of the T component in TNM staging measured manually or by CAD of CT scans, as compared to the pathologic tumor size and T staging after surgical resection in patients with NSCLC.

The purpose of this study was to assess the interobserver agreement of T-staging based on the seventh edition of TNM classification published by the International Association for the Study of Lung Cancer (IASLC) in patients with lung cancer based on preoperative CT scans.

MATERIALS AND METHODS

Twenty patients (M : F = 14 : 6; age range, 45-77 years; mean age, 62.6 years) who had undergone a lobectomy ($n = 18$) or pneumonectomy ($n = 2$) for primary lung cancer (8 squamous cell carcinomas, 10 adenocarcinomas, 1 large cell carcinoma, and 1 bronchioloalveolar carcinoma) at our institution between February 2009 and June 2011 were included in this study. Their CT and pathologic findings were retrospectively reviewed.

Chest CT scans were obtained using a 256-slice multidetector CT scanner (Brilliance iCT, Philips Healthcare, Andover, MA, USA), from the lung apices to the level of the middle portion of both kidneys. The scanning parameters were 120 kVp and 125 mA, with a beam width of 10-20 mm, a beam pitch of 0.915, and a reconstruction thickness of 2.5 mm. CT scanning was performed after the injection of 150 mL of iopamidol, an iodinated contrast medium (Pamiray 300; Dongkook Pharm. Co., Ltd., Seoul, Korea) at a rate of 2 mL/sec with a power injector (Stellant; MEDRAD, Inc., Indianola, PA, USA). Image

data were reconstructed with a soft tissue algorithm for mediastinal window images and with a bone algorithm for lung window images. The reconstructed images were directly interfaced with a picture archiving and communication system (Marosis; Infinitt Co., Ltd., Seoul, Korea), which displayed all image data on two monitors (1,536 × 2,048 matrix, 8-bit viewable gray scale and 60-ft-lambert luminescence). All images were reviewed on both mediastinal (width: 360 HU, level: 60 HU) and lung (width: 1,500 HU, level: -700 HU) windows.

We evaluated the location (central or peripheral; lobar segmentation) and morphology (smooth, lobulated or irregular) of each primary tumor. The maximum diameter of the primary tumor was measured by two thoracic radiologists with seven and nine years of experience, respectively, and by a CAD system (Extended Brilliance Workspace; Philips Health Care; Andover, MA, USA) and categorized according to the seventh edition TNM classification (6). In cases of manual measurement, the maximum diameter of each tumor was measured on an axial and lung window image. The observers were blinded to the clinical details and outcomes and evaluated the preoperative CT images independently. In cases of CAD measurement, all reconstructed thin-section (1-mm) CT images were transferred to a workstation, and the 2D diameter and 3D volumetric measurements were obtained using the lung nodule assessment application of the commercially available CAD system. We used an axial slab image as the main viewport, and coronal and sagittal image as the reference image. After our inspection indicated the existence of a nodule, we used the mark nodule tool to automatically segment the nodule. On the mark nodule tool, the pointer of a pencil-shaped cursor was clicked on the nodule, which was automatically accepted. Next, a volume-rendered image of the nodule and lesion measurement table including maximum diameter, was provided. We calculated the CT scanning interval, which was defined as the time interval from CT scanning to resection of the primary lung cancer.

All references were based on the pathologic tumor size and staging. The CT sizes and staging of the primary tumors were estimated by radiologists or by the CAD system and compared as a function of pathologic sizes and staging. The pathologic tumor size and staging was evaluated by two pathologists with twelve and four years of experience, respectively. The surgical specimens were all inflated transbronchially with

neutrally buffered formalin until the pleura was smooth. After overnight fixation, the lungs were sliced at 1-cm intervals in the parasagittal plane. The maximal superior-inferior, anteroposterior and mediolateral macroscopic dimensions of the primary lesion were documented by the pathologists. Of these three-dimensional diameters, the largest one was considered to be the reference tumor size.

Statistical analyses were performed with a statistical software (PASW Statistics 18.0; IBM Corporation; Armonk, NY, USA), and *p*-values less than 0.05 were considered to be statistically significant. We evaluated the interobserver difference and agreement by the Wilcoxon signed-rank test and the Spearman's rho correlation coefficient. The linear regression model was used in evaluating factors associated with the CT size of the primary tumors. CT T-staging was statistically evaluated by the marginal homogeneity test and Spearman's rho correlation coefficient.

