Nuclear Science and Techniques **20** (2009) 17–21

Comparison of diagnostic accuracy between ^{18}F-FDG PET and PET/CT for pulmonary neoplasm

CHEN Yangchun[1] CHEN Ping[1,*] TIAN Jiahe[2] CAI Xin[1]
YE Guangchun[1] DENG Huaifu[1] YANG Xiaofeng[3]

[1] *Department of PET/CT Center, No.1 Hospital Affiliated to Guangzhou Medical College, Guangzhou 510230, China*

[2] *Department of Nuclear Medicine, The Chinese PLA General Hospital, Beijing 100853, China*

[3] *Department of Nuclear Medicine, The People's Hospital of Xinjiang Uygur Autonomous Region, Urumuqi 830001, China*

Abstract Aimed at comparing diagnostic accuracy of ^{18}F-FDG PET with PET/CT for pulmonary neoplasm, a study based on multi-center clinical trial of the diagnoses, in randomized and semi-blind ways, was executed from January 2006 to June 2007. It included 55 patients, i.e. 16 with histopathologically proved lung tumors, 16 with tuberculosis and 23 with benign lesions (inflammation, pseudotumor, granuloma, fibrosis and others). The histopathologic and clinic results were served as reference standard. Statistical significances in pulmonary nodule diagnosis between ^{18}F-FDG PET and PET/CT were determined with 95% confidence interval obtained by ROC analysis. The ^{18}F-FDG PET detected lung neoplasm with a sensitivity of 87.5% (14/16), a specificity of 59.0% (23/39), an accuracy of 67.3% (37/55) and a positive-likelihood ratio of 2.13. The ^{18}F-FDG PET/CT detected lung neoplasm with a sensitivity of 93.8% (15/16), a specificity of 61.5% (24/39), an accuracy of 70.9% (39/55) and a positive-likelihood ratio of 2.43. The area under curves (AUC) of ^{18}F-FDG PET and PET/CT were 0.803±0.068 and 0.799±0.063, respectively. It can be concluded that the diagnostic accuracy for malignant pulmonary nodules between ^{18}F-FDG PET and PET/CT was not statistically different.

Key words ^{18}F-FDG, PET/CT, Pulmonary neoplasm, Receiver operating characteristic

1 Introduction

Characterization of pulmonary nodules is established by the indication of metabolic imaging with ^{18}F-FDG. The performance of ^{18}F-FDG PET imaging for pulmonary nodules was the same with PET/CT in a Chinese study[1], but it was different from PET/CT in an American study[2]. In this paper, a paired, open, prospective, randomized and semi-blind multicentre clinical trial was performed for pulmonary nodule patients with ^{18}F-FDG PET/CT[3]. The diagnostic accuracy of ^{18}F-FDG PET was compared with PET/CT for pulmonary neoplasm. Reasons for the difference between the Chinese and American studies were discussed.

2 Materials and methods

2.1 Population

All patients were imaged in a sequential order in one of the 6 PET centers from January 2006 to June 2007.

Age and gender of the patients were not concerned. The inclusion criteria[3] are as follows.

1) Patients of less than three pulmonary nodules;
2) No definite diagnosis before the study;
3) No specific treatment prior to the study;
4) No significant dysfunction or disorder of major organs;
5) Willing to follow the study protocols and to give written consensus for the trial;
6) Possible clinical outcome within foreseeable duration of follow-up.

Supported by Health Bureau of Guangzhou City (Grant No.2007-YB-159)

* Corresponding author. *E-mail address:* fudanzhsh@hotmail.com

Received date: 2008-09-02

The exclusion criteria[3] were as follows.

1) The diagnosis had been defined by other imaging or clinical data;
2) With severe illness or abnormal laboratory findings suggesting metabolic abnormality such as hyperglycemia;
3) Unlikely to comply with the study protocols;
4) Unable to provide necessary clinical data.

The data of a patient would be excluded from the final analysis when the quality of either radiotracer or PET/CT scans was questionable, or the diagnosis was in doubt in the final collective image reading.

2.2 Radiopharmaceuticals

^{18}F-FDG was synthesized in each of the PET centers. The raw materials and agents were purchased from the same supplier. The synthesis and quality control for each preparation of ^{18}F-FDG followed the standard operational protocols (SOPs) strictly and were subject to inspection. The labeling yield, radiochemical purity and specific radioactivity of the product were checked and recorded after each production. The products had to meet certain criteria, such as the labeling yield >10% and the radiochemical purity >95%.

The patient, who had been asked to fast over 4 h and to rest for 15 min, was injected with 5 MBq·kg^{-1} ^{18}F-FDG, and was imaged 60 min after injection.

