PERGAMON

European Journal of Cancer 36 (2000) 858–867

European Journal of Cancer

www.ejconline.com

Experience in qualitative and quantitative FDG PET in follow-up of patients with suspected recurrence from head and neck cancer

M. Lapela [a,b,*], A. Eigtved [c], S. Jyrkkiö [a,b], R. Grénman [d], T. Kurki [e], P. Lindholm [a,b], J. Nuutinen [a,b], E. Sutinen [a,b], O. Solin [b], I. Bjornskov [f], P. Bretlau [g], L. Friberg [c], S. Holm [c], M. Jensen [c], H. Sand Hansen [h], H. Minn [a,b]

[a] *Department of Oncology and Radiotherapy, University of Turku, University Central Hospital, Turku, Finland*
[b] *Turku PET Centre, University of Turku, University Central Hospital, Turku, Finland*
[c] *Cyclotron and PET Unit, Department of Clinical Physiology & Nuclear Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark*
[d] *Department of Otorhinolaryngology, University of Turku, University Central Hospital, Turku, Finland*
[e] *Medical Imaging Centre, University of Turku, University Central Hospital, Turku, Finland*
[f] *Department of Radiology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark*
[g] *Department of Otorhinolaryngology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark*
[h] *Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark*

Received 14 July 1999; received in revised form 16 December 1999; accepted 31 January 2000

Abstract

We evaluated positron emission tomography (PET) with 2-[fluorine-18]fluoro-2-deoxy-D-glucose (FDG) in the detection of recurrent head and neck cancer, and compared visual and quantitative interpretation of PET images for their accuracy in the identification of tumour recurrence. Sixty-two FDG PET studies were performed in 56 patients having a total of 81 lesions, which were clinically suspected for recurrent carcinoma of the head and neck. The PET images were interpreted visually, and tracer uptake was quantitated as the standardised uptake value adjusted to body weight (SUV). Sensitivity of visual interpretation of the PET images for the presence of malignancy ranged from 84 to 95%, and specificity from 84 to 93%, respectively, depending on the selected scheme for grading of the lesions. Malignant lesions accumulated significantly more FDG than the benign ones (the median SUVs were 6.8 and 3.3, respectively, $P < 0.001$). However, there was a wide overlap of the FDG uptake values between these two groups. Hence, the highest accuracy of quantitative analysis in correct identification of tumour recurrence (75% at Receiver Operating Curve analysis) was inferior to that of visual analysis (89%). FDG PET is feasible for the detection of recurrent head and neck cancer. Although quantitation of FDG uptake using SUV shows significantly higher tracer concentrations for malignant than benign lesions, the wide overlap of individual SUVs between these two groups is a serious concern in diagnostic evaluation. Therefore, in clinical practice it may be preferable to identify the presence of tumour recurrence within this patient group by qualitative interpretation of the PET images.

Keywords: Positron emission tomography; Head and neck cancer; Recurrent cancer

1. Introduction

Treatment of head and neck cancer may cause permanent distortion of normal neck anatomy in an unpredictable way. High-dose irradiation and radical resection of the primary and loco-regional tumour deposits may cause a variety of acute and late changes e.g. inflammation, fibrosis, oedema, or scarring [1]. Under these circumstances it can be difficult to distinguish recurrent disease from post-treatment sequelae. Whilst radiography, computed tomography (CT), and magnetic resonance imaging (MRI) provide important anatomical information, they are rather insensitive in distinguishing benign post-treatment reactive masses from viable tumours [2–5]. Since the capacity for functional recovery is diminished in the irradiated tissue, post-irradiation biopsies should be performed with caution. Therefore, non-invasive methods for accurate

* Corresponding author. Tel.: +358-40-741-2038; fax +358-2- 333-2450.
E-mail address: maria.lapela@leiras.fi (M. Lapela).

0959-8049/00/$ - see front matter
PII: S0959-8049(00)00037-X

assessment of post-treatment masses harbouring risk for cancer regrowth in the head and neck region would be of great value.

Positron emission tomography (PET) with fluorine-18-labelled fluorodeoxyglucose (FDG) has the ability of imaging metabolic changes that appear to be linked with malignancy [6]. Detection of primary head and neck cancer and metastatic cervical lymph nodes with FDG PET has been successful [7–9], as well as monitoring of the response to chemotherapy [10] and radiotherapy [11–14]. However, in head and neck imaging the greatest clinical expectations have been raised in the use of FDG PET in the post-therapy setting, as shown during the last few years [15,16]. These studies have indicated that FDG PET provides additional information to add to that obtained from conventional morphological imaging methods and may have an impact on the management of patients presenting with a suspected recurrence of head and neck cancer. Our previous study has suggested that quantitation of tracer uptake may enable the differentiation between malignant and benign lesions in patients with suspected recurrence of head and neck cancer [17]. However, controversy exists, whether qualitative analysis of images is relevant for diagnostic purposes or if quantification of tracer uptake is necessary in clinical practice [18].

The purpose of the study was to confirm the efficacy of FDG PET in differential diagnosis between malignant and benign lesions in a relatively large collaborative study consisting of patients previously treated for head and neck cancer. Specifically, the aim was to evaluate whether quantitation of FDG uptake is better than visual interpretation of the images to differentiate viable tumour from non-malignant tissue.

