In this issue of the *International Journal of Clinical Oncology*, Drs. Arai and Osuga, who are representative interventional oncologists in Japan, report topics of interest in interventional oncology that reveal the utility and significance of its existence.

Conflict of interest The author declares that he has no conflict of interest.

Ann Nucl Med
DOI 10.1007/s12149-013-0712-x

ORIGINAL ARTICLE

Observer variation study of the assessment and diagnosis of incidental colonic FDG uptake

Ryogo Minamimoto · Takashi Terauchi · Seishi Jinnouchi · Tsuyoshi Yoshida · Eriko Tsukamoto · Takuro Shimbo · Kimiteru Ito · Kimiichi Uno · Hitoshi Ohno · Kazuhiro Oguchi · Satoshi Kato · Koichiro Kaneko · Yoko Satoh · Tsuneo Tamaki · Tadaki Nakahara · Miyako Morooka · Tomio Inoue · Michio Senda

Received: 25 October 2012 / Accepted: 24 February 2013

Abstract
Purpose The aim of this study was to evaluate the interpretations of incidental colonic 18F-FDG uptake made by 10 experienced readers and to more clearly identify the pattern of suspicious colonic FDG uptake. The potential contributions of delayed FDG-PET scanning and of immune fecal occult blood testing (FOBT) in making a diagnosis were also analyzed.
Materials and methods Visual interpretations by 10 readers were made for 147 FDG uptake sites from 126 PET scans (cancer, 38 sites; adenoma, 43 sites; and no abnormality, 66 sites) with colonic FDG uptake. Assessments for the early FDG-PET images were (1) FDG uptake pattern, (2) FDG uptake degree, and (3) likelihood of malignancy. For the delayed images, the assessments were (1) change in the FDG uptake position, (2) change in FDG uptake degree, and (3) likelihood of malignancy. The results of FOBT were analyzed independently of the visual interpretations.
Results Interobserver agreement (κ) was 0.501 for assessing FDG uptake patterns, while agreement on assessing changes in uptake degree and changes in uptake position between early and delayed imaging were low ($\kappa =$ 0.213–0.229). Logistic regression analysis indicated that 'FDG uptake patterns' and 'FDG uptake degree' were significantly related to decide on the suspicion of malignancy ($p < 0.001$) and the final result ($p < 0.001$). "Small localized" and "large irregular localized" types had a high probability of a lesion regardless of either (1) FDG uptake degree or (2) variation in the uptake between the early and the delayed image. The delayed image decreased false-positive cases for some FDG uptake patterns, but it had

R. Minamimoto (✉) · M. Morooka
Division of Nuclear Medicine, Department of Radiology, National Center for Global Health and Medicine, 1-21-1, Toyama, Shinjyuku-ku, Tokyo 162-8655, Japan
e-mail: ryogominamimoto@yahoo.co.jp

R. Minamimoto · T. Inoue
Department of Radiology, Graduate School of Medicine, Yokohama City University, Yokohama, Japan

T. Terauchi
Screening Technology and Development Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan

S. Jinnouchi
Atsuchi Memorial Institute of Radiology, Atsuchi Memorial Clinic PET Center, Kagoshima, Japan

T. Yoshida · K. Kaneko
PET Imaging Center, Koga Hospital 21, Kurume, Japan

E. Tsukamoto
Central CI Clinic, Sapporo, Japan

T. Shimbo
Department of Clinical Research and Informatics, National Center for Global Health and Medicine, Tokyo, Japan

K. Ito
Department of Radiology, National Center of Neurology and Psychiatry, Tokyo, Japan

K. Uno
Gaienhigashi Clinic, Tokyo, Japan

H. Ohno
Department of Radiology, Saitama Medical Center, Saitama Medical School, Saitama, Japan

K. Oguchi
Positron Imaging Center, Aizawa Hospital, Matsumoto, Japan

Published online: 17 March 2013

little impact on distinguishing clearly between "cancer or adenoma" and "normal". The addition of FOBT had little impact on the diagnosis.

Conclusion There was highest agreement among readers with respect to the recognition of specified colonic FDG uptake patterns, and this pattern recognition had the most influence on the diagnosis. "Small localized" and "large irregular localized" types had a high probability of a lesion. The addition of delayed imaging and of FOBT results to the early imaging did not have much impact on the diagnosis.

Keywords Colonic FDG uptake · PET · Colorectal cancer · Observer validation study · Fecal occult blood testing (FOBT) · Dual-time-point PET

Introduction

Incidental colonic FDG uptake on PET examination is often problematic throughout the diagnostic process. Colonic FDG uptake does not always indicate a malignancy, since FDG uptake can be seen in a variety of conditions, such as physiological variations, adenomas, and inflammatory lesions [1]. Thus, PET is not recommended for screening of colonic malignancy but rather for cancer staging that is related to the detection of regional lymph nodes and distant metastases. On the other hand, colonic FDG uptake has been seen in 1.3 % [2] to 3.0 % [3] of PET scans, and half [2] to two-thirds [3] of these patients were subsequently diagnosed with an unsuspected colonic carcinoma or adenoma [4]. Therefore, it appears that colonic FDG uptake should not be ignored since it has the potential for indicating malignancy. Although the mechanism of FDG uptake in the colon is still unclear, colonic FDG uptake has been compared to pathological assessment [5–10]. However, colonic FDG uptake shows a variety of patterns in form and degree. Tatlidil et al. [11] reviewed the pattern of colonic FDG uptake in 27 patients by classifying them into 4 FDG uptake patterns and 4 grades of FDG uptake. Cases of nodular-focal FDG uptake consisted of abnormal lesions that were either malignant or benign. Those with segmental FDG uptake mainly consisted of colitis. Those with diffuse FDG uptake were noncancerous. However, the number of subjects in the cited study was relatively small, and the study was dominated by cases showing intense FDG uptake (78 % of cases). The results of the study are considered the reference for assessment of colonic FDG, but further evaluation with a larger number of cases and by current experienced readers are needed to illuminate the questions that remain about colonic FDG uptake.

The purpose of the present study was to evaluate the interpretations of incidental colonic FDG uptake made by 10 experienced readers and to more clearly identify the pattern of suspicious colonic FDG uptake. In addition, we wanted to examine whether a second FDG-PET time point or immune fecal occult blood testing (FOBT) could contribute to increasing the diagnostic accuracy of incidental colonic FDG uptake.

Materials and methods

Study design

All study protocols in this retrospective observation study were approved by the institutional review board of Yokohama City University. The cases for evaluating colorectal FDG uptake were taken from FDG-PET scans carried out as cancer screening of asymptomatic subjects at three PET centers in Japan. From 2003 to 2009, 24,295 PET scans were performed in the three PET centers. Abnormal incidental colonic FDG uptake was identified in 354 cases (1.5 %) by a physician specializing in nuclear medicine at each facility, and colon cancer was confirmed in 103 cases (0.4 %).

Inclusion criteria for our study were: (1) in addition to the early scan, a delayed scan at 120 min after FDG injection; (2) FOBT using fecal samples collected on two consecutive days; (3) age $\geq$40 years; (4) blood glucose level <120 mg/dl at the time of the PET scan; (5) no history of any cancer; and (6) final pathological diagnosis for colorectal lesion had been obtained by surgery and/or colonoscopy or by at least 2-year follow-up without cancer development.

According to the criteria, 147 sites in 126 subjects (71 males, 55 females; average age 61.2 years, range 42–87 years) with colorectal FDG uptake were selected for visual interpretation, consisting of 38 FDG uptake sites of

S. Kato
Department of Diagnostic Radiology, Kofu Municipal Hospital, Kofu, Japan

Y. Satoh
PET Center, Kofu Neurosurgical Hospital, Kofu, Japan

T. Tamaki
Department of Radiology, East Nagoya Imaging Diagnosis Center, Nagoya, Japan

T. Nakahara
Department of Radiology, School of Medicine, Keio University, Tokyo, Japan

M. Senda
Division of Molecular Imaging, Institute of Biomedical Research and Innovation, Kobe, Japan

cancer, 43 uptake sites of adenoma, and 66 uptake sites which were proven not to represent an abnormal lesion. The mean glucose level of the 126 subjects was 97.0 mg/dl.

The diagnosis of colorectal FDG uptake in the 147 sites was established by colonoscopy in 117 sites, surgery in 12 sites, and at least 2 years of observation in 18 sites. If the case had colonic FDG uptake on the first FDG scan and no lesion was confirmed within 2 or 3 years of observation after the first scan, any colorectal FDG uptake observed at the first scan was regarded as "normal".

PET image interpretation

All 147 sites with colorectal FDG uptake were evaluated on the basis of visual inspection by each of 10 experienced, board-certified, nuclear medicine physicians with various degrees of experience with PET imaging diagnosis: 5 years for 3 physicians (S.K., Y.S., M.M); 6 years for 2 physicians (K.I., H.O.); 8 years for one physician (K.O.); 9 years for one physician (K.K.); and 10 or more years for 3 physicians (K.U., T.N., T.T.). All reconstructed PET images were reviewed on a Vox-Base SP1000 workstation (J-MAC Systems, Sapporo, Japan), and the interpretation was based on transaxial, sagittal, coronal, and maximum intensity projection (MIP) views with grayscale images. The observers were requested to make an interpretation and assess the likelihood of malignancy either on the early or the delayed scan. If the image was obtained by PET/CT, only the PET image was provided for observers to evaluate.

At the time of PET image interpretation, the observers were provided with the patient's (1) age, (2) sex, (3) and blood glucose level, and (4) the FDG uptake region for interpretation. The aim of interpretation was not to screen the FDG uptake, but to assess the specific incidental colonic FDG uptake. Therefore, the FDG uptake for interpretation for each case was indicated on a print-out, which was only used to identify the FDG uptake region for interpretation; the visual assessment was performed based on the viewer images mentioned above.

Assessed for the early scan were: (1) FDG uptake pattern (small localized, large irregular localized, short segmental, and long segmental); (2) FDG uptake degree (very high, high, moderate, and slight); and (3) likelihood of malignancy (highly likely, likely, possibility of malignancy, possibility of benign lesion, no abnormality). Assessed for the delayed scan were: (1) change in FDG uptake degree (increased, stable, decreased, and unconfirmed; (2) change in the FDG uptake position (moved or unmoved); (3) and likelihood of malignancy (highly likely, likely, possibility of malignancy, possibility of benign lesion, no abnormality) based on the interpretation of the delayed scan.

