250

Available online at www.sciencedirect.com

SCIENCE DIRECT

ELSEVIER

JNMU

Journal of Nanjing Medical University, 2008, 22(4):250-254

www.elsevier.com/locate/jnmu

Research Paper

Research on consistency of identifying solitary pulmonary masses with CT

Qiuping Wang[a], Gang Niu[a], Yun Zhang[a], Yongqian Qiang[a], Zicheng Li[b], Youmin Guo[c,*]

[a] **The imaging Center, the First Affiliated Hospital, medical school of Xi'an Jiaotong University, Xi'an 710061, Shaanxi Province, China;** [b]**Imaging Center, Shen-mu County Hospital, Shen-mu 719300, Shaanxi Province,China**

[c]**The imaging Center, the Second Affiliated Hospital, medical school of Xi'an Jiaotong University, Xi'an 710004, Shaanxi Province, China**

Received 3 April, 2008

Abstract

Objective:To research on consistency of identifying solitary pulmonary masses with CT. Methods:Three observers with different working backgrounds in imaging diagnosis individually interpreted the same group images of solitary pulmonary mass, by 12 indexes of objective signs. The differences in interpretation resulted in ante- and post-interpretations were assessed by the x^2 test. The agreement of two interpretations from the same observer was confirmed with the kappa test. A double-blind method was adopted for analysis. Results:The agreement rates of ante- and post-interpreting from the three observers were respectively 82.65%(486/588) 80.27%(472/588) and 84.86% (499/588) while their interpreting results were generally accordant without significant difference(x^2 = 4.975, **df** = 2, P = 0.083) however there was difference between the observer 2 and observer 3(x^2 = 4.875, **df** = 1, **P** = 0.027). There were five indexes with **k** > 0.40 of ante- and post-interpreting results of the three observers, including clarity of nodule borderline, presence of sentus, uniformity of density, existence of cavity and calcification in pathological region, among them, the agreement rate of interpreting borderline and cavity was higher(**k** > 0.07); the blood vessel convergence poorer(0 < **k** ≤ 0.40); the other six CT signs of interpretation were slightly different. Conclusion:The ability to identify solitary pulmonary mass was inconsistent, and needs to be improved further.

Key words: pulmonary mass; CT sign; intro-observer; inter-observer; kappa index

INTRODUCTION

CT has become a powerful tool in examining pulmonary disease[1-3]. Identifying the characters of solitary pulmonary mass is key to the differential diagnosis of benign and malignant diseases. Detection of signs in CT plays a significant role in determining nature of disease[4-7]. Nowadays, imaging diagnosticians are asked to work out diagnosis and finish reports of CT within two hours in domestic hospitals. However we haven't seen anything clear in literature about how diagnostician's consistency fairs in interpreting CT. So, to investigate this we analyzed diagnosticians with different working backgrounds and the same diagnostician's reading of the CT signs of the pulmonary mass in the same group at a different time.

MATERIALS AND METHODS

Image-reader

Three image-readers from different hospitals voluntarily participated in interpreting the signs of solitary pulmonary mass. Observer 1 had been working as a vice chief physician at chest imaging diagnosis for 24 years in class 3A Hospitals; Observer 2 had been working as a vice chief physician in imaging diagnosis for 24 years in class 2A Hospitals; Observer 3 had been working as an attending physician at chest imaging diagnosis for 9 years in class 3A Hospitals. Three observers were working as usual without receiving normalization training about reading solitary pulmonary mass images.

*Corresponding auther.

E-mail address: cjr.guoyoumin@vip.163.com

Instrumentations

CT scanner was Prospeed spiral CT scanner purchased from GE Company with 5 mm, 10 mm of thickness of slice image. The film was printed by the solid state infrared laser camera. Q/SYYC003-92 viewer was purchased by Yuhua medical apparatus and instruments factory in Shantou.