2. Patients and methods

2.1. Patients

Patients were enrolled from a population seen at the Department of Otorhinolaryngology in Turku University Hospital, Turku, Finland or at the Department of Oncology at Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. The criterion for eligibility was clinically suspected (due to a new mass, intense oedema or fibrosis, increasing pain, delayed mucosal healing, persistent wound, fistula or chronic inflammation or neurological signs) but not yet proven recurrence of head and neck carcinoma during follow-up after definitive surgical and/or radiation treatment. In addition to the physical examination and clinical history, otorhinolaryngological status including indirect pharyngo- and laryngoscopy, bimanual palpation and endoscopic examination were performed using either fibre- and/or rigid scopes when required. Patients were recruited between November 1994 and August 1997. In addition, 14 eligible patients reported in a previous study [17] were included. The degree of suspicion of recurrence for each of these lesions was graded as either strong (clinical grade 2) or weak (clinical grade 1) based on clinical status and patient history by the responsible clinician. After the PET study, a histological or cytological specimen confirming the presence of carcinoma, a negative biopsy or more than 5 months with no evidence of emerging disease was required for the definition of disease status. Each patient gave a written informed consent, and the study protocol was approved by the local ethical committees in Turku and Copenhagen.

The patient and tumour characteristics are presented in Table 1. 56 patients, 16 females and 40 males (mean age: 61 years; range: 34–79 years), underwent PET imaging with FDG, 4 of them twice and 1 patient three times in consecutive recurrence states. Hence, a total of 62 PET studies were performed. All patients were previously treated with a curative intent for head and neck carcinoma. 53 patients had received megavoltage radiation therapy to a total dose of over 60 Gy (range: 62–74 Gy; median: 66 Gy) to the primary tumour area and 1 patient was treated with a total dose of 54 Gy, respectively. In the remaining 2 patients, the exact dose of external radiation therapy to the primary tumour was unknown. 50 patients were also irradiated for bilateral or unilateral cervical nodes (median dose, 50 Gy; range, 32–68 Gy). Before PET, 35 patients had undergone radical surgery, 27 patients for the primary disease, 6 patients for the primary and a recurrent tumour and 2 patients for a previous recurrence. Cervical lymphadenectomy was done in 25 patients. PET imaging of the currently suspected recurrence was performed more than 8 weeks (median: 7 months; range: 2–253 months) after completion of previous therapy. The body mass index (BMI) calculated as weight in kilograms divided by the square of height in metres varied from 15.2 to 40.2 kg/m^2 (median, 23 kg/m^2). None of the patients was known to have diabetes, and all of them fasted for a minimum of 6 h prior to PET imaging. Serum glucose concentration was monitored prior to PET study in 50 patients, and ranged from 3.5 to 7.3 mmol/l (median: 5.5 mmol/l).

2.2. PET study

At Turku PET Center, PET studies were performed with an eight-ring ECAT 931/08 (Siemens/CTI Corp., Knoxville, TN, USA) or General Electric (GE) Advance scanner (General Electric Medical Systems, Milwaukee, WI, USA). In Rigshospitalet, GE Advance and GE 4096 scanners were used. Using the previously reported methods for image reconstruction and scatter correction the devices have a measured axial and spatial resolution of 6.7 mm and 6.5 mm (ECAT), 5 mm and 5 mm (GE Advance) and 6 mm and 6 mm (GE 4096), respectively.

Table 1
Patient and tumour characteristics

	No. of patients (%)
Gender	
Male	40 (71)
Female	16 (29)
Nationality	
Finnish	46 (82)
Danish	10 (18)
Site of primary tumour	
Oral cavity	22 (39)
Pharynx	9 (16)
Larynx	16 (29)
Salivary glands	6 (11)
Unknown primary	3 (5)
Primary tumour size	
T1	6 (11)
T2	22 (39)
T3	12 (21)
T4	12 (21)
TX	4 (7)
Primary nodal status	
N0	33 (59)
N1	9 (16)
N2	11 (20)
N3	2 (4)
NX	1 (2)
Histology	
Squamous cell carcinoma	48 (86)
Adenocarcinoma	2 (4)
Adenoid cystic carcinoma	2 (4)
Lymphoepithelial carcinoma	1 (2)
Transitional cell carcinoma	1 (2)
Acinar cell carcinoma	1 (2)
Mucoepidermoid carcinoma	1 (2)
Grade	
G1	23 (41)
G2	15 (27)
G3	8 (14)
GX	10 (18)
Surgery	
Radical tumour resection	35 (63)
Neck dissection	
Unilateral	22 (39)
Bilateral	3 (5)
No surgery	17 (30)
Radiotherapy	
Conventional	36 (64)
Hyperfractionated	17 (30)
Accelerated	3 (5)

Transmission scanning was performed with a removable ring or pin source containing germanium-68. A mean dose of 340 MBq (range: 228–429 MBq) of FDG was injected intravenously into the cubital vein. Emission data were obtained as 5–10-min frames beginning 35–60 min after injection. PET imaging was performed using a one bed position for each patient. All data were corrected for deadtime, decay and photon attenuation.