RESULTS

Table 1 summarizes tumor characteristics, sizes, and stag-

ing in 20 patients with lung cancer. The morphology of all tumors was irregular. The maximum diameters of the primary tumors were 2.8 ± 1.4 cm, 2.9 ± 1.6 cm, 3.0 ± 1.6 cm, 3.4 ± 1.1 cm, as measured by pathological examination, radiologic observer 1, radiologic observer 2, and the CAD system, respectively. The pathologic staging of 20 non-small cell lung cancers included stage 1a (*n* = 6), stage 1b (*n* = 6), stage 2a (*n* = 6) and stage 2b (*n* = 2). The time interval from CT scanning to pathologic analysis ranged from 3 to 32 days (mean interval, 16.1 days).

There was statistically significant interobserver agreement for CT size between the two radiologists (*p* = 0.984) and between the pathologic size and the CT size estimated by the manual measurement (radiologist 1, *p* = 0.679; radiologist 2, *p* = 0.558). There was a high correlation of CT size measurement between the two radiologists (Fig. 1), and a statistically significant interobserver difference between each radiologist and the CAD system (radiologist 1, *p* = 0.005; radiologist 2, *p* = 0.006), and between the pathologic size and the CT size estimated by the CAD system (*p* = 0.003). The correlation of CT size measurement in CAD was lower than in radiologist

Table 1. Tumor Characteristics, Size and Staging in 20 Patients with Lung Cancers

No	Sex	Age	Pathology	Location	Tumor Size (cm)				T Staging				
					Pathologic	Radiologist 1	Radiologist 2	CAD	Pathologic	Radiologist 1	Radiologist 2	CAD	
1	M	77	AD	RUL	Peripheral	0.8	2.1	2.2	2.3	1a	1b	1b	1b
2	M	57	AD	RLL	Peripheral	1.2	1.6	1.3	1.7	1a	1a	1a	1a
3	M	66	SQ	LUL	Peripheral	1.2	1.0	1.2	1.5	1a	1a	1a	1a
4	M	47	AD	RML	Peripheral	1.6	1.7	1.5	2.7	1a	1a	1a	1b
5	F	61	AD	RLL	Peripheral	1.7	1.7	1.6	2.5	1a	1a	1a	1b
6	M	67	AD	LUL	Peripheral	2.0	2.2	4.2	3.9	1a	1b	2a	2a
7	M	74	AD	LUL	Peripheral	2.2	1.9	2.3	2.9	1b	1a	1b	1b
8	M	56	AD	LLL	Peripheral	2.2	2.2	2.8	2.7	1b	1b	1b	1b
9	M	56	SQ	LUL	Central	2.4	2.3	1.9	3.7	1b	1b	1a	2a
10	F	69	AD	LUL	Peripheral	2.4	2.1	2.2	3.2	1b	1b	1b	2a
11	F	48	AD	LLL	Peripheral	2.6	2.0	2.3	2.9	1b	1a	1b	1b
12	F	45	BAC	RUL	Peripheral	2.7	2.6	2.4	2.8	1b	1b	1b	1b
13	M	64	SQ	LUL	Peripheral	3.1	2.7	2.4	2.9	2a	1b	1b	1b
14	M	68	LC	RUL	Peripheral	3.5	4.3	4.2	4.7	2a	2a	2a	2a
15	M	67	SQ	RLL	Peripheral	3.6	3.1	3.2	4.0	2a	2a	2a	2a
16	F	65	SQ	RUL	Peripheral	3.8	4.1	3.9	3.5	2a	2a	2a	2a
17	M	59	SQ	LUL	Central	3.9	4.4	4.6	5.0	2a	2a	2a	2a
18	F	66	AD	RLL	Central	4.1	4.5	4.4	4.8	2a	2a	2a	2a
19	M	66	SQ	LUL	Peripheral	5.5	3.8	3.5	4.5	2b	2a	2a	2a
20	M	75	SQ	RLL	Peripheral	6.2	8.2	8.2	5.8	2b	3	3	2b

Note.—AD = adenocarcinoma, BAC = bronchioloalveolar carcinoma, CAD = computer-aided diagnosis, LC = large cell carcinoma, LLL = left lower lobe, LUL = left upper lobe, RLL = right lower lobe, RML = right middle lobe, RUL = right upper lobe, SQ = squamous cell carcinoma

1, but higher than in radiologist 2 (Fig. 1). However, there was no statistically significant interobserver difference between pathologic and CT staging by manual measurement (radiologist 1, $p = 0.705$; radiologist 2, $p = 0.739$) using the CAD system ($p = 0.132$). Table 2 summarizes the correlation coefficient of tumor size measurement and T-staging between observers and pathology. The location of the primary tumors was peripheral ($n = 17$) or central ($n = 3$), but could not be statistically analyzed on interobserver variability because of the very small sample size.