Data

Eligible standard :① CT images obtained from the patients who were pathologically diagnosed; ②All images and the data were taken within 10 days from when the samples underwent surgery and the paracentesis were pathologically diagnosed; ③ Pulmonary nodule was less than 5 cm in diameter; ④ No mediastinal lymphadenovarix; ⑤The CT image slice was 5mm and 10mm in thickness. *Excluded standard*: ①This included severe constructed defects that affected the observation; ②Other pulmonary diseases influenced the observation; ③ During the course of reading CT image, the image slice was lost and damaged, for this reason, three observers were not able to make a diagnosis. There were 124 cases of CT image slices of solitary pulmonary masses confirmed in pathologically diagnosis, among them, 54 cases were eligible with five excluded(leaving 49 cases for the published result).

Methods

A double blind method was constructed. Assistants were appointed to randomly number the collected CT image slices and to tabulate the data; including 12 indexes with single-choice questions. The CT image slices and tables were then handed over to the three observers(each observer didn't know the other participants). Without any clinical information, the observer interpreted independently 49 cases of CT slices in the same group, and read each CT slices in detail under viewer, including the judgment of borderline, outline, sublobe, spinous process, sentus, density, vacuole sign, air bronchogram, cavity, calcification, vascular convergence sign, and pleural indentation sign. Finally the results were transferred to the prepared table. The reading time was unlimited. The CT image slices and the tables were taken back once the reading was finished. The process above was repeated at the interval of 20 days. Another non-observer checked each item in the table and picked out the table which had missing and inconsistency items in the table, and then those tables, along with CT slices, were returned to original observer who re-judged them. According to the requirements, two participants input the data in all tables into the previous prepared database.

Statistical analysis

All data from the three observers were statistically processed with SPSS12.0. The agreement rate of the two observations among three observers was performed with χ^2 test. The interpretation of CT image slices of solitary pulmonary masses in the same group and the assessment of agreement of two observations were evaluated with kappa value(k).

Judgment of results[8]: That k value is set between +1 and-1. $k \leqslant 0$ showed the result agreement was less than chance agreement between two observation. $0<k\leqslant 0.40$ indicated poor of agreement of two observation; $0.40 < k \leqslant 0.60$ moderate; $0.60 < k \leqslant 0.80$ higher; $k > 0.80$ excellent; $P < 0.05$ hinted a difference in statistics; $P < 0.01$ significant difference.

RESULTS

Twelve signs of each case in 49 cases of solitary pulmonary masses were individually interpreted twice. The agreement rate of two observations of three observers was respectively 82.65%(486/588) 80.27%(472/588) and 84.86%(499/588); the final agreement rate of three observers was 82.60%(1457/1764). Although there were individual differences in the interpretation of three observers(see Table 1 in detail) there was no statistical difference(χ^2 = 4.317, **P** = 0.116).

Table 1 The χ^2 test of agreement on reading 12 common CT signs of three observers

Observers	agreement	disagreement	χ^2	df	*P* (two-tailed)
Observer 1	486(82.65%)	102(17.35%)	1.056[a]	1	0.304
Observer 2	472(80.27%)	116(19.73%)	1.104[b]	1	0.293
Observer 3	499(84.86%)	89(15.14%)	4.307[c]	1	0.038
sum	1457(82.60%)	307(17.40%)	4.314[d]	2	0.116

a:χ^2 value of agreement of two interpretations between observer 1 and 3; b:χ^2 value of agreement of two interpretations between observer 1 and 2; c:χ^2 value of agreement of two interpretations between observer 2 and 3; d:χ^2 value of agreement of two interpretations between ante-and post-interpretation of three observers.

In 12 CT signs of solitary pulmonary mass, among three observers, there were 2 indexes with k >0.70 in two observations: clarity of borderline and presence of cavity; 3 indexes with $0.40 < k \leqslant 0.60$:sentus, density and calcification inside. As far agreement of three observations in two interpretations, there was only one index with $k \leqslant 0.40$: the interpretation of vascular convergence signs; There were 6 indexes with $k > 0.40$ in at least one observer(see Table 2).