In Turku, the FDG synthesis was a modification of the method reported by Hamacher and colleagues [19]. The FDG synthesis in Copenhagen was performed using the FDG Microlab from General Electric Medical Systems. The radiochemical purity of FDG was >92%, the radionuclidic purity >99% and the specific activity at the end of synthesis was 74 GBq/mmol (>2 Ci/mmol) [20].

2.3. Methods for PET analysis

2.3.1. Visual analysis

Acquisition data obtained as a 5-min frame between 55 and 60 min postinjection was used for PET analysis. Hard copies of emission images equipped with colour scale, as well as the respective transmission images, were extracted for visual evaluation. Judgement of the scan readings was made by three investigators with information on initial clinical history and site of clinical suspicion but without knowledge about current CT, histological, surgical or follow-up findings. The visual analysis was graded as clearly positive (visual grade 2) when high FDG uptake as a focal hot spot was seen, or equivocal (visual grade 1) when slightly enhanced accumulation of FDG compared with the expected normal distribution in the region was detected. Otherwise the image was interpreted as negative (visual grade 0) for malignancy.

2.3.2. Standardised uptake values

The same PET frames that were evaluated in visual analysis were used for quantification of FDG uptake by using a semi-automated algorithm for defining a region of interest (ROI) in a tissue with maximum uptake. The algorithm determines maximum average radioactivity in a rectangular ROI of predetermined size (5×5 pixel) within a large investigator-defined ROI comprising the clinically suspected anatomical area. The average counts in each ROI were used for further calculations.

Tracer accumulation in the ROIs was reported as the standardised uptake value (SUV), which is the radioactivity concentration in a ROI at a fixed time point divided by the injected dose and the patient's weight. In addition, SUVs corrected for the lean body mass (unless more than the actual weight) were calculated as SUV_{lean}s [21].

2.4. CT imaging

Neck CT corresponding to PET imaging was achieved in 52 cases (46 patients) on a GE CT Pace system within a median of 12 days (range: 0–75 days) from the PET study. Scans were obtained from the base of the skull to the thoracic inlet at a slice thickness of 5 mm and 5 mm collimation except for laryngeal carcinomas, which were scanned at the level of suspected tumour at a slice

thickness of 3 mm and 3 mm collimation. Intravenous contrast (Iopromid 300 mg I/ml, Schering, Berlin, Germany) was administered in all studies as bolus injections of 100–120 ml. A similar three-point rating scale as for visual PET analysis was used for image interpretation. In CT studies, any abnormal masses with enhancement or interval soft tissue increases compared with previous CT scans were considered malignant (CT grade 2), whereas oedema and soft tissue swelling were interpreted as inconclusive for malignancy (CT grade 1). Lymph nodes were interpreted as malignant (CT grade 2) if CT showed central necrosis or the minimal axial diameter was greater than 12 mm, and as equivocal (CT grade 1) if the minimal axial diameter was 8–12 mm and no necrosis was present. Otherwise the CT scan was interpreted as negative (CT grade 0) for malignancy. The scans were interpreted with the same clinical information that was available for the visual PET analysis.

2.5. Statistical analysis

The quantitative results between different groups were compared with Wilcoxon rank sum test for unpaired data or with Kruskall–Wallis analysis of variance. Standard receiver-operating characteristics (ROC) curves were generated for assessment of diagnostic accuracy. All P values were two-tailed.

3. Results

Sixty-two FDG PET studies in 56 patients were performed because of clinical suspicion of recurrent head and neck cancer. Recurrent disease was suspected if a mass was found ($n=17$), if oedema ($n=9$) or fibrosis ($n=16$) was intense and persistent, if the patient complained about increasing pain ($n=6$), if mucosal healing was delayed ($n=1$), or if persistent wound ($n=5$), fistula ($n=3$) or chronic inflammation ($n=2$) was present. One patient presented with a paresis of hypoglossal nerve, and 2 had trismus. An additional 19 sites in the neck of 15 patients were suspected for disease recurrence, most of which were presented as fibrosis (fibrosis, $n=14$, a mass, $n=4$, or fistula, $n=1$). Hence, the diagnostic efficacy of different modalities was compared in a total of 81 lesions. Fine-needle biopsy was obtained in 37 lesions within a median of 16 days (range: 1–181 days), and reoperation was performed in 29 cases within a median of 22 days (range: 1–296 days) after the PET study. Subsequently, 37 (46%) of the 81 lesions proved to be malignant, based on histopathological or cytological tissue evaluation. 3 patients presented with a histologically confirmed recurrent cancer 6, 7 and 9 months, respectively, after the PET study. However, these patients were also included in the statistical analysis as having recurrent cancer at the time of the study. Forty-four (54%) suspected regions have not presented with malignancy, confirmed either by surgery ($n=4$), by fine-needle biopsy ($n=12$), or by follow-up ($n=28$) over a median of 15.8 months (range: 5.6–58 months). When considering only the major suspected region in each study ($n=62$), half of the lesions ($n=31$) have later been shown to be malignant.

3.1. Clinical evaluation

The clinical suspicion of recurrent disease was defined as described above. Forty-seven out of the total 81 lesions were graded as 1, indicating weak suspicion, for which PET imaging was requested to exclude recurrence. In 11 of these 47 regions (23%), viable cancer was later found. For the remaining 34 regions the referring physician requested a PET study to confirm the strong clinical suspicion (clinical grade 2) of cancer. In 26 of these cases (76%) recurrent disease has been found.