Before starting the study, these image interpretation criteria were discussed by all readers. Observers shared a common understanding that "small localized FDG uptake" was small solitary nodular FDG uptake, "large irregular localized FDG uptake" was large solitary FDG uptake indicating the existence of a mass lesion, "short segmental FDG uptake" was FDG uptake along a short reach of colon, and "long segmental FDG uptake" was FDG uptake along a long reach of colon. The representative FDG-PET images for every FDG uptake pattern and FDG uptake degree, some of which are shown in Fig. 1, were selected by agreement of all readers. All readers carried out the interpretation of FDG-PET images by consulting these images.

For each case, a 'consensus interpretation' was determined in one of two ways. It was the choice selected by five (the majority) of the readers, if that number or more agreed on a given choice. If no such agreement existed, the case was then re-reviewed by the three most experienced physicians (S.J., T.Y., E.T.) together; the choice they agreed on became the 'consensus interpretation'. Then, the interpretations of "likelihood of malignancy" were reclassified into just two groups. The first group was called "suspicious for malignancy" and was composed of cases probably malignant (the two original designations "highly likely" and "likely") and cases only suspicious of malignancy (the original designation "possibility of malignancy"). The second group was called "malignancy unlikely" and was composed of the two original designations "possibility of benign lesion" and "no abnormality".

Quantitative analysis was performed by using maximum standardized uptake value (SUV_{max}), with regions of interest (ROIs) drawn over each colonic FDG uptake for interpretation. The analysis was performed independently to the interpretation of the observers. It was aimed at the verification of the result of visual observation. The SUV was classified according to 4 categories (very high, high, moderate, and slight).

"The category FOBT-positive" was defined as at least one positive FOBT test among the two FOBT tests. If the consensus interpretation in PET imaging was combined with FOBT, "positive" was defined as at least one positive result. The results of FOBT testing were not provided to the observers. Likewise, FOBT tests were analyzed independently.

PET scans

Of the 126 PET scans, 93 had been performed using the GE Advance Nxi in two-dimensional mode, with emission scans and transmission scans per bed of 2 and 1 min, respectively, starting at 60 min for the early scan and 120 min for the delayed scan after administration of 3.7 MBq/kg FDG. The scans were corrected for decay, scatter, and randoms and reconstructed using ordered

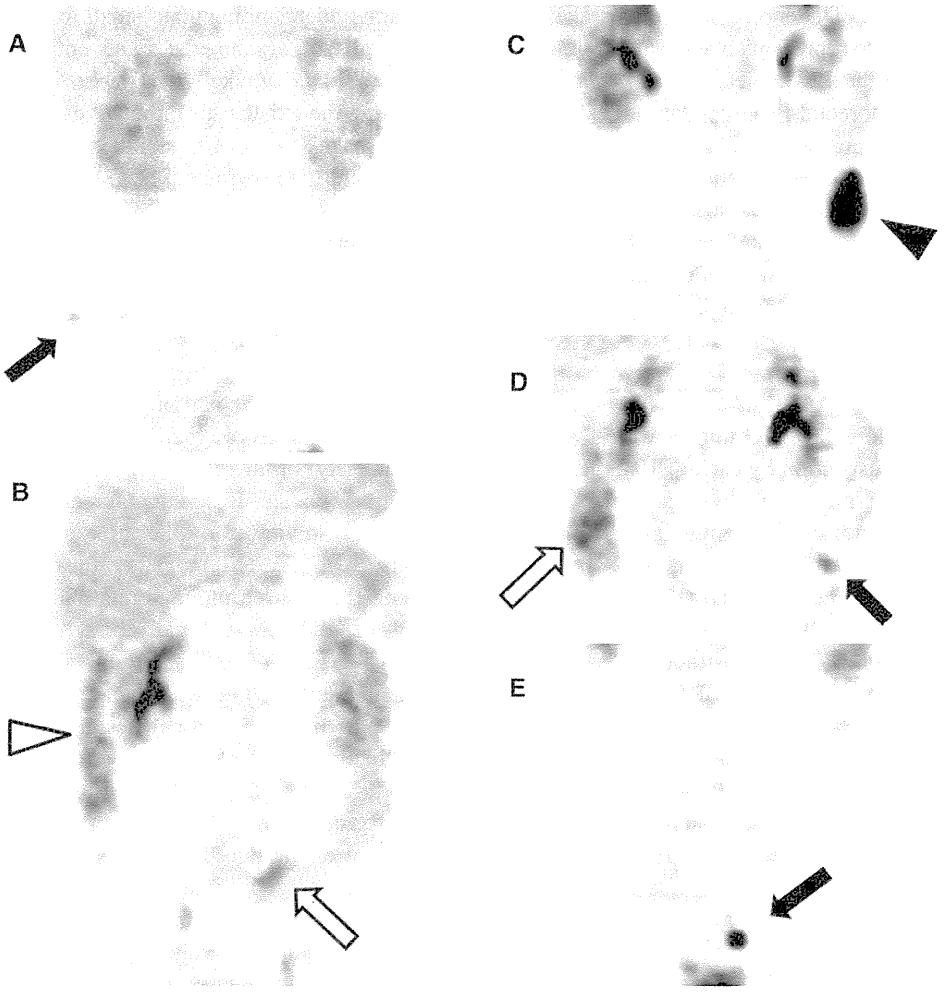

Fig. 1 **a** (*arrow*) Uptake pattern: small localized, Uptake degree: slight. **b** (*white arrow head*) Uptake pattern: long segmental, Uptake degree: moderate (*white arrow*) Uptake pattern: short segmental, Uptake degree: moderate. **c** (*arrow head*) Uptake pattern: large irregular localized, Uptake degree: very high. **d** (*white arrow*) Uptake pattern: short segmental, Uptake degree: moderate (*arrow*) Uptake pattern: small localized, Uptake degree: moderate. **e** (*arrow*) Uptake pattern: small localized, Uptake degree: high

subset expectation maximization (OSEM) with two iterations. In the other 33 cases, 2 PET/CT scanners were used: the GE Discovery ST Elite performance in three-dimensional mode with emission scans per bed of 2 min and reconstructed using OSEM; and the Philips Gemini GXL in three-dimensional mode with emission scans per bed of 2 min and reconstructed using 3D-line of response (LOR). PET/CT scanning was carried out with the same timing as above after 4.5 MBq/kg FDG administration. All PET studies were performed with at least 4 h of fasting prior to the FDG injection.

Statistical analysis

Levels of interobserver agreement were quantified using kappa values (κ values) [12]. The higher the kappa value, the better is the agreement between the observers. Bootstrapping was used to calculate 95 % confidence intervals. The averages of SUV_{max} for colonic FDG uptake were expressed in terms of mean ± standard deviation (SD), and Mann–Whitney *U* test was used to evaluate the differences between categories of visual interpretation. The percent change in the SUV_{max} of colonic FDG uptake between early and delayed image was calculated for the purpose of comparing the visual interpretation of FDG uptake change in the delayed image with the change ratio of SUV_{max}. The Mann–Whitney *U* test was applied to determine the significant difference of SUV_{max} change in each category of visual interpretation.

The differences in the PPV and NPV were analyzed using a marginal logistic regression model. The sensitivity and specificity obtained were regarded as relative values because of the selection bias in the study cases. Thus, the terms "relative sensitivity" and "relative specificity" are

used in our report. The McNemar paired test was used to evaluate the differences in sensitivity and specificity between early and delayed images in terms of visual analysis and location-based analysis.

A logistic regression analysis was performed with the following factors: "FDG uptake pattern" and "FDG uptake degree" for suspecting malignancy in the early image and the final result, and "FDG uptake patterns", "FDG uptake degree", "change in the FDG uptake position in the delayed image", and "change in FDG uptake degree in the delayed image" for suspecting malignancy in the delayed image and the final result.

A Chi-square test was performed with the following status: "the case showing moved colonic FDG uptake", or "decreased or unconfirmed FDG uptake in delayed image" and the case without showing these change in delayed image, and the case with change in FDG uptake position and the case without change in FDG uptake position.

Statistical analyses were performed using the statistical package Stata (version IC 11; Stata Corp, TX, USA). Values of $p < 0.05$ were considered significant.

Results

Re-review by the three most experienced physicians to reach an agreement was carried out in 5 cases for FDG uptake patterns, 8 cases for FDG uptake degree, 11 cases for FDG uptake degree change on the delayed image, 8 cases for change of FDG uptake position on the delayed image, 8 cases for likelihood of malignancy on the early image, and 1 case for likelihood of malignancy on the delayed image.

Interobserver agreement among the 10 readers for each index is summarized in Table 1. The interobserver κ values was 0.501 for assessing the FDG uptake pattern on the early image, This agreement was better than that for uptake degree on the early image or for either of the two categories on the delayed image. However, the interobserver κ value for suspected malignancy was higher with the delayed image ($\kappa = 0.518$) than with the early image ($\kappa = 0.489$).

Table 1 Result of interobserver agreement

Image	Interpretation	κ	95 % CI
Early	FDG uptake pattern	0.501	0.456–0.516
	FDG uptake degree	0.310	0.286–0.326
	Suspected malignancy	0.489	0.473–0.511
Delayed	Change of uptake degree	0.229	0.209–0.241
	Change of uptake position	0.213	0.181–0.218
	Suspected malignancy	0.518	0.504–0.523

The average of SUV_{max} according to consensus interpretation of FDG uptake degree was 21.7 ± 9.0 (range 11.2–41.3) at "very high", 9.5 ± 3.0 (6.4–16.8) at "high", 5.5 ± 1.1 (3.9–8.2) at "moderate", and 4.1 ± 0.7 (2.9–5.8) at "slight". Statistical significance was confirmed in FDG uptake degree between "very high" and "high" ($p < 0.001$), "high" and "moderate" ($p < 0.001$) and "moderate" and "slight" ($p < 0.001$). These significances indicate that FDG uptake degree was clearly differentiated by consensus interpretation. For consensus interpretation for change in FDG uptake degree, the average of change ratio (%) for SUV_{max} between the early image and the delayed image was 32.1 % (range −34.0 to 113.3 %) at "increased", 1.8 % (−30.8 to 44.4 %) at "stable", and −30.2 % (−73.1 to 13.8 %) at "decreased". Statistical significance was confirmed in "increased" and "stable" ($p < 0.001$) and "stable" and "decreased" ($p < 0.001$). These significances indicate change of FDG uptake in delayed imaging was clearly differentiated by consensus interpretation.