2.3 Equipment and acquisition parameters

The PET/CT and radiotracer production were based on products of the GE Healthcare, with the Discovery ST PET/CT scanner, the MiniTrace cyclotron and the TracerLab FxFN synthesizer. An obligatory standardized quality-control program was followed by all PET centers and was subject to the organizer's inspections. To protect the patients from undue radiation dose, a low-dose CT scan was acquired with the following settings: automatic adjustment of 120 kV and 100–250 mA, rotation of 0.8 s, slice thickness of 2.5 mm and a pitch of 1.25. The PET scanner has a 15.7-cm axial field width and a spatial resolution of 4 mm (FWHM) at 1 cm from the center. PET images of 3 or 4 bed positions covering the entire chest were acquired by 2.5 min per bed position in 3-dimensional mode. In some cases, the whole-body imaging was performed from skull base to upper thighs. The images were reconstructed by a Fourier rebinding iterative algorithm.

2.4 Image interpretation

The CT images were displayed as 3.75-mm cross-axial slices. Morphologic features of the nodule(s) were checked, such as size, density, cavity, calcification, notch on margin, speculated margin or plural contraction, etc. The CT value was assessed, and the PET images were inspected with the maximum standardized uptake value (SUVmax) determined from a circular region of interest (ROI) over the entire lesion. Uptake of a lesion was scored as,

0, No uptake.
1, Uptake lower than that of the mediastinum.
2, Uptake equal to or greater than that of the mediastinum but lower than that of the liver.
3, Obvious uptake higher than that of the liver and
4, Very strong uptake[2,3].

The threshold for malignancy was set as SUVmax $\geq$2.5[1,3] or score$\geq$2[2,3].

2.5 Data collection and verification

The original copies of working sheet, the imaging data and the data records for each patient were sealed individually and were sent to the organizing center over the Internet or by means of CD-ROM. All sets of the serially numbered forms, used or not used, should be returned to the organizing center No correction or modification was allowed. The following datasets were collected from every center:

1) One original copy of PET/CT working sheet.
2) One package of raw image data from the PET/CT imaging.
3) The original follow-up records, with the date and type of surgical procedure and the pathologic diagnosis or the date and findings on follow-up.
4) The original copy of radiopharmaceutical production sheet, with the information on the production, quality control, and the raw material and agents used for each synthesis.
5) The signed consent form from every subject.
6) A summary of all cases, successful or failed, with the relevant information and explanations.
7) A summary of the execution of the multi-center clinical trial (MCCT) by each PET center.

All the data were verified before further processing by physicians, physicists, radiochemists, administrators, and inspectors in the organizing center. Any noncompliance with the MCCT protocols resulted in exclusion of the patient's data.

2.6 Collective image reading

Two sessions of blinded, collective-image reading were carried out in this trial. The first session, aimed at verifying the interpretation criteria, was organized six months after initiating the trial with three independent readers of CT or PET professional background. The second reading session took place on completion of the trial with an expanded team of readers. Only the results of the second reading session were analyzed and reported in this article. Nine readers, four CT professionals and five nuclear medicine professionals took part in the final collective reading. They, who had 1–4 years working experience in PET/CT when the trial began, were responsible for the primary PET/CT-image interpretation in their own imaging center.

In the collective reading, the images were reconstructed and assessed using the central workstation. ^{18}F-FDG PET and ^{18}F-FDG PET/CT images of a patient were read in pairs. In each round of the reading, the images were randomized, with the image heading being masked until viewing them, and the readers were unaware of the patient's history. The images were projected onto a screen, with the display windows and angles being adjusted to a reader's request, assisted by an independent operator. No discussion among the readers was allowed. Each reader had to make his/her own judgment on each subject and score the images. The recording sheet of every reader was collected before the next round of reading. The imaging diagnosis was determined by a consensus of at least five of the nine readers, and the corresponding score was determined by averaging the total. The results were statistically analyzed in light of the standard of truth.

2.7 Endpoint and standard of truth

The endpoint of this trial was determined as either the pathologic evidence obtained from surgical processes or the clinical conclusion derived from the therapeutic response or from imaging or laboratory findings in follow-up over one year after imaging. Therefore, the standard of truth was the histological diagnosis or the validated clinical evidence derived at the end of 12-month follow-up.

2.8 Statistical analysis

The results are presented as the mean ± SD. The accuracy, sensitivity and specificity for PET and PET/CT were calculated to evaluate statistical significances with the paired-x^2 statistics. An ROC (receiver-operating-characteristic) analysis was done on the PET/CT and PET interpretations.