3.2. Visual PET analysis

Three observers interpreted the PET images in consensus visually by scoring each of the 81 suspected regions. Thirty-four areas were judged to be clearly positive (visual grade 2) for malignancy (Fig. 1), three of which proved to be false-positives (radiation injury, $n=2$, Fig. 2, and inflammatory tissue, $n=1$). Time interval between the end of radiation therapy and PET scanning was 5–8 months in these patients. Eight regions were difficult to assess clearly positively or negatively and were considered to be equivocal (visual grade 1). In histological and/or cytological evaluation, four of these had malignant, and three benign tissue (fibrosis, ulceration, and radiation injury, one of each). The remaining equivocal region has not shown any evidence of cancer during a follow-up of 36 months and is thus scored as benign. Thirty-nine lesions were interpreted as negative (visual grade 0) at qualitative PET analysis. Two of these were false negatives at surgery, one was a metastatic lymph node and the other a necrotic ulcerative recurrence in the subglottic larynx.

Sensitivity for the visual interpretation of PET images was 40/42; 95% and specificity 35/42; 84%, when clear and equivocal lesions (visual grade 2 or 1, $n=42$) were considered as positive for malignancy. If only lesions with high uptake (visual grade 2, $n=34$) were considered as positive, sensitivity decreased to 84% but specificity increased to 93%. Positive predictive value was 91% and negative predictive value 87% when grade 2 lesions were considered. Overall accuracy of qualitative FDG PET in differentiation between malignant and benign post-treatment lesions was 89%, independent of whether high (visual grade 2) or high to intermediate (visual grade 1 and 2) uptake was selected in rating the lesions as positive.

Qualitative PET results were evaluated also based on the strength of clinical suspicion for cancer (Fig. 3). Six out of 34 (18%) clinically highly suspicious regions (clinical grade 2) were negative (visual grade 0) at qualitative PET analysis. Four of these were correctly interpreted at PET: three proved to be benign by

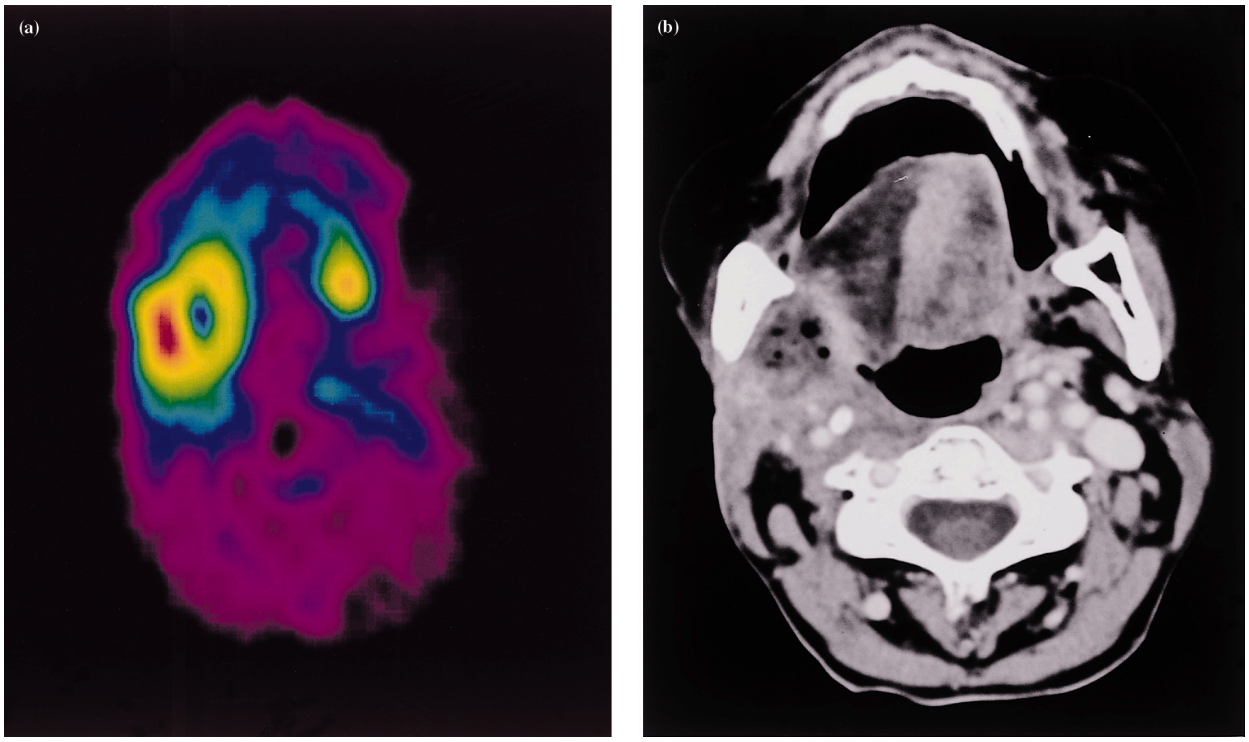

Fig. 1. (a) An FDG PET image of a patient with a histologically verified recurrent adenocystic carcinoma of the oral cavity and in the neck previously treated with radiotherapy (7 years ago) and surgery. High tracer uptake visually scored as grade 2 was found both in the oral cavity (SUV, 3.6) and in the neck (SUV, 5.1). (b) Soft tissue swelling in neck at right (CT grade 1) was detected in a corresponding CT image, but nothing pathological in the oral cavity.