The results of the visual analyses based on 10 readers' interpretations and consensus interpretations are summarized in Table 2. The result was independent of experience with PET imaging diagnosis. The delayed image increased the relative specificity ($p < 0.02$) compared to early imaging, but resulted in little impact on PPV ($p = 0.50$), NPV ($p = 0.72$), and relative sensitivity. If their interpretation was applied to detect "cancer and adenoma", no significant difference was observed between the early and delayed images. The combination of FOBT showed no statistically significant difference in PPV and NPV compared to any PET image. The combination of FOBT and the early image for detecting cancer decreased relative specificity ($p < 0.001$) compared to the early image alone. The combination for detecting cancer and adenoma increased relative sensitivity ($p < 0.001$) but decreased specificity ($p < 0.005$) compared to the early image alone. The combination of FOBT and the delayed image for detecting cancer decreased relative specificity ($p < 0.001$) compared to the delayed image alone. The combination for detecting cancer and adenoma increased relative sensitivity ($p < 0.003$) compared to the delayed image alone (Table 2).

Logistic regression analysis based on consensus interpretation showed that 'FDG uptake pattern' and 'FDG uptake degree' were significantly related to the decision about suspecting malignancy and the final result ($p < 0.001$). The change in the FDG uptake degree on the delayed image influenced the suspicion of malignancy on the combined delayed image ($p < 0.001$), in contrast to the lack of effect of FDG uptake position change on the delayed image ($p = 0.08$). The delayed image did not have a significant relation to the final result (FDG uptake degree changed on the delayed image, $p = 0.79$; FDG

Table 2 Results of visual analysis

Lesion type	Imaging	Interpretation	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)
Cancer	Early	10 readers	29.1–40.3	86.1–97.4	71.1–97.4	22.0–57.8
		Consensus	33.0	92.7	92.1	34.9
		Consensus + FOBT	31.7	100.0	100.0	22.8
	Delayed	10 readers	28.1–49.2	82.9–94.7	76.3–94.7	20.1–57.8
		Consensus	37.6	94.4	92.1	46.8
		Consensus + FOBT	33.0	97.1	97.4	31.2
Cancer and adenoma	Early	10 readers	59.8–74.0	62.5–79.5	65.4–92.6	28.0–69.7
		Consensus	65.1	70.7	86.4	43.9
		Consensus + FOBT	62.5	77.8	92.6	31.8
	Delayed	10 readers	57.9–84.5	53.8–71.1	65.4–87.6	22.7–83.3
		Consensus	69.9	70.4	80.2	57.6
		Consensus + FOBT	64.3	74.3	88.9	39.4

The results from 10 readers are represented by the range

PPV positive-predictive value, *NPV* negative-predictive value, *FOBT* fecal occult blood test

Table 3 Results of logistic regression analysis

Image	FDG uptake	p value	Subject
Early	FDG uptake patterns	<0.001	Suspected malignancy in early image
		<0.001	Final result
	FDG uptake degree	<0.001	Suspected malignancy in early image
		<0.001	Final result
Delayed	Change of uptake degree	<0.001	Suspected malignancy in early image
		0.79	Final result
	Change of uptake position	0.08	Suspected malignancy in early image
		0.37	Final result

uptake position changed on the delayed image, $p = 0.37$) (Table 3).

The results of interpretations based on location are summarized in Table 4. Region-based analysis on the early image had a low PPV in the all areas of the colon, except the rectal region. The delayed image tended to increase PPV and relative specificity for judging "cancer" and "cancer and adenoma". However, no statistically significant difference between the early and delayed images was found for all the regions.

The results for the lesions based on visual interpretations are summarized in Table 5. Small localized FDG uptake included a large number of cancers, but the relative specificity was low because the FDG uptake pattern included a large number of adenomas. The large irregular localized type and the very high FDG uptake were relatively specific to cancer.

Delayed imaging did not improve the results in cases of small and large, irregular, localized FDG uptake (Table 6). The incidences of "decreased", "unconfirmed", and "moved" classifications for the delayed image were small in number, but tended to exclude normal cases preferentially. If the cases with "decreased" "unconfirmed" or "moved" were excluded (Table 6, exclusion A), 40.0 % (16/40) of normal cases within short segmental pattern or long segmental pattern, and 27.6 % (16/58) of normal cases showing over moderate FDG uptake could be reduced. However, delayed imaging did not contribute to distinguishing the cancerous from the noncancerous case for any FDG uptake pattern observed in the early imaging ($p \geq 0.16$).

Discussion

The present study assessed observer variation in characterizing incidental colonic FDG uptake. Interobserver agreement was highest for assessing the FDG uptake pattern, and "small localized" and "large irregular localized" patterns had high probabilities of colon cancer regardless of FDG uptake degree and variation in the uptake on the delayed image. The delayed image contributed to decreasing false-positives when either "short segmental" or "long segmental" FDG uptake on the early image was confirmed.

Colonic FDG uptake shows a variety of patterns in form and degree. In the present study, the result of interpretation was not based on individual experience but rather a consensus judgment. Thereby, we sought to establish a strategy for diagnosing incidental colonic FDG uptake that was robust. Tatlidil et al. [11] reviewed the patterns of colonic FDG uptake by classifying them into 4 FDG uptake patterns and 4 grades of FDG uptake. However, determining

Table 4 Results of region-based analysis based on consensus interpretation

Lesion	Region	Number				Early				Delay			
		To	Ca	Ad	Nor	PPV (%)	NPV (%)	Sen (%)	Spe (%)	PPV (%)	NPV (%)	Sen (%)	Spe (%)
Cancer	Cecum	14	2	1	11	22.2	100.0	100.0	41.7	33.3	100.0	100.0	66.7
	Ascending	38	10	11	17	35.7	100.0	100.0	35.7	43.5	100.0	100.0	53.6
	Transverse	7	2	2	3	25.0	66.7	50.0	40.0	25.0	66.7	50.0	40.0
	Descending	21	5	7	9	26.7	83.3	80.0	31.3	30.8	87.5	80.0	43.8
	Sigmoid	44	9	15	20	27.3	100.0	100.0	31.4	28.6	93.8	88.9	42.9
	Rectum	23	10	7	6	52.9	83.3	90.0	38.5	52.6	100.0	100.0	30.8
Cancer and adenoma	Cecum	14	2	1	11	33.3	100.0	100.0	45.5	50.0	100.0	100.0	72.7
	Ascending	38	10	11	17	57.1	50.0	76.2	29.4	69.6	66.7	76.2	58.8
	Transverse	7	2	2	3	75.0	66.7	75.0	66.7	75.0	66.7	75.0	66.7
	Descending	21	5	7	9	66.7	66.7	83.3	44.4	76.9	75.0	83.3	66.7
	Sigmoid	44	9	15	20	69.7	90.9	95.8	50.0	64.3	62.5	75.0	50.0
	Rectum	23	10	7	6	82.4	50.0	82.4	50.0	78.9	50.0	88.2	33.3

To total, *Ca* cancer, *Ad* adenoma, *Nor* normal, *PPV* positive-predictive value, *NPV* negative-predictive value, *Sen* sensitivity, *Spe* specificity

Table 5 Results of consensus visual interpretation based on FDG uptake pattern

Interpretation	FDG uptake pattern														
	Small localized			Large irregular			Short segmental			Long segmental			Total		
	Ca	Ad	Nor	Ca	Ad	Nor	Ca	Ad	*N*	Ca	Ad	Nor	Ca	Ad	Nor
Total	22	31	22	13	2	4	2	6	28	1	4	12	38	43	66
FDG uptake degree															
Very high	1	1	0	6	0	0	0	0	1	0	0	0	7	1	1
High	15	16	7	7	2	4	2	4	15	0	2	5	24	24	31
Moderate	3	11	13	0	0	0	0	2	9	0	1	4	3	14	26
Slight	3	3	2	0	0	0	0	0	3	1	1	3	4	4	8
Change of uptake degree															
Increased	13	25	13	9	1	2	2	2	13	1	2	8	25	30	36
Stable	7	3	8	3	1	1	0	2	7	0	2	4	10	8	20
Decreased	1	3	1	0	0	1	0	2	5	0	0	0	1	5	7
Unconfirmed	1	0	0	1	0	0	0	0	3	0	0	0	2	0	3
Change of uptake position															
Moved	2	4	0	0	0	0	0	3	10	0	0	3	2	7	13
Unmoved	20	27	22	13	2	4	2	3	18	1	4	9	36	36	53

Ca cancer, *Ad* adenoma, *Nor* normal

whether the colonic FDG uptake pattern and FDG uptake degree can be common criteria among readers remained a question. In the present study, 10 readers showed agreement, with $\kappa = 0.501$, for the FDG uptake pattern, for which the interpretation criteria are the more reliable. Confusion of the FDG uptake pattern between "short segmental" and "large irregular localized" might remain an issue. We assumed that the "short segmental" FDG uptake pattern showed FDG accumulation in the shape of the colon, while "large irregular localized" was defined as FDG accumulation beyond the regular colon shape, suggesting the presence of a large mass lesion. In these two FDG uptake patterns, FDG uptake degree was interpreted as being higher in cancer than in other conditions. Among the cases showing the "short segmental" or "large irregular localized" FDG uptake pattern, very high or high FDG uptake had a PPV of 39.0 % and a relative sensitivity of 100.0 % for detecting cancer. Thus, the degree of FDG uptake appears to have some impact on the diagnosis. The present results showed that the delayed image contributed to decreasing false-positives if the FDG uptake pattern was either the "short segmental" or "long segmental" type.

Table 6 Results of consensus visual interpretation in delayed imaging

Interpretation	Total				Exclusion (A)					Exclusion (B)				
	Ca	Ad	Nor	Total	Ca	Ad	Nor	Total	Rate (%)	Ca	Ad	Nor	Total	Rate (%)
FDG uptake patterns														
Small localized ($N = 75$)	22	31	22	75	20	26	21	67	4.5	20	27	22	69	0.0
Large irregular ($N = 19$)	13	2	4	19	12	2	3	17	25.0	13	2	4	19	0.0
Short segmental ($N = 36$)	2	6	28	36	2	2	15	19	46.4	2	3	18	23	35.7
Long segmental ($N = 17$)	1	4	12	17	1	4	9	14	25.0	1	4	9	14	25.0
FDG uptake degree														
Very high ($N = 9$)	7	1	1	9	7	1	0	8	100.0	7	1	0	8	100.0
High ($N = 79$)	24	24	31	79	23	17	24	64	22.6	24	19	27	70	12.9
Moderate ($N = 43$)	3	14	26	43	3	14	18	35	30.8	3	14	19	36	26.9
Slight ($N = 16$)	4	4	8	16	2	2	6	10	25.0	2	2	7	11	12.5

Total and after applying exclusion criteria

Exclusion (A): exclusion of decreased, unconfirmed, or moved in delayed image

Exclusion (B): exclusion of moved in delayed image

Rate (%): reduction rate of normal cases by delayed imaging

Ca cancer, *Ad* adenoma, *Nor* normal

Although the delayed image had little influence on cases classified as "large localized" on the early image, FDG uptake that did not move on the delayed image might contribute substantially to reducing false-positive cases when the FDG uptake pattern was difficult to differentiate between "short segmental" and "large irregular localized".