Table 2 The results of interpreting CT signs in two reading of three observers

First interpretation		Second interpretation											
		observer 1				observer 2				observer 3			
		yes	no	k	p	yes	no	k	p	yes	no	k	p
Borderline clarity	yes	43	1	0.777	0.000	36	0	0.702	0.000	34	1	0.791	0.000
	no	1	4			5	8			3	11		
Outline smooth	yes	5	1	0.668	0.000	35	3	0.128	0.321	11	4	0.792	0.000
	no	3	40			9	2			0	34		
Sublobe	yes	34	2	0.463	0.001	17	6	0.273	0.048	14	6	0.381	0.007
	no	7	6			12	14			9	20		
Spinous process	yes	4	3	0.235	0.071	0	2	-0.072	0.493	19	4	0.632	0.000
	no	10	32			9	38			5	21		
Sentus	yes	22	5	0.630	0.000	11	5	0.461	0.001	16	2	0.703	0.000
	no	4	18			7	26			5	26		
Density	yes	31	6	0.478	0.001	33	8	0.709	0.000	21	7	0.593	0.000
	no	4	8			2	6			3	18		
Vacuole sign	yes	0	3	-0.052	0.712	1	4	0.109	0.445	7	2	0.787	0.000
	no	2	44			4	40			1	39		
Air bronchogram	yes	2	2	0.335	0.016	0	0	0.000		5	5	0.505	0.000
	no	4	41			7	42			2	37		
Cavity	yes	2	1	0.790	0.000	2	0	0.790	0.000	2	0	0.790	0.000
	no	0	46			1	46			1	46		
Calcification	yes	1	0	1.000	0.000	1	1	0.657	0.000	2	1	0.459	0.001
	no	0	48			0	47			3	43		
Vascular convergence	yes	12	0	0.317	0.002	2	0	0.285	0.004	18	14	0.288	0.028
	no	19	18			8	39			4	13		
Pleural indentation	yes	18	0	0.176	0.029	4	2	0.075	0.413	13	4	0.680	0.000
	no	24	7			21	22			3	29		
Sum	yes	174	24	0.636	0.000	142	31	0.565	0.000	162	50	0.668	0.000
	no	78	312			85	330			39	337		

mean k = 0.623, P = 0.000.

For three observers, the k value of reading results for 12 CT signs of all CT images in two observations respectively were 0.636, 0.565, 0.668, and mean 0.623(P = 0.000). There were ten with k > 0.40 in 12 indexes (83.33%) for observer 3; there were eight indexes(66.67%) with k > 0.60; there were five indexes (41.67%)with k > 0.40 in 12 indexes, for observer 2, there were four indexes(33.33%) with k > 0.60; the agreement of observer 1 was moderate: there were seven (58.33%) with k > 0.40 in 12 indexes and five indexes (41.67%)with k > 0.60.

DISCUSSION

Present status of identifying CT signs of solitary pulmonary mass

In the research, the agreement rates of two interpretations taken for 12 signs of 49 cases of solitary pulmonary masses were from 80.27% to 84.86%, average rate 82.60%(1457/1764) and k values ranging from 0.565 to 0.668(P = 0.000) average k value 0.623(P = 0.000). In spite of the same reading environment and situation, disagreement ranged from 15.14% to 19.73% (average 17.40%). The research of Wakeley[9] and Leslie *et al.*[10] on intra-observer difference revealed disagreement rates of interpreting MR and CT slices for intra-observers from 19.5% to 40%, which was accordant with the results in this paper.

The results showed the disagreement rate of observer 2 in two interpretations was the highest(19.73%)58.33% (7/12) CT signs in two interpretations disagreed(k < 0.40). In the course of interpretation, there was a lower agreement rate of the observer 2 than that of the other two observers, and there was the statistical difference of interpreting CT signs from observer 3(χ^2 = 4.307, **P** = 0.038). So for what reason caused the differences between observer 2 and the other observers? Observer 2 was a senior diagnostician on imaging in class 2A hospital, who was asked to diagnose all kinds of CT signs without subject classification. The observers 1 and 3 worked as diagnosticians for imaging in clinic, and also had duties of teaching and research of chest imaging diagnosis, in a class 3A hospital. With the requirement of teaching, they have a strong sense of mastering theory and standardizing interpretation of CT signs. Through continual training over a long time, they developed the permanent habit of reading, while observer 2 had a different work situation, working in a primary hospital, where the doctors and patients pay more attention to qualitative diagnosis, and less requirement for describing pathological changes which lead to a vague standard for diagnosis and causing a lack of knowledge in CT signs. Resulting from this, a poor stability in interpreting CT images appeared, so the agreement rate during two interpretations was lower. In addition, a lack of professional division in the primary hospital led to distraction of physician and poor knowledge, which was another cause of poor stability on interpreting CT signs of solitary pulmonary mass for observer 2.