DISCUSSION

Tumor size measurement is very important for the accurate staging of non-small cell lung cancers, determination of management methods, and evaluation of the response to chemotherapy (7). In the seventh TNM classification, the T1 descriptor was divided into T1a (≤ 2 cm) and T1b (≥ 2 -3 cm), and the T2 descriptor into T2a (> 3 -5 cm) and T2b (> 5 -7 cm) (8). Further sub-classifications depending on the tumor size can increase the impact on staging in patients with lung

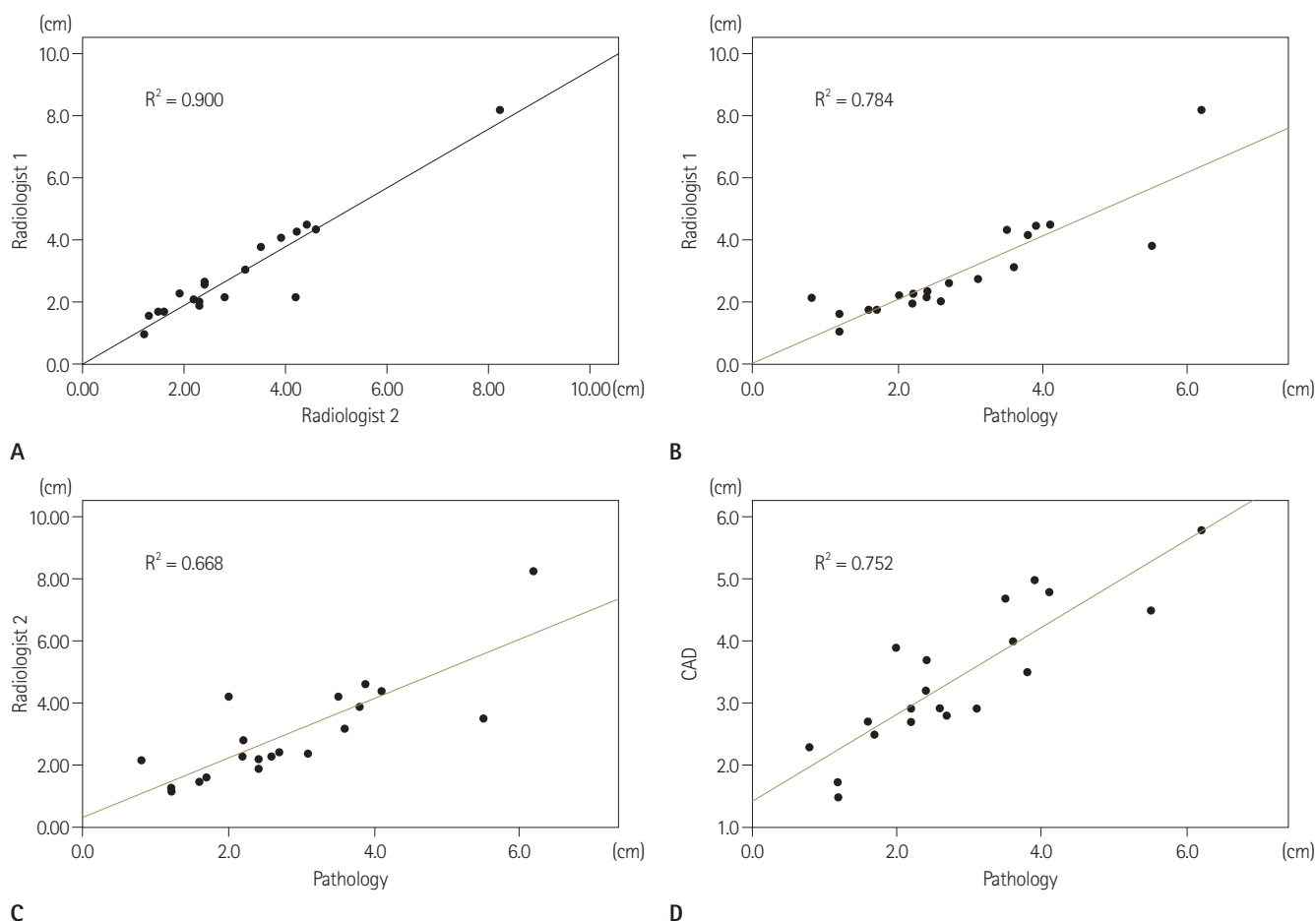


Fig. 1. Comparison of tumor sizes between two radiologists (A), radiologist 1 and pathology (B), radiologist 2 and pathology (C), and CAD and pathology (D).