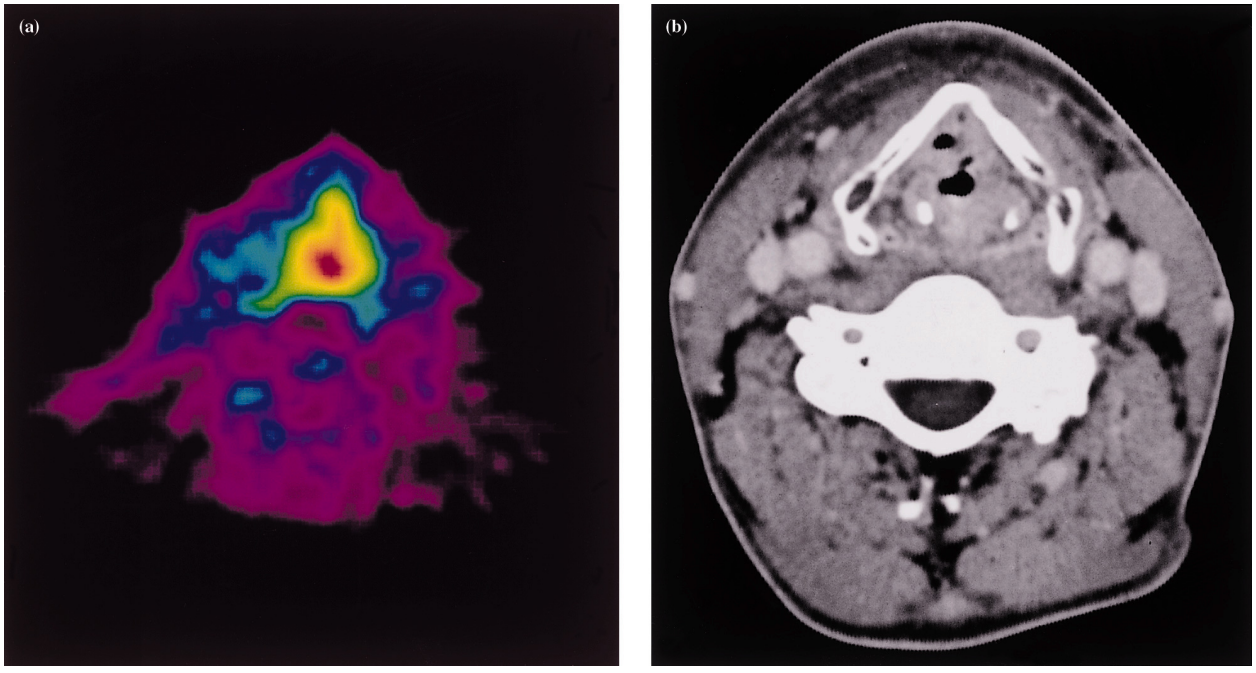

Fig. 2. (a) An FDG PET image of patient with a history of definitively radiated glottic carcinoma (5 months prior to the PET study) presenting with intense laryngeal fibrosis. The PET image was visually graded as strongly positive for malignancy (visual grade 2, SUV, 5.7). Histologically, only fibrosis and radiation injury was found. (b) Diffuse soft tissue enhancement (CT grade 1) was present in a corresponding CT image.

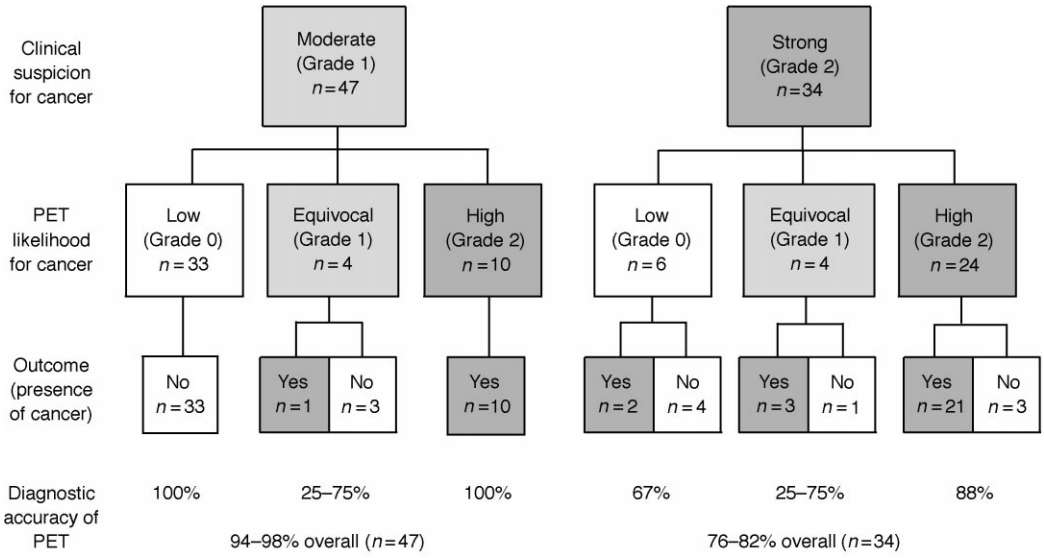

Fig. 3. Outcome of lesions clinically suspected for recurrence of head and neck cancer.