Research on the stability of interpreting CT signs of solitary pulmonary

Among the common CT signs of solitary pulmonary masses described in the paper[11], the observers had a better agreement of two observations as to the clarity of the borderline and cavity inside($k > 0.700$) with better repeatability and stable recognization. Among them, there was the highest agreement rates of identifying of clarity of borderline, ranging from 89.80%(44/49) to 95.92%(47/49) mean 92.52%(136/147); the three observers had 97.96%(48/49) agreement on recognizing cavities; a moderate agreement on recognizing sentus and calcification($k > 0.450$) with 80.95%, 79.59%, and 96.60%.

As for vascular convergence sign, there were a poorer agreement with $k = 0.285$ to 0.317($P < 0.05$) of three observers in k value. In identifying borderline clarity or not, presence of sublobe and spinous process, vacuole sign and air bronchogram, pleural indentation sign, there were different agreements in the three observers. Among them, k value of spinous process, vacuole sign, air bronchogram, and pleural indentation sign was less than 0.100. The results described above predicted that three observers had an unstable ability to reach a consensus in identifying CT signs. The reasons were as follows: (1)Vascular convergence sign and vacuole sign were new signs after CT, specially HRCT was applied in clinic. Although spinous process, air bronchogram, and pleural indentation sign had been described little in basic sign of traditional chest images, there were differences of content and morphology from conventional chest radiography owing to application of CT in clinic [12-16]. (2)There are a lot of descriptions about these signs in literature, but a lack of normalization standard for their diagnosis, furthermore, more book based description and less practical instruction in reading standards. Imaging diagnosis physicians were so short of dependence that they read CT images vaguely with poor repetitiveness. For instance, one observer confused masses-adhered pleura with pleural indentation signs[6,7,17,18]. (3)Observers had undefined standards for certain images, and displayed a lack of understanding in details. E.g. in identifying vacuole signs, observer 2 found out 5 cases of cavities in two observations, but only one agreement, in the other 4 cases of disagreement; he believed to be a vacuole at first time, and then confirmed there was no vacuole; then he thought not to have vacuole at first time, and then to be vacuoles. The same situation also was found in the observer 3 during diagnosing vacuole signs. There were 29 cases to be thought as vacuole in two interpretations, among them, 15 cases were diagnosed as vacuole only one time, the first time 6 cases were diagnosed out; 9 cases diagnosed in the second time. Owing to the occurrences of unstable interpretations, it is reasoned that imaging diagnosticians were vague, unfamiliar and lower-sensitive in standard for interpretation.

According to the description above, for recognizing CT signs of solitary pulmonary masses, diagnosticians showed considerable errors in reading images, with up to 19.73 % disagreement between two readings[19-21]. To aim directly at the problem, it is necessary to establish a scientific clinical medical quality system and assessment system. On the basis of traditional teaching, it is necessary to make the use of charts, quantitative data, and semi-quantitative information to identify CT findings. Combining the concept of CT diagnosis and some extensive, comprehensive analysis forms of standardizing diagnosis, should assist in creating a solid foundation for reliable computer-aided diagnosis.

References

[1] Ost D, Fein AM, Feinsilver SH. Clinical practice. The solitary pulmonary nodule. *N Engl J Med* 2003; 348: 2535-42.

[2] Tan BB, Flaherty KR, Kazerooni EA, Lannettoni MD. The solitary pulmonary nodule. *Chest* 2003; 123 : 89S-96S.