3 Results

The first patient was imaged in January 2006, and the follow-up of the last patient was completed in June 2007. Fifty-five patients (22 women aged 17–72 and 33 men aged 31–82) met the criteria and were included for further analysis. They were 16 lung tumor patients, 16 tuberculosis patients, and 23 benign lesion patients (inflammation, pseudotumor, granuloma, and other benign conditions). The diagnosis was confirmed *via* surgical processes (operation or biopsy) in 27 patients or *via* other clinical processes in 28 patients. No side effects were reported with either radiopharmaceutical or in PET/CT scanning. The mean nodule size was (27.6±26.5) mm in 16 tumor patients and (18.5 ± 9.4) mm in 39 non-tumor patients. By semi-quantitative analysis, the SUVmax for all malignant lesions averaged at 10.4±6.7, for benign lesion averaged at 5.8 ± 4.5.

The lung neoplasm was detected by PET with a sensitivity of 87.5%, a specificity of 59.0%, an accuracy of 67.3% and a positive-likelihood of 2.13 (Table 1). The false-negative findings included adenocarcinoma (n=1) and bronchioalveolar carcinoma (n=1) with nodules of 8 mm and 29 mm, respectively. The false-positive findings included tuberculosis (n=9), inflammation (n=4), granuloma (n=2), and benign lesions (n=1).

The lung neoplasm was detected by PET/CT with a sensitivity of 93.8%, a specificity of 61.5%, an accuracy of 70.9% and a positive-likelihood of 2.43

(Table 1). A false-negative finding of bronchioalveolar carcinoma, which was also mistaken by PET, was found. The false-positive findings were tuberculosis (n=6), inflammation (n=6), granuloma (n=2), and benign lesion (n=1).

With the paired-χ^2 statistics, no difference was found (Table1) between PET and PET/CT in accuracy, sensitivity, and specificity (P>0.05).

Table 1 ^{18}F-FDG PET and PET/CT in characterization of lung tumors in China and America

	Ding Q Y (China)		Kim S K (America)		This work	
^{18}F-FDG	PET	PET/CT	PET	PET/CT	PET	PET/CT
Sensitivity (%)	86.7(26/30)	90.0(27/30)	69.0(20/29)	96.6(28/29)*	87.5(14/16)	93.8(15/16)
Specificity (%)	90.0(27/30)	93.3(28/30)	84.6(11/13)	84.6(11/13)	59.0(23/39)	61.5(24/39)
Accuracy (%)	88.3(53/60)	91.7(55/60)	73.8(31/42)	92.9(39/42)*	67.3(37/55)	70.9(39/55)

An ROC analysis was performed for PET and PET/CT (Fig. 1). The areas under curve (AUC) were 0.803 ± 0.068 for PET (95% confidence in 0.670~0.936 interval) and 0.799 ± 0.063 for PET/CT (95% confidence in 0.675~0.923 intervals), hence similar performance of the ^{18}F-FDG PET and PET/CT.

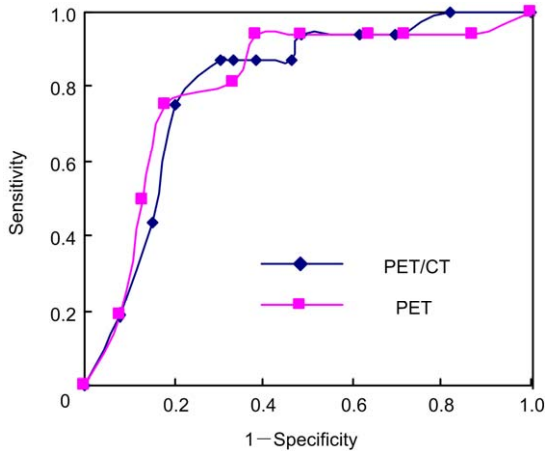

Fig.1 ROC results, PET and PET/CT curves based on consensus interpretation with AUC of 0.803 for PET and 0.799 for PET/CT.

4 Discussion

The cutoff for malignancy was set as SUVmax ≥2.5, the same as in Ref.[1]. The PET/CT sensitivity of 93.8% and the PET sensitivity of 87.5% in our study are similar to Ref.[1]. The threshold for malignancy was set as SUVmax ≥2.0 in Ref.[2]. The PET in Ref.[2] may have the highest sensitivity of the three studies, but its sensitivity is just 69.0%. This may be due to smaller nodule diameter in Ref.[2], where nine lesions were <10 mm in axial diameter.

The object affects the specificity and sensitivity of PET or PET/CT. The objects coming from different cross sections may lead to different results. The data in this study came from a paired, open, prospective, randomized and semi-blind multicentre clinical trial, which could exclude the possibility of any artificial variance as much as possible. Most of the false-positive cases were tuberculosis (9/16 by PET, 5/16 by PET/CT) in this study with lower specific than the other two studies that performed retrospectively.