The measurement of the standardized uptake value (SUV) is valuable to define the degree of FDG uptake. However, SUV is affected by factors such as the size of the ROI, the reconstructed resolution of the PET image, body mass index, and serum glucose level; it appears to be difficult to define a diagnostic cut-off value for FDG uptake by SUV [13]. "Very high" or "high" FDG uptake did include the majority of cancers and adenomas. For the two classifications of "very high and high" and "moderate and slight" for FDG uptake, interobserver agreement among the 10 readers was estimated to be 0.485 (95 % CI, 0.436–0.505), which might be regarded as acceptable.

On visual analysis, the rectosigmoid colon, cecum, and proximal ascending colon show higher levels and a more variable pattern of FDG uptake compared with other colonic regions [14]. This result might be related to the present result that showed low PPV for the colon excepting the rectum.

Colonic adenomas are regarded as precursors of colorectal cancer [15], and an adenoma–carcinoma sequence is a widely accepted theory [16]. Thus, screening and removal of adenomatous polyps are significant aspects of prevention and improvement of prognosis [17]. The establishment of reliable screening tests for adenomas increases the chance of patients undergoing colonoscopic examination or surgery to remove advanced adenomas before they become malignant [18–20]. However, adenomas have a low probability of malignant transformation, taking an average of 10 years, and the probability depends on the size. Therefore, the clinical importance of screening adenomas is to detect advanced adenomas, defined as adenomas with a diameter of 10 mm or greater, or high-grade dysplasia. The usefulness of FDG-PET for colonic adenoma detection has been previously reported [5, 6, 21, 22], and the sensitivity of FDG-PET depended on the size and histologic grade of the adenoma [1, 11]. Thus, FDG-PET has a high probability for detecting adenomas that should be removed. The present results show that FDG-PET has a high relative sensitivity for either cancer or adenoma. Small and large irregular localized FDG uptake patterns were common in cancers and adenomas. Moreover, large irregular localized FDG uptake was frequently seen in cancer. The FDG uptake degree was higher in cancer than in adenoma, but they overlapped considerably in "high accumulation", which also included normal areas. These results indicate that the FDG uptake pattern of cancer has much in common with that of adenoma, resulting in difficulty differentiating between cancer and adenoma.

PET/CT provides exact localization of FDG uptake, which is more accurate than PET and/or CT [23]. PET/CT contributes to the evaluation of colonic FDG uptake by not only providing the anatomical location, but also by showing the CT finding [1, 4]. Since a large colorectal tumor can be identified by the CT finding, PET/CT might contribute to evaluation of cases showing "short segmental" or "large irregular localized" uptake. However, the CT image is not sufficient if the suspected malignant lesion is not well defined on the image [24]; thus, the findings of

characteristic colonic FDG uptake can help in developing a more accurate method.

Randomized trials have shown that FOBT reduces colorectal cancer mortality by up to 16 % [25–28]. However, the sensitivity and specificity of FOBT have remained problematic, even though FOBT is a noninvasive and cost-effective test [29]. In the present study, the result of FOBT alone scored a higher PPV and relative specificity, and a lower NPV and relative sensitivity than FDG-PET. The result of combined FDG-PET and FOBT tended to increase NPV and relative sensitivity but had lower PPV and relative specificity than the individual tests. Thus, the contribution of FOBT as an index for colorectal FDG uptake is low in terms of lowering the NPV and specificity, leading to unnecessary further testing.

Dual-time-point PET imaging has been reported to have the potential for differentiating between malignant and benign lesions, because FDG uptake in malignant lesions tends to be higher on delayed imaging than on the early imaging [30–35]. The potential for dual-time-point FDG-PET imaging for the colonic region has not been sufficiently evaluated. The present results show that delayed imaging decreased the relative sensitivity for detecting cancer in the sigmoid colon. Toriihara et al. [36] reported that physiological FDG uptake on delayed imaging was increased in the colon region, and it might have affected the present result. Moreover, the present results show that dual-time-point FDG-PET imaging does not contribute to detecting adenoma and differentiating between cancer and adenoma. FDG uptake was relatively lower for adenoma than for cancer, so that FDG uptake for adenoma may be affected by variations in physiological FDG uptake.

This study involved only subjects who had incidental colonic FDG uptake on PET examinations. Thus, the present result does not reflect on the potential of FDG-PET for screening for colon lesions, because PET-negative colon lesions were not included in the present study. In addition, since delayed images can sometimes clearly show a colon lesion regardless of being PET-negative on the early image, the present conclusions about delayed imaging are limited to lesions with incidental FDG uptake detected on an early image.

A limitation of this study was the absence of cases with symptomatic, non-neoplastic diseases, such as inflammatory enterocolitis, collagen colitis, and diverticulitis. Intraobserver variability was not examined, and a prospective study for validation should be done.

Conclusion

The present study assessed incidental colonic FDG uptake by studying observer variation. Interobserver agreement was highest for assessing FDG uptake patterns, and "small localized" and "large irregular localized" types were associated with high probabilities of colon cancer regardless of FDG uptake degree and variations in the uptake on delayed images. Although the delayed image decreased false-positive cases for some FDG uptake patterns on the early PET image, delayed imaging and also the addition of an FOBT result to FDG-PET did not have much impact on the patient diagnosis afforded by colonic FDG uptake.

Acknowledgments This work was supported in part by the National Cancer Center Research and Development Fund 23-A-25. The authors thank Dr. Hirokazu Takahashi from Yokohama City University and Dr. Yoko Miyata from National Center for Global Health and Medicine for valuable advice toward this article. The authors also thank Hiromitsu Daisaki, Ph.D., for management and evaluation of this study.

References

1. Prabhakar HB, Sahani DV, Fischman AJ, Mueller PR, Blake MA. Bowel hot spots at PET-CT. Radiographics. 2007;27:145–59.
2. Israel O, Yefremov N, Bar-Shalom R, Kagana O, Frenkel A, Keidar Z, et al. PET/CT detection of unexpected gastrointestinal foci of 18F-FDG uptake: incidence, localization patterns, and clinical significance. J Nucl Med. 2005;46:758–62.
3. Kamel EM, Thumshirn M, Truninger K, Schiesser M, Fried M, Padberg B, et al. Significance of incidental 18F-FDG accumulations in the gastrointestinal tract in PET/CT: correlation with endoscopic and histopathologic results. J Nucl Med. 2004;45: 1804–10.
4. Lee JC, Hartnett GF, Hughes BG, Ravi Kumar AS. The segmental distribution and clinical significance of colorectal fluorodeoxyglucose uptake incidentally detected on PET-CT. Nucl Med Commun. 2009;30:333–7.
5. Drenth JP, Nagengast FM, Oyen WJ. Evaluation of (pre-) malignant colonic abnormalities: endoscopic validation of FDG-PET findings. Eur J Nucl Med. 2001;28:1766–9.
6. Agress H Jr, Cooper BZ. Detection of clinically unexpected malignant and premalignant tumors with whole-body FDG PET: histopathologic comparison. Radiology. 2004;230:417–22.
7. Gutman F, Alberini JL, Wartski M, Vilain D, Le Stanc E, Sarandi F, et al. Incidental colonic focal lesions detected by FDG PET/CT. AJR Am J Roentgenol. 2005;185:495–500.
8. Even-Sapir E, Lerman H, Gutman M, Lievshitz G, Zuriel L, Polliack A, et al. The presentation of malignant tumours and premalignant lesions incidentally found on PET-CT. Eur J Nucl MedMol Imaging. 2006;33:541–52.
9. Pandit-Taskar N, Schöder H, Gonen M, Larson SM, Yeung HW. Clinical significance of unexplained abnormal focal FDG uptake in the abdomen during whole-body PET. AJR Am J Roentgenol. 2004;183:1143–7.
10. Nakajo M, Jinnouchi S, Tashiro Y, Shirahama H, Sato E, Koriyama C, et al. Effect of clinicopathologic factors on visibility of colorectal polyps with FDG PET. AJR Am J Roentgenol. 2009;92:754–60.
11. Tatlidil R, Jadvar H, Bading JR, Conti PS. Incidental colonic fluorodeoxyglucose uptake: correlation with colonoscopic and histopathologic findings. Radiology. 2002;224:783–7.
12. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.