histology, and 1 patient has had no evidence of disease 6 months since the PET study. Two lesions with a strong clinical suspicion but negative PET images were histologically malignant and thus, false-negatives at PET (the two false-negative lesions mentioned above). Forty-seven regions were analysed visually with PET in a situation where the clinician ordered a PET study with a mild suspicion of recurrence (clinical grade 1). PET was correctly graded as normal (visual grade 0) in 33 and equivocal (visual grade 1) in three of the 36 lesions which were later found to be free of disease. Visual PET was correctly positive in 10 regions that were classified high (grade 2). One equivocal classification (grade 1) with only a weak clinical suspicion was positive on histological or cytological verification of viable cancer in subsequent surgery or biopsy.

3.3. Quantitative PET analysis

FDG uptake was quantitated in all of the 81 clinically suspicious regions. The median SUV of all lesions was 4.2 (range: 1.5–36.9), and the median SUV_{lean} was 3.9 (range: 1.1–36.9). The SUVs of the malignant lesions ($n=37$) ranged from 2.1 to 36.9 (median: 6.8), and that of the benign lesions ($n=44$) from 1.5 to 9.3 (median: 3.3; Fig. 4). The difference between these two groups was statistically significant ($P<0.001$). The median SUV_{lean} of malignant and benign lesions were 6.3 (range: 2.1–36.9) and 3.1 (range: 1.1–9.3), respectively ($P<0.001$). If only the single most suspected location in each case was considered, the median SUV of the malignant lesions was 6.8 ($n=31$; range: 2.1–36.9), and that of the benign lesions 3.5 ($n=31$; range: 2.1–9.3), and this difference was also statistically significant ($P<0.0003$). No correlation between serum glucose concentration and the false-negative or positive cases was seen in SUV or visual analyses (data not shown).

SUVs for the visually false-positive lesions were 9.3, 5.7 and 4.9, respectively. All of these values were within the two middle quartiles of the SUVs of the malignant lesions. SUVs for the two false-negative lesions were 2.1 and 5.3, both less than the median SUV of the malignant lesions. Even for experienced investigators, 8 out of 81 images were visually difficult to interpret as either malignant or benign. Unfortunately, we did not find quantitative assessment of FDG uptake very helpful for the diagnosis of these visually 'borderline' cases. In fact, SUVs of these equivocal lesions were comparable for both cases which proved to be malignant (range: 2.4 to 4.0, $n=4$) and benign (range: 2.5–8.8, $n=4$).

The calculated tracer uptake values were analysed also with (ROC) curve to obtain the optimal cut-off

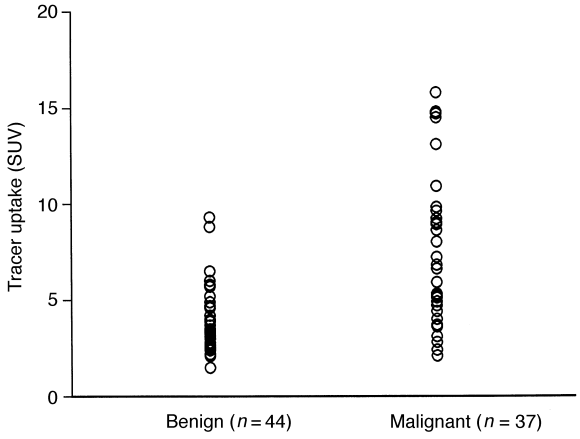

Fig. 4. Tracer uptake in tissues presented as SUVs in malignant ($n=37$) and benign ($n=44$) lesions. One patient with an exceptionally high SUV 36.9 has been excluded from the graph to enhance comparison of other lesions

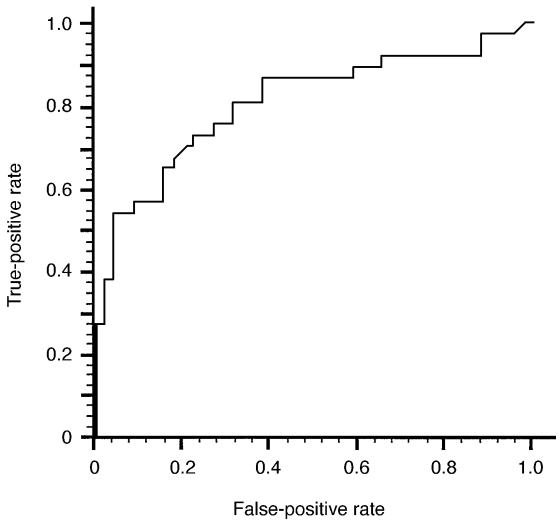

Fig. 5. Receiver-operating characteristic curve (ROC) of quantitative FDG PET readings (SUV). The respective curve for SUV_{lean} was almost identical (data not shown). Note the lack of a steep edge in the upper left corner, indicating the absence of a useful cut-off value for differentiation between malignant and benign lesions.

value of SUV (or SUV_{lean}) to differentiate between malignant and benign post-treatment abnormalities (Fig. 5). In our series, however, we could not demonstrate a better accuracy with SUV analysis than that obtained by visual evaluation only. The highest proportion of correct interpretation was 75% at a probability level where sensitivity was 68% and specificity was 82%. Thus, the accuracy of quantitative analysis was inferior to visual analysis, independent of the level of confidence chosen by the observer.