Note.—CAD = computer-aided diagnosis

Table 2. Correlation Coefficient of Tumor Size Measurement and T Staging

	Pathology		Radiologist 1		Radiologist 2	
	Size	T Staging	Size	T Staging	Size	T Staging
Radiologist 1	0.903	0.834				
Radiologist 2	0.811	0.763	0.905	0.873		
CAD	0.866	0.689	0.877	0.845	0.862	0.782

Note.—Correlation is significant at the 0.01 level.

CAD = computer-aided diagnosis

cancer (1). No study of interobserver variability of the T component based on the new seventh TNM classification has been reported. We assessed the influence of interobserver variability on the sub-classification of the new T staging based on the seventh edition of the TNM classification published by the IASLC in patients with lung cancer.

A number of studies have investigated interobserver variations in tumor size estimation by manual measurement (3, 9, 10). In a CT variability study of tumor size, Macpherson et al. (9) reported significant interobserver variation in defining gross tumor volume. In contrast, Wormanns et al. (10) reported a good correlation in the assessment of lesion size on CT scans between two radiologists. In our study, high interobserver agreement was observed in the manual measurement of the maximum tumor size using CT. Nevertheless, it is still required to complement the limitations of measurement differences resulting from manual measurement.

CAD can be used to consistently measure tumor size because measurements made by CAD systems are more accurate and reproducible than manual measurements in the assessment of lesion size (4, 5). In a CT study of nodule characteristics, Iwano et al. (4) reported that there was significant interobserver variability among radiologists for manual, but not CAD, measurements of maximal diameter. In a treatment response study using CT, Marten et al. (5) concluded that automated volumetry using CAD allowed for better reproducibility of response evaluations and should be preferred over manual unidimensional measurements. In our study, however, there was a statistically significant difference in the measurement of tumor size using a CAD system compared to the pathologic tumor size. This system causes errors in three-dimensional manual measurements such that radiologists can be overwhelmed by large thin-section CT data including axial, sagittal, and coronal images. Whereas, unidimensional measurements do not cause errors because the relatively small amount of axial data can decrease interobserver variation. Fortunately, there was no statistically significant difference in lung cancer staging based on the CAD system. Based on our findings, we suggest that improvement may be required for the clinical application of automatic measurement using CAD systems.

In variability and agreement regarding software-generated volume measurement, Wang et al. (2) reported that nodule

morphology, location, and size influenced volume measurement variability, particularly for juxtavascular and irregular nodules. However, Nishino et al. (3) reported that location, morphology, or adjacent atelectasis had no significant impact on inter- or intraobserver variability. In our study, we could not evaluate the effect of nodule morphology on the variability of size measurements because all nodules had irregular shapes and nodule location could not be statistically analyzed on the basis of measurement variability because of the very small sample size.

Our study had some limitations. First, the sample size was small and diverse tumor sizes were not included. Second, we measured the maximum diameters of tumors on unidimensional axial CT images. However, this is unlikely to affect measurement errors and observer agreement because the greatest diameters of the tumors were also measured on axial CT images in most previous studies (5, 11). Third, tumor size was measured by a CAD system after we defined the nodule attenuation using a region of interest. In a future study, nodule detection and the measurement of nodule size could be completely automated by improvement of the CAD software.

In conclusion, the T component of the lung cancer staging system was sub-classified by variable size criteria with divisions at 2 cm, 3 cm, 5 cm, and 7 cm in the seventh edition of the TNM staging system published by the IASLC. Fortunately, in our present study, there were no differences in the measurements of tumor size among radiologists and in the assessment of T-staging by radiologists and a CAD system. However, we think that many more sample sizes, including multiple tumor stages, should be studied, and the improvement of CAD systems may be required in order to apply automatic measurements in clinical practice.