13. Keyes JW Jr. SUV: standard uptake or silly useless value? J Nucl Med. 1995;36:1836–9.
14. Kim S, Chung JK, Kim BT, Kim SJ, Jeong JM, Lee DS, et al. Relationship between gastrointestinal F-18-fluorodeoxyglucose accumulation and gastrointestinal symptoms in whole-body PET. Clin Positron Imaging. 1999;2:273–80.
15. Bond JH. Polyp guideline: diagnosis, treatment, and surveillance for patients with colorectal polyps. Practice Parameters Committee of the American College of Gastroenterology. Am J Gastroenterol. 2000;95:3053–63.
16. Chen CD, Yen MF, Wang WM, Kim SJ, Jeong JM, Lee DS, et al. A case-cohort study for the disease natural history of adenoma–carcinoma and de novo carcinoma and surveillance of colon and rectum after polypectomy: implication for efficacy of colonoscopy. Br J Cancer. 2003;88:1866–73.
17. Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, et al. Prevention of colorectal cancer by colonoscopic polypectomy. N Engl J Med. 1993;329:1977–81.
18. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H, Chejfec G, et al. Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. N Engl J Med. 2000;343:162–8.
19. Read TE, Read JD, Butterly LF. Importance of adenomas 5 mm or less in diameter that are detected by sigmoidoscopy. N Engl J Med. 1997;336:8–12.
20. Pickhardt PJ, Choi JR, Hwang I, Butler JA, Puckett ML, Hildebrandt HA, et al. Computed tomographic virtual colonoscopy to screen for colorectal neoplasia in asymptomatic adults. N Engl J Med. 2003;349:2191–200.
21. Chen YK, Kao CH, Liao AC, Shen YY, Su CT. Colorectal cancer screening in asymptomatic adults: the role of FDG PET scan. Anticancer Res. 2003;23:4357–61.
22. Yasuda S, Fujii H, Nakahara T, Nishiumi N, Takahashi W, Ide M, et al. 18F-FDG PET detection of colonic adenomas. J Nucl Med. 2001;42:989–92.
23. von Schulthess GK. Positron emission tomography versus positron emission tomography/computed tomography: from "unclear" to "new-clear" medicine. Mol Imaging Biol. 2004;6:183–7.
24. Kostakoglu L, Hardoff R, Mirtcheva R, Goldsmith SJ. PET-CT fusion imaging in differentiating physiologic from pathologic FDG uptake. Radiographics. 2004;24:1411–31.
25. Mandel JS, Church TR, Bond JH, Ederer F, Geisser MS, Mongin SJ, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. N Engl J Med. 2000;343:1603–7.
26. Kronborg O, Fenger C, Olsen J, Jørgensen OD, Søndergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet. 1996;348:1467–71.
27. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet. 1996;348:1472–7.
28. Kramer BS, Gohagan JK, Prorok PC, editors. Cancer screening. New York: Marcel Dekker; 1999. p. 143–93.
29. Kubota K, Itoh M, Ozaki K, Ono S, Tashiro M, Yamaguchi K, et al. Advantage of delayed whole-body FDG-PET imaging for tumour detection. Eur J Nucl Med. 2001;28:696–703.
30. Hustinx R, Smith RJ, Benard F, Rosenthal DI, Machtay M, Farber LA, et al. Dual time point fluorine-18 fluorodeoxyglucose positron emission tomography: a potential method to differentiate malignancy from inflammation and normal tissue in the head and neck. Eur J Nucl Med. 1999;26:1345–8.
31. Nakamoto Y, Higashi T, Sakahara H, Tamaki N, Kogire M, Doi R, et al. Delayed 18F-fluoro-2-deoxy-D-glucose positron emission tomography scan for differentiation between malignant and benign lesions in the pancreas. Cancer. 2000;89:2547–54.
32. Lodge MA, Lucas JD, Marsden PK, Cronin BF, O'Doherty MJ, Smith MA. A PET study of 18FDG uptake in soft tissue masses. Eur J Nucl Med. 1999;26:22–30.
33. Zhuang H, Pourdehnad M, Lambright ES, Yamamoto AJ, Lanuti M, Li P, Mozley PD, et al. Dual time point 18F-FDG PET imaging for differentiating malignant from inflammatory processes. J Nucl Med. 2001;42:1412–7.
34. Matthies A, Hickeson M, Cuchiara A, Alavi A. Dual time point 18F-FDG PET for the evaluation of pulmonary nodules. J Nucl Med. 2002;43:871–5.
35. Kumar R, Loving VA, Chauhan A, Zhuang H, Mitchell S, Alavi A. Potential of dual-time-point imaging to improve breast cancer diagnosis with 18F-FDG PET. J Nucl Med. 2005;46:1819–24.
36. Toriihara A, Yoshida K, Umehara I, Shibuya H. Normal variants of bowel FDG uptake in dual-time-point PET/CT imaging. Ann Nucl Med. 2001;25:173–8.

Ann Nucl Med (2013) 27:225–232
DOI 10.1007/s12149-012-0676-2

ORIGINAL ARTICLE

Standardization of image quality across multiple centers by optimization of acquisition and reconstruction parameters with interim FDG-PET/CT for evaluating diffuse large B cell lymphoma

Hiromitsu Daisaki · Ukihide Tateishi · Takashi Terauchi · Mitsuaki Tatsumi · Kazufumi Suzuki · Naoki Shimada · Hiroyuki Nishida · Akihiko Numata · Koji Kato · Koichi Akashi · Mine Harada

Received: 11 September 2012 / Accepted: 4 December 2012 / Published online: 22 December 2012

Abstract
Objective A multicenter trial is currently underway using FDG-PET/CT to evaluate diffuse large B cell lymphoma in Japan (JSCT NHL10). Standardization of the image quality between the participating centers is a fundamental aspect of the study. Within the framework of JSCT NHL10, standardization of the image quality was attempted by optimizing the acquisition and reconstruction conditions using mid-therapy FDG-PET/CT for diffuse large B cell lymphoma. This report describes the procedures and results of this attempt.
Methods The acquisition protocols and imaging quality were initially determined at each center and again after modification. The image quality was based on performance with an ^{18}F-filled National Electrical Manufacturers Association standards body phantom. We determined that the acquisition duration and reconstruction parameters of each scanner evaluated were in compliance with the

Present Address:
H. Daisaki (✉)
Nihon Medi-Physics Co., Ltd., Tokyo, Japan
e-mail: hdaisaki@gmail.com

H. Daisaki · T. Terauchi · N. Shimada
Screening Technology and Development Division,
Research Center for Cancer Prevention and Screening,
National Cancer Center, Tokyo, Japan

T. Terauchi
e-mail: tterauch@ncc.go.jp

N. Shimada
e-mail: nashimad@ncc.go.jp

U. Tateishi
Department of Radiology, Yokohama City University
Graduate School of Medicine, Yokohama, Japan
e-mail: utateish@yokohama-cu.ac.jp

M. Tatsumi
Department of Radiology,
Osaka University Graduate School of Medicine,
Suita, Osaka, Japan
e-mail: m-tatsumi@radiol.med.osaka-u.ac.jp

K. Suzuki
Department of Radiology, Dokkyo Medical University Hospital,
Tochigi, Japan
e-mail: kazufumi@dokkyomed.ac.jp

H. Nishida
Division of Molecular Imaging,
Institute of Biomedical Research and Innovation, Kobe, Japan
e-mail: nishida@fbri.org

A. Numata · K. Kato · K. Akashi
Department of Medicine and Biosystemic Science,
Kyushu University Graduate School of Medical Sciences,
Fukuoka, Japan
e-mail: akihiko@intmed1.med.kyushu-u.ac.jp

K. Kato
e-mail: kojikato@intmed1.med.kyushu-u.ac.jp

K. Akashi
e-mail: akashi@med.kyushu-u.ac.jp

M. Harada
National Hospital Organization, Omuta Hospital,
Omuta, Japan
e-mail: mharada@karatsu.saga.med.or.jp

M. Harada
Japan Study Group for Cell Therapy and Transplantation,
Tochigi, Japan

Japanese guideline for the oncology FDG-PET/CT data acquisition protocol: synopsis of Version 1.0 (the Guideline) based on the results of the phantom experiments performed by the Core Laboratory.

Results A total of 18 centers (19 scanners) participated in this trial. The center's default protocol was unchanged for 9 scanners (47.4 %) and changed for 10 scanners (52.6 %). Both acquisition duration and reconstruction parameters were changed in 3 (15.8 %) of 10 scanners and the acquisition duration alone was changed in 7 (36.8 %) scanners. Also, the accuracy of the standardized uptake value (SUV) was evaluated with the acceptable level 1.0 ± 0.1, resulting in readjustment and recalibration in 2 scanners (10.5 %), which were confirmed to attain the acceptable accuracy after the required readjustment. As of August 2012, 21 patients have undergone an FDG-PET/CT examination under the acquisition protocols determined by the Core Laboratory. Evaluation of the image quality using several physical parameters confirmed that the accumulated data were of sufficient quality.

Conclusions Optimization of the acquisition protocol, in compliance with the guideline, was successfully achieved by the Core Laboratory in the framework of JSCT NHL10 to accumulate equivalent quality data across multiple centers. The progress of the trial was greatly facilitated by support from the Japan Society of Nuclear Medicine Working Group for Investigation of Response Evaluation Criteria in Malignant Tumors Using Standardized PET/CT (Principal Investigator: Ukihide Tateishi, MD., PhD).

Keywords FDG · PET/CT · Diffuse large B cell lymphoma · Standardization · Multicenter study

Introduction

Positron emission tomography (PET) using 2-fluorine-18-fluoro-2-deoxy-D glucose (FDG) is an indispensable imaging modality for diagnosing malignant lymphoma. FDG-PET imaging has been reported to be effective in staging and evaluating the response to treatment. Its effectiveness in monitoring post-treatment recurrence/exacerbation and evaluating the prognosis at mid-therapy was recently investigated in a clinical trial [1].

FDG-PET is performed during treatment of diffuse large B cell lymphoma (DLBCL) because the response to treatment revealed by FDG-PET images can reflect prognosis. If FDG-PET images show a poor response after treatment, the planned treatment can potentially be altered to improve the outcome [2–4]. On the basis of this hypothesis, several stratified clinical trials were performed on patients with aggressive lymphoma with a poor prognosis. Aggressive chemotherapy or high-dose chemotherapy in combination with autotransplantation was employed when FDG-PET revealed a poor response to treatment [5, 6].

However, the data from several reports do not support the correlation between mid-therapy FDG-PET observations and prognosis [7]. One of the possible reasons for the negative results may be related to the details of FDG-PET data acquisition and assessment [8].

In Japan, a multicenter trial, "Japan Study Group for Cell Therapy and Transplantation (JSCT) NHL 10", is underway to investigate the effectiveness and safety of a stratified treatment protocol, in which high-risk CD20 positive DLBCL patients undergo two courses of R-CHOP chemotherapy before evaluation by FDG-PET. Patients with poor response to treatment (positive cases) receive high-dose chemotherapy followed by autologous stem cell transplantation (PET positive treatment) while patients with sufficient response to treatment (negative patients) receive eight courses of R-CHOP chemotherapy (PET negative treatment).

FDG-PET image quality depends on many factors including scanner type, administered dose, acquisition protocol, and image reconstruction parameters. Therefore, the quality of the PET images between multiple centers can only be guaranteed through standardization of the imaging parameters and appropriate quality control (QC) of the scanners. In the US, the Clinical Trial Network (CTN) organized by the Society of Nuclear Medicine (SNM) is attempting to standardize the synthesis of PET radiopharmaceuticals and the scanning protocol [9]. In addition to the CTN, the American College of Radiology Imaging Network (ACRIN) [10] is active in the field and the European Association of Nuclear Medicine (EANM) operates a similar network [8]. However, a network that focuses on standardizing image quality through scanning protocols or QC of the acquired data is still under development in Japan [11].

Japanese guideline for the oncology FDG-PET/CT data acquisition protocol: synopsis of Version 1.0 (the Guideline) was established [12], though there is no Core Laboratory to promote standardization that is comparable to those in the US and Europe. As a result, we established a trial-specific Core Laboratory for JSCT NHL10 comprising two board certified PET nuclear medicine physicians and three board certified nuclear medicine technologists. The purpose of the laboratory was to standardize FDG-PET image quality and quantification on the basis of the guideline, as well as to evaluate the validity of the standardization. This report describes the procedure and result of this attempt. The Japan Society of Nuclear Medicine Working Group for Investigation of Response Evaluation Criteria in Malignant Tumors Using Standardized PET/CT

(Principal Investigator: Ukihide Tateishi, MD., PhD) jointly conducted this study.