[3] Ding QY, Hua YQ, Zhang GZ, Zhao J, Guan YH, Ge XJ, *et al*. A controlled study of position-emission- tomography and position-emission-tom-ography/computed tomography in differ-ential diagnosis of solitary pulmonary nodules eport of 60 cases. *Chin Med J (Engl)* 2005;118: 1572-6.

[4] Trotman-Dickenson B, Baumert B. Multidetector-row CT of the solitary pulmonary nodule(Review). Semin Roentgenol, 2003, 38: 158-167.

[5] Leef JL 3 rd, Klein IS. The solitary pulmonary nodule. *Radiol Clin North Am* 2002; 40: 123-1431.

[6] Wang XH, Chen H, Ma DQ, Gao PY, Zhou XH. Application of artificial neural networks in the CT study of solitary pulmonary nodule. *Chin J Radiol* (*in Chinese*) 2005; 40: 377-82.

[7] Chen G, Ma DQ, Li TY. Detection of Pathognomonic Signs of Small Peripheral Lung Cancers on MPR with Multi-slice Spiral CT. *J Clin Radio*, 2004; 23:109-12.

[8] Yan Y, Hua L, Zhang J. The study of diagnostic concordance with Kappa value and estimated index. *Chin J Health Statistics* (*in Chinese*) 2007;24(3):313-5.

[9] Wakeley CJ, Jones AM, Kabala JE, Prince D, Goddard PR. Audit of the value of double reading magnetic resonance imaging films. *Br J Radiol*, 1995; 68: 358-60.

[10] Leslie A, Jones A, Goddard P. The Influence of clinical information on the reporting of CT by radiologists. *Br J Radiol* 2000; 73:1052-55.

[11] Qiang YQ, Guo YM, Li X, Wang QP, Chen H, Cui DW. The diagnostic rules of peripheral lung cancer preliminary study based on

data mining technique. *JNMU*, 2007, 21:190-5.

[12] Liu WG, Zou LG, Liao CW, Wang WX, Sun QR, Dai SH *et al*. Value of pleural indentation in CT diagnosis of peripheral pulmonary nodules. *Acta Acad Med Mil Tertiae* 2006;28:9-11.

[13] Qiao PG, Li GJ, Lu H, Yang Y, Xing XD. Manifestation and diagnostic value of the vascular convergence sign of peripheral pulmonary cancer on MPR with spiral CT. *Bull Acad Mil Med Sci* 2007; 31:59-61.

[14] Wang XH, Ma DQ, Zhou XH. The Diagnostic Value and Pathological Basis of Air Space in Small Peripheral Lung Carcinoma. *Chin J Med Imaging* (*in Chinese*) 2005, 13:427-30.

[15] Wu FY, Liu W, Zhang SQ, Yu TF. Radiological analysis of image on geriatric pulmonary tuberculosis. *JNMU* 2005;19(5):265-6.

[16] Wang XH, Ma DQ, Chen H. Value of Spiculation on CT in Diagnosis of Small Peripheral Lung Carcinoma. *Chin J Med Imaging* (*in Chinese*) 2006; 14:127-30.

[17] Yang J, Ma DQ. The differential diagnosis of pulmonary inflammatory mass and peripheral lung cancer by computed tomography. *Chin J Radio* (*in Chinese*) 2002; 36:235-9.

[18] Zhang BY, Chen ZQ, Lu GM. Computer aided detection of pulmonary nodules. *J Chin Clin Med Imaging* (*in Chinese*) 2006; 17:148-51.

[19] Fitz GR. Radiological error: analysis, standard setting, targeted instruction, and teamworking. *Eur Radio* 2005;15:1760 -7.

[20] Santiago LM，Craig CB. Evidence-based Radiology: Review and Dissemination. *Radiology* 2007;244:331-6.

[21] Van Rijn JC, Klemetso N, Reitsma JB, Majoie CB LM, Hulsmans FJ, Peul WC, *et al*. Observer variation in MRI evaluation of patients suspected of lumbar disk herniation. *AJR* 2005;184(1):299-303.