REFERENCES

1. Tsim S, O'Dowd CA, Milroy R, Davidson S. Staging of non-small cell lung cancer (NSCLC): a review. *Respir Med* 2010; 104:1767-1774
2. Wang Y, van Klaveren RJ, van der Zaag-Loonen HJ, de Bock GH, Gietema HA, Xu DM, et al. Effect of nodule characteristics on variability of semiautomated volume measurements in pulmonary nodules detected in a lung can-

- cer screening program. *Radiology* 2008;248:625-631
3. Nishino M, Guo M, Jackman DM, DiPiro PJ, Yap JT, Ho TK, et al. CT tumor volume measurement in advanced non-small-cell lung cancer: performance characteristics of an emerging clinical tool. *Acad Radiol* 2011;18:54-62
4. Iwano S, Okada T, Koike W, Matsuo K, Toya R, Yamazaki M, et al. Semi-automatic volumetric measurement of lung cancer using multi-detector CT effects of nodule characteristics. *Acad Radiol* 2009;16:1179-1186
5. Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C. Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. *Eur Radiol* 2006;16:781-790
6. Rusch VW, Asamura H, Watanabe H, Giroux DJ, Rami-Porta R, Goldstraw P; Members of IASLC Staging Committee. The IASLC lung cancer staging project: a proposal for a new international lymph node map in the forthcoming seventh edition of the TNM classification for lung cancer. *J Thorac Oncol* 2009;4:568-577
7. Lee HY, Lee KS, Hwang HS, Lee JW, Ahn MJ, Park K, et al. Molecularly targeted therapy using bevacizumab for non-small cell lung cancer: a pilot study for the new CT response criteria. *Korean J Radiol* 2010;11:618-626
8. Tanoue LT, Detterbeck FC. New TNM classification for non-small-cell lung cancer. *Expert Rev Anticancer Ther* 2009;9:413-423
9. Macpherson RE, Higgins GS, Murchison JT, Wallace WA, Price A, Gaffney S, et al. Non-small-cell lung cancer dimensions: CT-pathological correlation and interobserver variation. *Br J Radiol* 2009;82:421-425
10. Wormanns D, Diederich S, Lentschig MG, Winter F, Heindel W. Spiral CT of pulmonary nodules: interobserver variation in assessment of lesion size. *Eur Radiol* 2000;10:710-713
11. Zhao B, James LP, Moskowitz CS, Guo P, Ginsberg MS, Lefkowitz RA, et al. Evaluating variability in tumor measurements from same-day repeat CT scans of patients with non-small cell lung cancer. *Radiology* 2009;252:263-272

International Association for the Study of Lung Cancer의 7차 개정판 TNM 분류의 T 요소에서 종양 크기의 평가: 폐암 환자에서 영상의학과 의사와 Computer-Aided Diagnosis System의 일치도¹

김진경¹ · 정세민¹ · 서재승^{1,2} · 이선진¹ · 한 현²

목적: 본 연구는 폐암 환자에서 7차 개정판 TNM 분류에 근거한 원발 종양 크기 평가에 있어서 영상의학과 의사와 computer-aided diagnosis (이하 CAD) system의 일치도를 평가하였다.

대상과 방법: 폐암으로 진단되어 엽절제 또는 폐절제를 시행한 20명의 환자를 대상으로 2명의 영상의학과 의사와 CAD system을 이용하여 CT에서 원발 종양의 최대 직경을 측정하였고 7차 개정판 TNM 분류에 따라 T 병기를 결정하였다. 원발 종양의 CT 크기와 T 병기가 각각 병리 크기와 병기와 비교되었고 영상의학과 의사, CAD System 그리고 병리 결과 사이의 원발 종양 크기와 T 병기에 대한 차이를 통계학적으로 분석하였다.

결과: 2명의 영상의학과 의사 사이에서 측정된 원발 종양의 CT 크기, 영상의학과 의사에 의해 측정된 원발 종양의 CT 크기와 병리 크기 및 영상의학과 의사와 CAD를 이용하여 결정된 원발 종양의 CT 병기와 병리 병기 사이에 유의한 차이는 없었다. 그러나 CAD로 측정된 원발 종양의 CT 크기와 병리 크기 사이에 유의한 차이가 있었다($p = 0.003$).

결론: 원발 종양의 크기는 영상의학과 의사 사이에 통계적으로 유의한 차이는 없었고 영상의학과 의사와 CAD에 의한 원발 종양의 CT 병기와 병리 병기 사이에 유의한 차이는 없었다.

¹중앙대학교 의과대학 영상의학과의학교실, ²강원대학교 의과대학 영상의학과의학교실