Materials and methods

Accreditation of the PET/CT scanners and determination of data acquisition protocol

The Core Laboratory was utilized to determine if FDG-PET data were acquired under appropriate conditions and if appropriate QC was implemented for the scanner. To ascertain the validity of the acquisition protocol and image reconstruction parameters, phantom experiments were conducted in compliance with the guideline. The investigated items were: (1) scanner manufacturer and model type; (2) frequency and content of QC; (3) availability of FDG; (4) injection method; (5) injected dose; (6) uptake duration; (7) data acquisition mode; (8) scan duration; (9) availability of list mode; (10) image reconstruction method; (11) image reconstruction parameters; (12) history of phantom experiments; and (13) evaluation of phantom experiment results. If possible, centers provided existing phantom experimental data to the Core Laboratory. In centers with no existing phantom experimental data, center's staff carried out the experiment and/or Core Laboratory members visited the centers to provide technical support during acquisition of phantom experiment data.

The National Electrical Manufacturers Association (NEMA) standards body phantom [13] and ^{18}F-solution were used as part of experiments in which PET data on a simulated torso of the human body were acquired with varied scan duration. The obtained images were both physically and visually evaluated. Phantom background activity concentration was controlled at 2.65 kBq/ml and the hot sphere activity was controlled at four times the background activity. Emission scan duration was 12 min with three-dimensional (3D) list-mode acquisition. When the list mode was not available, the default acquisition duration used in the clinical PET/CT examinations was used to acquire the data. Computed tomography (CT) for attenuation correction was performed with default acquisition parameters of each center. List-mode data were histogramed into a sinogram of 1–10 min/bed position and the image was reconstructed by the center's default method and parameters.

Phantom data were stored in the DICOM format and submitted to the Core Laboratory. PET images submitted to the Core Laboratory were viewed with the viewer OsiriX MD for physical analysis. Then, the phantom noise equivalent count ($NEC_{phantom}$), visualization of the 10-mm sphere (visual score), image noise (N_{10mm}), % contrast ($Q_{H,10mm}/N_{10mm}$), and relative recovery coefficient (RC) were evaluated and compared with the reference values recommended in the Guideline [12]. Visualization of the 10-mm sphere was evaluated in the Core Laboratory. On the basis of these analyses of phantom data, the imaging conditions recommended for the clinical trial were determined for each scanner.

Although the guideline does not provide any recommendations regarding the accuracy of the SUV, the values were obtained for 12 regions of interest (ROIs) defined in the background area of a phantom image with a 30-min scan duration to evaluate accuracy with an acceptable level of 1.0 ± 0.1 [10].

Quality assurance of the patient data

As of August 2012, 21 patients have undergone FDG-PET examination with the acquisition conditions determined on the basis of phantom experiments (Table 1). The institutional review board approved the investigations in each center. All patients gave informed consents. The patient group comprised 11 males and 10 females with a mean height of 160.7 ± 9.6 cm (range 146.4–185.0 cm), a mean weight of 50.5 ± 11.2 kg (range 31.6–73.0 kg), and a body mass index (BMI) of 19.4 ± 3.0 kg/m^2 (range 13.2–25.0 kg/m^2). The mean administered dose was 270.5 ± 38.4 MBq (range 211.0–324.8 MBq), while uptake duration was 60.3 ± 2.6 min (54.0–64.0 min), and blood glucose level was 103.0 ± 10.6 mg/dl (89.0–117.0 mg/dl). Patient image data were submitted to the Core Laboratory immediately after examination to check for the presence of artifacts attributable to the scanner, body movement during examination, or poor registration of the PET and CT images. The patient noise equivalent count ($NEC_{patient}$), patient noise equivalent count density ($NEC_{density}$), and mean signal-to-noise ratio (SNR) within the liver ROI (liver SNR) were determined to establish if the patient data were in compliance with the Guideline.

Results

A total of 18 centers (19 scanners) were enrolled in this trial (Table 1). Only one center had conducted phantom experiments to determine the scan and image reconstruction conditions before participation in this trial. The other 17 centers carried out the phantom experiments especially for this trial. One of the centers performed the phantom experiment independently while the others had support from the Core Laboratory members who performed the phantom experiments and obtained the data. The following PET/CT models were used: General Electric—3 Discovery STs, 3 Discovery STEPs, 4 Discovery STEs, and 1 Discovery 690; Siemens—4 Biographs, 1 Biograph True

Table 1 List of center, prefecture, scanner and number of patients

No.	Center	Prefecture	Scanner	Number of patients
1	Anjo Kosei Hp.	Aichi	Biograph 16 (Siemens)	
2	Central CI clinic Hp.	Hokkaido	Discovery STEP (GE)	
3	Gifu Univ. Hp.	Gifu	Biograph 16 (Siemens)	2
4	Ehime Pref. Central Hp.	Ehime	Discovery STE (GE)	
5	Kusunoki Hp.	Gunma	Discovery ST (GE)	3
6	National Tokyo Medical Center	Tokyo	Biograph 40 (Siemens)	
7	Sapporo Goryokaku Hp.	Hokkaido	Discovery STEP (GE)	
8	Sapporo Medical Univ. Hp.	Hokkaido	Discovery STEP (GE)	1
9	Kameda Clinic Hp.	Kanagawa	Discovery STE (GE)	
10	Toranomon Hp.	Tokyo	Aquiduo (Toshiba)	
11	Kawasaki Medical Univ. Hp.	Okayama	Discovery STE (GE)	5
12	Kyushu Univ. Hp.	Fukuoka	Discovery STE (GE)	6
13	Kyushu Univ. Hp.	Fukuoka	Biograph mCT (Siemens)	
14	Oike Clinic	Kyoto	Biograph 64 (Siemens)	
15	Izumo City Hp.	Shimane	True Point Biograph 64 (Siemens)	
16	Kurume Univ. Hp.	Fukuoka	GEMINI GX-L (Philips)	4
17	Kanazawa Advanced Medical Center	Kanazawa	Discovery 690 (GE)	
18	Takinomiya General Hp.	Kagawa	Discovery ST (GE)	
19	Kusatsu General Hp.	Shiga	Discovery ST (GE)	

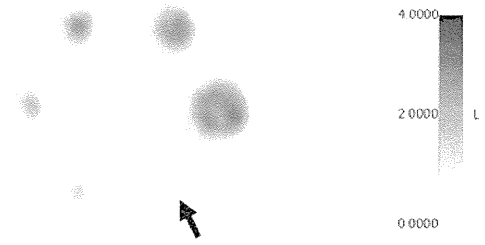

Fig. 1 Example of a phantom image that could not visualize a 10-mm hot sphere with a SUV = 4

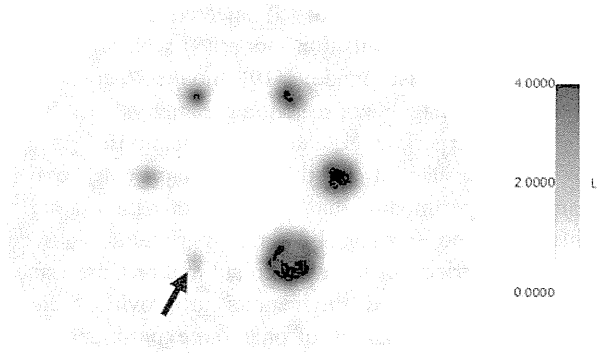

Fig. 2 Example of a phantom image that visualized a 10-mm hot sphere with a SUV = 4

Point, 1 Biograph mCT; Philips—1 GEMINI GX-L; and Toshiba Medical—1 Aquiduo.

Analysis by the Core Laboratory confirmed visualization of a 10-mm hot sphere in 9 scanners (47.4 %) with the respective centers' own clinical default settings without a need for modification. Examples of the phantom images that could and could not visualize the 10-mm hot sphere are shown in Figs. 1 and 2, respectively. Three scanners required modification of the scan duration and image reconstruction parameters (15.8 %), and 7 scanners

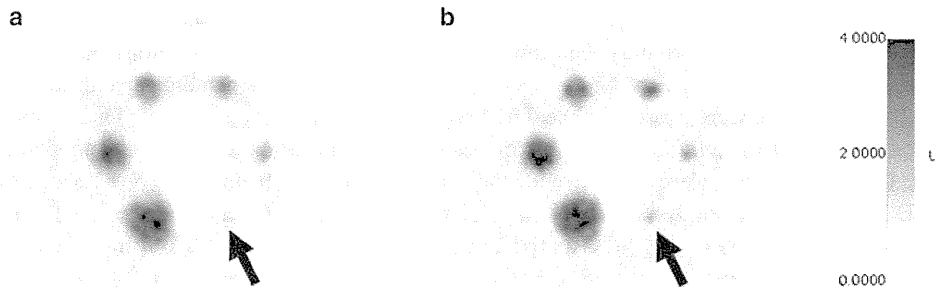

Fig. 3 Phantom images that **a** initially could not visualize a 10-mm hot sphere with a SUV = 4 and **b** visualized the hot sphere after the conditions were revised. The protocol revisions included changes in the scan duration from 2 min to 4 min/bed and image reconstruction parameters from 2 iterations to 3 iterations

Table 2 List of scanners, scan durations and injected dose

Scanner	Scan duration (min)		Injected dose (MBq/kg or MBq)	
	Initial	Revision	Initial	Revision
Biograph Sensation16	2	3	3.0 MBq/kg	3.7 MBq/kg
Discovery STEP	2.43	NC	4.5 MBq/kg	NC
Biograph 16	2	4	185 MBq (deliver FDG)	NC
Discovery STE	3	NC	3.0 MBq/kg	3.7 MBq/kg
Discovery ST	2	4	185 MBq (deliver FDG)	NC
Biograph 40	3	4	185 MBq (deliver FDG)	NC
Discovery STEP	3	NC	185 MBq (<50 kg), 220 MBq (50–80 kg), 260 MBq (>80 kg)	3.7 MBq/kg
Discovery STEP	3	NC	185 MBq (deliver FDG)	NC
Discovery STE	3	NC	4.3 MBq/kg	NC
Aquiduo	2	4	185 MBq (deliver FDG)	NC
Discovery STE	2	NC	185 MBq (deliver FDG)	NC
Discovery STE	3	NC	185 MBq (deliver FDG)	NC
Biograph mCT	2	NC	185 MBq (deliver FDG)	NC
Biograph 64	2	3	4.5 MBq/kg	NC
True Point Biograph 64	2	3	185 MBq (deliver FDG)	NC
GEMINI GX-L	2	4	4.4 MBq/kg	NC
Discovery 690	2	NC	4.0 MBq/kg	NC
Discovery ST	2.5	4	185 MBq (deliver FDG)	NC
Discovery ST	2	3	3.0 MBq/kg	3.7 MBq/kg

The initial and post-revision scan durations and injected doses are shown

NC No revision parameters

required modification of the scan duration (36.8 %). A phantom image that could not visualize the 10-mm hot sphere before the scanning protocol was modified is shown in Fig. 3a. Figure 3b shows an image that visualized the hot sphere after modification of the scanning protocol. No serious artifacts were observed in the PET images produced by any of the scanners.