As diagnostic value of ^{18}F-FDG PET is known to be restricted by tuberculosis in lung neoplasm diagnosis[3], whether ^{18}F-FDG PET can be used in tuberculosis or not is still controversial. According to the same diagnosis standards with lung neoplasm, the tuberculosis was detected by ^{18}F-FDG PET with a sensitivity of 56.3%, a specificity of 37.5%, and a positive-likelihood of 1.04 in this study. This suggests that ^{18}F-FDG PET is useless in tuberculosis diagnosis. And ^{18}F-FDG PET/CT was also useless because of its positive- likelihood of < 1 (only 0.61).

We tried to discuss the utility of ^{18}F-FDG PET in diagnosing lung neoplasm in China by calculating its positive forecast rate and comparing the rate with that in America. The positive forecast rate (PF) of ^{18}F-FDG PET can be obtained with the Bayes's theory[4] of PF=Sen×Inc/[Sen×Inc+(1-Spe)×(1-Inc)], where Sen is the sensitivity of ^{18}F-FDG PET, Spe is the specificity of ^{18}F-FDG PET and Inc is the lung neoplasm (or tuberculosis) incidence.

The sensitivity and specificity of ^{18}F-FDG PET are used in the study. It is estimated that annual lung neoplasm incidence in China and the US is 37.8 and

57.3 per 100 000 population, respectively[5-7]. From the WHO database, annual tuberculosis incidence in China and the US was 99 and 4 per 100 000 population in 2006[8], respectively. And the PF rates calculated by the Bayer's theory are listed in Table 2. The rate of lung tumor to tuberculosis is 0.62 in China, while it is 23.4 in America, suggesting that ^{18}F-FDG PET scan is more useful to detect lung neoplasm in America than in China.

Table 2 Lung neoplasm/tuberculosis incidence (per 100 000) and the positive forecast rate (per 100 000) of 18F-FDG PET on lung neoplasm/tuberculosis in China and the USA

Item	China	USA
Lung neoplasm incidence	37.8	57.3
TB incidence	99	4
PF of lung neoplasm (*L*)	64.5	97.7
PF of TB (*T*)	103	4.2
L/*T*	0.62	23.4

Prior study showed that the threshold for malignancy set as SUVmax≥6.0 could improve specificity of PET[9]. By that standard, lung tumor versus tuberculosis was detected by PET with a sensitivity of 75.0% versus 43.8%, and specificity of 64.1% versus 51.3%. According to Bayer's theory, when a Chinese patient is ^{18}F-FDG PET scan positive, the person with the probability of lung tumor to tuberculosis is 0.89. So the role of ^{18}F-FDG PET could be improved by raising the threshold.

However this study is based on only 55 lung nodule patients, but the diagnosis of only about half (27) of them was confirmed *via* surgical processes. Apparently, a perfect conclusion shall be based on larger numbers of patients.

5 Conclusion

Lung neoplasm can be distinguished from non-tumor nodules by ^{18}F-FDG PET or PET/CT. ^{18}F-FDG PET scan is more useful to detect lung tumor in America than in China. There was no difference in detecting lung tumor between PET and PET/CT in China. The role of ^{18}F-FDG PET could be improved by raising the threshold in China.

Acknowledgement

We thank the staff members of the six PET centers. We thank ZHAO Zhoushe, WU Wenkai, LI Hongli and WANG Shuang for their help of supervising the study.

References

1 Ding Q Y, Hua Y Q, Guan Y H, *et al.* Chin Nucl Med, 2005, **25:** 261–263 (in Chinese).

2 Kim S K, Allen-Auerbach M, Goldin J, *et al.* J Nucl Med, 2007, **48:** 214–220.

3 Tian J H, Yang X F, Yu L J, *et al.* J Nucl Med, 2008, **49:** 186–194.

4 Pan Z H Y. Clinical nuclear medicine. Beijing: Atomic Energy Press, 1994, 109–118 (in Chinese).

5 NCCN clinical practice guideline in oncology: non-small cell lung cancer. Fort Washington: National Comprehensive Cancer Network, 2007 (in Chinese). http://www.nccn-china.org/images/PDF/NSCLC%202007%20Guideline%20Chinese%20Edition%20%20(DRAFT-2)%20NCCN-APPROVED%2007-17-07.pdf.

6 National economic and social development of statistical bulletin. Beijing: National Bureau of Statistics of China, 2008 (in Chinese). http://www.stats.gov.cn/tjgb/ndtjgb/qgndtjgb/t2008 0228 _402464933.htm.2008.2.28.

7 The U.S. population clock. Washington: U.S. Census Bureau, Population Division, 2008. http://www.census.gov/population/www/popclockus.html.

8 The estimated TB incidence in China versus America. World Health Organization, Global Tuberculosis Database, 2008. http://www.who.int/globalatlas/dataQuery/reportData.asp?rptType=1.2008.3.17.

9 Chen P, Tian J H, Yang X F, *et al.* Chin Nucl Med, 2008, **28:** 151–153 (in Chinese).