3.4. Computed tomography

CT was available in 52 cases (46 patients) including 69 of the 81 analysed lesions within the field of view. Nineteen regions were interpreted as clearly pathological (CT grade 2) at CT, and all of them proved to be malignant at surgery or fine-needle biopsy. The findings of 32 locations at CT were considered to be benign post-treatment changes (CT grade 0). In three of them malignant tissue was later confirmed. Eighteen lesions could not be definitely judged as either malignant or benign based on the criterion used for this analysis and, hence, were graded as equivocal (CT grade 1). Ten of these lesions turned out to be malignant and 8 benign post-treatment abnormalities. Sensitivity of CT in detecting malignancy was 29/32; 91% and specificity 29/37; 78%, if clear and equivocal findings (CT grade 2 and 1; $n=37$) were considered as malignant (Table 2). Correspondingly, if equivocal lesions were treated as normal findings and only clearly pathological were considered as positive (CT grade 0 and 1 versus CT grade 2; $n=19$), sensitivity decreased to 59%, but specificity increased to 100%. For comparison, if diagnostic accuracy of qualitative PET analysis was calculated only for cases where both CT and PET were performed ($n=69$), the sensitivities and specificities would be 97% and 83% (visual grade 0 versus visual grade 1 and 2), and 84% and 92% (visual grade 0 and 1 versus visual grade 2), respectively (Table 2).

4. Discussion

The feasibility of FDG PET in the identification of viable tumour tissue in the head and neck area after treatment was first demonstrated in a patient with recurrence in irradiated mandible [22]. Soon FDG PET was shown to have comparable or better sensitivity and specificity than MRI/CT for detection of loco-regional recurrences of head and neck cancer. Rege and colleagues demonstrated a specificity of 100% and sensitivity of 90% for PET compared with 59% and 70% for MRI, respectively, in 18 patients imaged 2.5–192 weeks after radiation therapy [15]. McGuirt and associates reported an accuracy of 85% for PET and 42% for CT/MRI in 13 patients with a history of laryngeal carcinoma [23]. Wong and coworkers found sensitivities and specificities of 75% and 80% for CT and/or MRI, respectively, and those of 100% and 100% for FDG PET, respectively, in 12 patients at the primary site and in 13 patients at nodal sites [24]. Anzai and colleagues evaluated 12 patients with FDG PET and detected 7 of 8 recurrent tumours, and PET was truly negative in 4 of 4 patients, whereas MRI/CT detected only 2 of 8 and was negative in 3 of 4 patients [25]. In the present study

Table 2
Comparison of FDG PET and CT for accuracy of detection of head and neck cancer with two schemes for grading of malignancy

Grade	0 and 1 versus 2[a]			0 versus 1 and 2[b]		
	Sensitivity %	Specificity %	Accuracy %	Sensitivity %	Specificity %	Accuracy %
FDG PET	84	93	88	95	84	90
CT	59	100	81	91	78	84

The number of lesions (and studies) where both FDG PET and CT were available were 69 (52, respectively). The slight difference in performance figures between this table and Results is due to omitting here cases where only PET (but not CT) was available.

[a] Lesions with strong suspicion (grade 2) for recurrence are rated as positive for malignancy.

[b] Lesions with equivocal and strong suspicion (grade 1 and 2) for recurrence are rated as positive for malignancy.

with a larger number of patients, similar results were demonstrated, thus supporting the clinical utility of FDG PET in the detection of recurrent head and neck cancer. This is in accordance with Keyes and associates who consider FDG PET to be of high accuracy, but limited usefulness, in primary head and neck cancer, the post-therapy setting being the single application where FDG PET appears to be advantageous [26].