Table 3 List of scanners, image reconstruction methods and parameters

Scanner	Image reconstruction		
	Method	Initial parameter	Revision parameter
Biograph 16	FORE-OSEM	iter:3, sub:8, filter: 5 mm, matrix:128	NC
Discovery STEP	VUE-POINT+	iter:2, sub:21, filter: 5.14 mm, matrix:128	NC
Biograph 16	FORE-OSEM	iter:2, sub:8, filter: 5 mm, matrix:256	iter:3, sub:8, filter: 5 mm, matrix:256
Discovery STE	VUE-POINT+	iter:2, sub:20, filter: 4.29 mm, matrix:128	NC
Discovery ST	FORE-OSEM	iter:5, sub:16, filter: 5.14 mm, matrix:128	NC
Biograph 40	FORE-OSEM	iter:4, sub:8, filter: 5 mm, matrix:168	NC
Discovery STEP	VUE-POINT+	iter:2, sub:21, filter: 6 mm, matrix:128	NC
Discovery STEP	VUE-POINT+	iter:2, sub:21, filter: 5.14 mm, matrix:128	NC
Discovery STE	VUE-POINT+	iter:2, sub:21, filter: 6 mm, matrix:128	NC
Aquiduo	FORE-OSEM	iter:4, sub:14, filter: 8 mm, matrix:128	NC
Discovery STE	VUE-POINT+	iter:2, sub:20, filter: 5.14 mm, matrix:128	NC
Discovery STE	VUE-POINT+	iter:2, sub:28, filter: 6 mm, matrix:128	NC
Biograph mCT	TrueX with TF	iter:2, sub:21, filter: 6 mm, matrix:256	NC
Biograph 64	FORE-OSEM	iter:3, sub:8, filter: 6.5 mm, matrix:168	iter:4, sub:8, filter: 6.5 mm, matrix:256
True Point Biograph 64	TrueX	iter:3, sub:21, filter: 4 mm, matrix:168	NC
GEMINI GX-L	LOR-RAMLA	smooth	HQ mode, sharp
Discovery 690	VUE-POINT FX	iter:2, sub:18, filter: 5 mm, matrix:128	NC
Discovery ST	FORE-OSEM	iter:4, sub:16, filter: 5.14 mm, matrix:128	NC
Discovery ST	FORE-OSEM	iter:5, sub:16, filter: 5.14 mm, matrix:128	NC

The initial and post-revision image reconstruction parameters are shown

NC No revision parameters

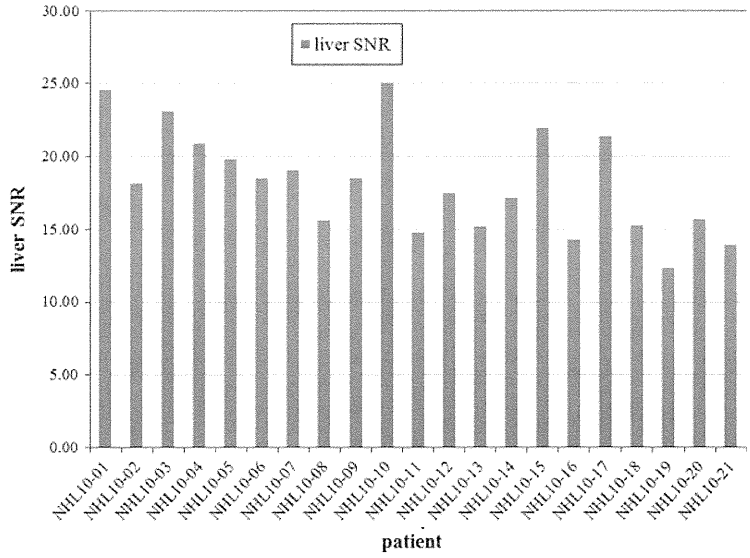

Fig. 4 Liver SNR in 21 cases

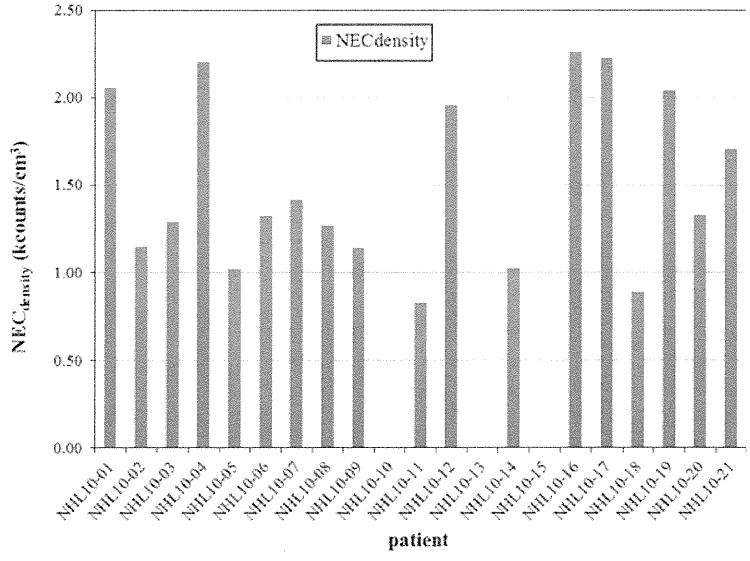

Fig. 5 $NEC_{density}$ in 21 cases

SUV was out of the acceptable range of 1.0 ± 0.1 in 2 scanners (10.5 %). Therefore, maintenance services including normalization, adjustment of gain value for the detector, QA of the dose calibrator, and recalibration were performed to ensure SUV accuracy. After this procedure, phantom data were acquired again and the SUV accuracy was confirmed to be within the acceptable range for this trial.

10-mm spheres were visualized in 3.1 ± 0.8 min on average (range 1–4 min). The mean physical indexes at that time were: $NEC_{phantom}$ = 7.7 ± 2.4 Mcounts (range 4.6–14.2 Mcounts), N_{10mm} = 6.0 ± 1.3 % (range 4.3–9.3 %), $Q_{H,10mm}$ = 16.7 ± 4.3 % (range 10.4–24.2 %), $Q_{H,10mm}/N_{10mm}$ = 2.8 ± 0.7 (range 1.8–4.1). The mean relative RC for 10-mm sphere was 0.44 ± 0.05 (range 0.38–0.55). The mean background SUV was 0.97 ± 0.04 (range 0.92–1.05). The injected dose and scan duration of each scanner are shown in Table 2, and reconstruction methods with the parameters are shown in Table 3 with or without revision.

No significant artifact was observed within any of the acquired patient images and compliance with the protocol was confirmed with the header data in each PET image. With the mean $NEC_{patient}$ = 46.8 ± 13.8 Mcounts/m (range 31.4–75.5 Mcounts/m), mean $NEC_{density}$ = 1.21 ± 0.45 kcounts/cm^3 (range 0.69–2.21 kcounts/cm^3), and mean liver SNR = 19.3 ± 3.1 (range 14.8–25.1), the physical indexes were in the reference ranges recommended in the Guideline (Figs. 4, 5, 6). Count data were not recorded in the DICOM header in three cases for which the liver SNR was used for confirmation. Typical clinical images with $NEC_{density}$ = 0.86 (kcounts/cm^3),

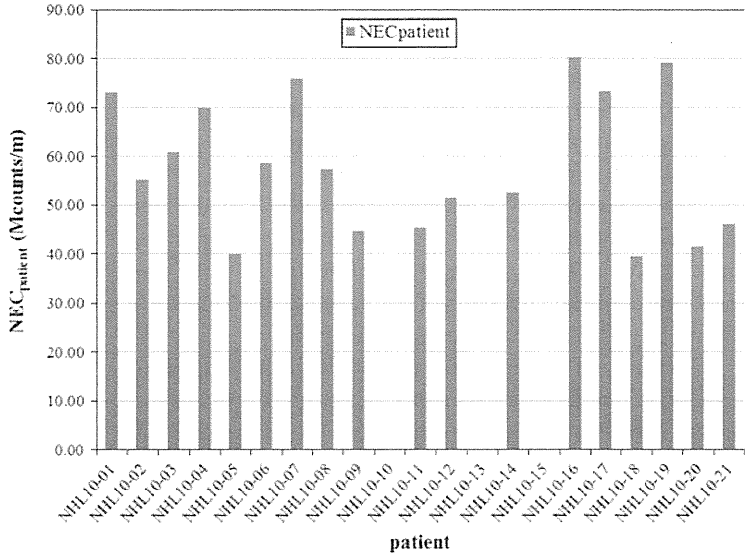

Fig. 6 $NEC_{patient}$ in 21 cases

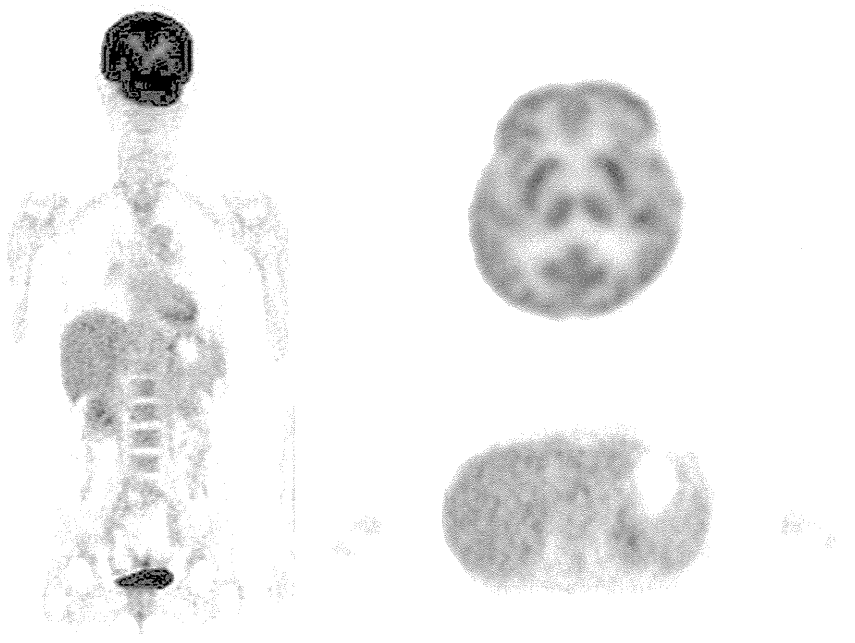

Fig. 7 Clinical FDG-PET images (*left* coronal, *upper right* brain transverse, *lower right* liver transverse)

$NEC_{patient}$ = 33.7 (Mcounts/m), and liver SNR = 18.5 are shown in Figs. 6, 7.

Discussion

We attempted to standardize FDG-PET image quality by optimizing the scan duration and reconstruction parameters on the basis of the guideline. Phantom experiments involved determining the acquisition/reconstruction parameters by scanner model and then verifying the validity of the conditions. Centers that already had default protocols only required the validation step. Though all of the participating centers in this trial had already performed clinical FDG-PET/CT examinations, we acquired the data for this trial using the list mode during validation in order to evaluate the data with varied acquisition durations. This procedure allowed for the simulation of PET images with varied acquisition durations, which enabled determination of optimal acquisition durations with only one phantom experiment. This was necessary when the default scanning protocol did not meet the recommendation of the guideline. However, multiple phantom experiments with varied acquisition durations

were required when utilizing PET/CT scanners with no list mode available.

With the acquisition duration that permitted visualization of the 10-mm hot sphere in the phantom experiment, the mean $NEC_{phantom}$ and N_{10mm} values were below the values recommended in the guideline while the mean $Q_{H,10mm}/N_{10mm}$ value of 2.81 was far beyond the recommended value. Visualization of a lesion depends on the correlation between image noise and contrast or $Q_{H,10mm}/N_{10mm}$. Therefore, the scanning protocol determined on the basis of the phantom experiments in this trial can guarantee sufficient image quality.

Although not included in the guideline, this trial evaluated SUV (calibration) accuracy. The Core Laboratory of the ACRIN uses a cylindrical phantom for 18F or 68Ge to evaluate the calibration accuracy within 10 % or less than 10 %. In this trial, we applied the same allowance range for calibration evaluation and determined that two scanners (10.2 %) did not meet the allowance range. Our phantom shape was different from ACRIN's but the influence should be minor. One of the two scanners required an adjustment and the other required an adjustment of the dose calibrator. Both scanners attained the acceptable accuracy after recalibration. This indicated that the phantom experiment was useful for identifying the physical error or calibration process error of the scanner.

Phantom experiments resulted in prolonged data acquisition duration in 10 centers. As of August 2012, 21 patients had undergone FDG-PET examination for this trial. The QA parameters $NEC_{patient}$, $NEC_{density}$, and liver SNR were all beyond the levels recommended in the guideline, suggesting that the accumulated image data have sufficient quality for the trial.

While there is currently (August 2012) no Core Laboratory representing the entire academic community in Japan, this trial was directed by the special Core Laboratory established for JSCT NHL10 and was facilitated by the collaboration of the Japan Society of Nuclear Medicine Working Group members (principal investigator: Ukihide Tateishi, MD., PhD).

The phantom experiments were performed based on the guideline to optimize the acquisition protocol and standardize the image quality used in the multicenter trial to assess mid-therapy FDG-PET/CT. Phantom experiments enabled accumulation of sufficient quality image data across multiple centers.

Acknowledgments This work was supported in part by Grand-in-Aid for Research Foundation for Community Medicine, Japan. The authors would like to thank Sone Teruki (Kawasaki Medical Univ. Hp.), Masatoshi Ishibashi (Kurume Univ. Hp.), Kiyosumi Maeda (Kusatsu General Hp.), Naoya Yama (Sapporo Medical Univ. Hp.), Eriko Tsukamoto (Central CI Clinic), Tomohito Kaji (Sapporo Goryokaku Hp.), Toshihiro Ouchi (Kameda General Hp.), Yukio Funatsu (Kusunoki Hp.), Shunji Okae (Anjo Kosei Hp.), Masayuki Kanematsu (Gifu Univ. Hp.), Eiji Tadamura (Oike Clinic), Shinsuke Matsuno (Takinomiya General Hp.), Hirotaka Maruno (Toranomon Hp.), Hiroyuki Kuroda (Izumo city Hp.), Hitoshi Miki (Ehime Prefectural Hp.), Kazuhito Toya (National Tokyo Medical Center), Koichi Abe (Kyushu Univ. Hp.), Seigo Kinuya (Kanazawa Univ. Hp.) and Takafumi Mochizuki (Kanazawa Advanced Medical Center) for performing FDG-PET/CT examinations.

Conflict of interest Daisaki is employed in Nihon Medi-Physics Co., Ltd., Japan, from April 2012 which delivers 18F-FDG although this study was almost performed in National Cancer Center, Japan.

References

1. Seam P, Juweid ME, Cheson BD. The role of FDG-PET scans in patients with lymphoma. Blood. 2007;110(10):3507–16.
2. Spaepen K, Stroobants S, Dupont P, Vandenberghe P, Thomas J, de Groot T, et al. Early restaging positron emission tomography with (18)F-fluorodeoxyglucose predicts outcome in patients with aggressive non-Hodgkin's lymphoma. Ann Oncol. 2002;13(9):1356–63.
3. Gallamini A, Rigacci L, Merli F, Nassi L, Bosi A, Capodanno I, et al. The predictive value of positron emission tomography scanning performed after two courses of standard therapy on treatment outcome in advanced stage Hodgkin's disease. Haematologica. 2006;91(4):475–81.
4. Casasnovas RO, Meignan M, Berriolo-Riedinger A, Bardet S, Julian A, Thieblemont C, et al. SUVmax reduction improves early prognosis value of interim positron emission tomography scans in diffuse large B-cell lymphoma. Blood. 2011;118(1):37–43.
5. Haioun C, Itti E, Rahmouni A, Brice P, Rain JD, Belhadj K, et al. [18F]fluoro-2-deoxy-D-glucose positron emission tomography (FDG-PET) in aggressive lymphoma: an early prognostic tool for predicting patient outcome. Blood. 2005;106(4):1376–81.
6. Kasamon YL, Wahl RL, Ziessman HA, Blackford AL, Goodman SN, Fidyk CA, et al. Phase II study of risk-adapted therapy of newly diagnosed, aggressive non-Hodgkin lymphoma based on midtreatment FDG-PET scanning. Biol Blood Marrow Transplant. 2009;15(2):242–8.
7. Pregno P, Chiappella A, Bellò M, Botto B, Ferrero S, Franceschetti S, et al. Interim 18-FDG-PET/CT failed to predict the outcome in diffuse large B-cell lymphoma patients treated at the diagnosis with rituximab-CHOP. Blood. 2012;119(9):2066–73.
8. Zijlstra JM, Boellaard R, Hoekstra OS. Interim positron emission tomography scan in multi-center studies: optimization of visual and quantitative assessments. Leuk Lymphoma. 2009;50(11):1748–9.
9. Kurihara C. Report of the participation in the US Society of Nuclear Medicine Clinical Trial Network Summit meeting. Clin Eval. 2010;38(3):623–8.
10. Scheuermann JS, Saffer JR, Karp JS, Levering AM, Siegel BA. Qualification of PET scanners for use in multicenter cancer clinical trials: The American College of Radiology Imaging Network Experience. J Nucl Med. 2009;50(7):1187–93.
11. http://www.jsnm.org/japanese/11-12-13. Accessed 31 Jul 2012.
12. Fukukita H, Senda M, Terauchi T, Suzuki K, Daisaki H, Matsumoto K, et al. Japanese guideline for the oncology FDG-PET/CT data acquisition protocol: synopsis of Version 1.0. Ann Nucl Med. 2010;24(4):325–34.
13. National Electrical Manufacturers Association. NEMA Standards Publication NU 2-2007: performance measurement of positron emission tomographs. Rosslyn, VA: National Electrical Manufacturers Association; 2007.

Hindawi Publishing Corporation
Pulmonary Medicine
Volume 2012, Article ID 196173, 5 pages
doi:10.1155/2012/196173

Review Article

Lung Radiofrequency Ablation: Potential as a Therapy to Oligometastasis and Oligo-Recurrence

Takao Hiraki and Susumu Kanazawa

Department of Radiology, Okayama University Medical School, 2-5-1 Shikatocho, Kitaku, Okayama 700-8558, Japan

Correspondence should be addressed to Takao Hiraki, takaoh@tc4.so-net.ne.jp

Received 13 August 2012; Accepted 1 October 2012

Academic Editor: Yuzuru Niibe

The early results (e.g., patient survival) of RFA for the treatment of patients with NSCLC and pulmonary metastasis from various primary lesions including colorectal cancer, lung cancer, hepatocellular carcinoma, renal cell carcinoma, and sarcoma appear encouraging and suggest the potential to offer long-term survival for the patients with oligo-recurrence or oligometastasis of lung cancer. The usefulness of RFA for oligo-recurrence or oligometastasis of lung cancer should be clarified by prospective studies in the future.

1. Introduction

Primary lung cancer is the most common malignancy and the leading cause of death from cancer worldwide. In addition, the lungs are the second most frequent site of metastasis from extrathoracic cancers and the only site of metastasis in 20% of such cases. Surgical resection is the first-line treatment for nonsmall-cell lung cancers (NSCLC) and offers the best treatment opportunity. Surgery is also accepted as a treatment option for carefully selected patients with metastatic lung cancer. However, surgical resection is not suitable for many patients mainly because of the advanced stage of cancer, compromised lung function, and/or comorbidities. Although chemotherapy, radiation therapy, or a combination of these serves as alternative treatments for such patients, complete remission of the disease is rarely achieved. Therefore, research that focused on alternative therapies for lung cancer has been extensive in the past decades; such therapies include stereotactic radiation therapy, cryoablation, laser ablation, and radiofrequency (RFA).

RFA causes focal coagulation necrosis of tissue by delivery of energy in the form of an alternating electrical current with a frequency of 460 to 500 kHz in the range of radio waves. The location of the ablative effect is determined by the precise placement of the radiofrequency electrode, usually using imaging guidance. The radiofrequency electrical current is concentrated near the noninsulated tip of the electrode, and the circuit is completed by returning either to electrical grounding pads usually located on the patient's thighs. The alternating electrical current causes ionic dipolar molecules in surrounding tissue and fluids to agitate, resulting in frictional heating that is greatest adjacent to the noninsulated portion of the electrode. The heat energy is then distributed radially to surrounding tissues. When radiofrequency current is applied in a slow, controlled fashion, the tissue heating is local, typically ellipsoid in shape, and predictable in distribution.

At first, RFA was noted as a therapy for hepatocellular carcinoma. The favorable outcomes of the RFA in the liver have encouraged the application of this technique to cancer in other organs. In 2000, Dupuy et al. [1] firstly reported clinical application of this technique in the lung. Since then, RFA has been gaining popularity rapidly as a treatment of lung cancer. RFA of lung cancer is usually performed under CT-guidance and the techniques are quite simple and similar to those used for CT-guided lung biopsy. Herein, we review clinical outcomes of RFA of lung cancer and discuss the potential to be used as a therapy to oligometastasis and oligo-recurrence.