We have previously demonstrated that quantification of tracer accumulation by a simple single scan approach (SUV) is comparable with a more elaborate kinetic modelling requiring knowledge of tracer input function [27]. Controversy exists, however, whether quantification of tracer uptake is necessary at all in clinical practice [18]. We have assumed that quantification of uptake might assist in correct image interpretation by overcoming some practical problems such as subjectivity of visual analysis and interference of physiologically enhanced tracer uptake sometimes seen in muscle, mucosa and inflammatory tissue. In the study by McGuirt and colleagues, tracer uptake by tumour and non-malignant tissues showed no overlap; the SUVs of the recurrent tumours ranged from 4.9 to 10.7, and those of non-recurrent regions from 2.4 to 4.7 [23]. Anzai and colleagues found that the mean SUV ($\pm$S.D.) of recurrent tumours was higher (5.5 ± 1.1) than that of postsurgery or postradiation changes (2.9 ± 1.1), but there was one case of overlap [25]. Our preliminary experience of 17 FDG PET studies in 15 patients indicated that malignant lesions had significantly higher FDG accumulation than benign ones, but some overlap between these two groups did exist [17]. In the current study, we wanted to investigate further whether an optimal cut-off level for FDG uptake that would differentiate benign and malignant lesions could be found by studying a sufficiently large patient population. However, the ROC analysis demonstrated the limited accuracy of a quantitative approach and indicated a poor likelihood for finding a clinically feasible cut-off value for FDG uptake index if SUV or SUV_{lean} is adapted. These results are seemingly contrasted by findings in our previous study [17] but simply reflect the fact that small sample sizes run a risk of making premature conclusions. Although tracer uptake between recurrent cancer and benign post-treatment tissue was significantly different, it is evident from the current series that statistical significance does not always translate to clinical usefulness. The highest uptake in a benign lesion (SUV 9.3) was much higher than 6.8, the median SUV of the malignant lesions. Likewise, uptake in the three malignant lesions was within the limits of the lowest quartile of the benign lesions. Indeed, the majority of SUVs as well as SUV_{lean}s of both groups was overlapped (Fig. 3). In spite of this, the visual interpretation of PET images correctly assessed the diagnosis in 68 out of 81 cases. Only five incorrect judgements were made of which three were false positives and two false negatives and 8/81 were visually difficult to interpret. From the SUV values for these regions quantitative assessment of FDG uptake was of little diagnostic value. Whilst diagnostic accuracy in the visual analysis ranged from 88% to 90% (Table 2) and that of quantitative analysis with SUV was found to be at its highest at 75% in ROC analysis we can see little justification for the calculation of SUVs in clinical practice (data not shown). Hence, the present study emphasises the fact that visual analysis is not inferior to quantitative analysis in the correct identification of post-treatment masses in the head and neck area.

It is acknowledged that quantitative analysis has to be performed with caution in lesions that are small in comparison with the spatial resolution of the PET device. Considering post-treatment malignant and benign tumours in the head and neck, definition of the exact size of the lesion can be a complicated task. Squamous cell carcinoma often has a highly infiltrative nature and demarcation of these neoplasms as well as that of benign lesions such as fibrosis and oedema may be poor in irradiated and operated tissue. Moreover, the measured uptake of small lesions would be artificially reduced in both malignant and benign tumours which appear to manifest themselves at comparable sizes Although correction for recovery is essential in the measurement of true radioactivity concentrations in small objects, it is unlikely that the current observations of the equal usefulness of qualitative and quantitative analyses of PET studies would be different even if correction factors had been used.

In our preliminary study, we suggested that FDG PET would be most useful in cases where CT is suspicious but PET is negative, or where CT is negative but a clearly increased focal FDG uptake is seen at PET. In the current series, CT was rated equivocal or weekly suspicious in 18 lesions. Out of these cases, visual PET was correctly positive in 9 out of the 10 malignant lesions, and correctly negative in 6 out of the 8 benign lesions. Three malignant lesions were not found in the CT scan; two of them were clearly visible in FDG PET and one was equivocal at visual PET analysis. The only malignant lesion negative at PET analysis where a CT scan was available was judged as clearly pathological by CT. Three lesions were false-positives in the PET analysis; CT scan was correctly negative in one and equivocal in two of these. Thus comparison of the CT and FDG PET techniques in this series supports their use as complementary, rather than alternative, imaging methods.

Knowledge of the relative metabolic variations in the normal structures of the head and neck first reported by Jabour and colleagues [28] and demonstrated in whole-body scans by Engel and associates [29] is essential when abnormal foci of tracer uptake have to be identified in PET images. The well known drawback of increased FDG uptake in inflammatory tissue was also

seen in this series, where inflammation and radiation injury without neoplastic cells was found in all three false-positive PET lesions. If a patient presents with findings suggestive of a chronic inflammation and on top of that has a recurrence, a series of follow-up PET scans combined with close clinical observation might be the approach of choice. It is also strongly recommended to wait until 2–3 months after radiotherapy before doing FDG PET for a patient with head and neck cancer [30]. This will allow acute tissue reactions within the tumour area to subside and facilitates the interpretation of the PET study.

5. Conclusion

In the current study of 56 patients with suspected recurrence of head and neck cancer, FDG PET was found to be feasible for the detection of malignancy. Visual analysis of PET images indicated that FDG PET is better than CT in correct interpretation of findings. We did not find quantitation of FDG uptake to add to the visual analysis in the differential diagnosis between post-treatment malignant and benign lesions. However, SUVs may still be used in selected cases where detailed characteristics on patient or tumour metabolism are needed, e.g. if comparison of sequential scans is desired. FDG PET can be recommended in a clinical setting when physical examination does not make clear the nature of patient's symptom(s) or a suspicious lesion is found but biopsy is negative. PET may also be valuable if MRI or CT is inconclusive, and positive PET can be used to guide diagnostic biopsies and/or surgery to confirm the recurrence. In contrast, a close clinical follow-up of patients with negative or indeterminate PET but persistent progressive symptoms is mandatory.

Acknowledgements

This study was financially supported by Nordisk Cancer Union.

References

1. Million RR, Cassisi NJ, eds. *Management of Head and Neck Cancer: A Multidisciplinary Approach*, 2nd edn. Philadelphia, J.B. Lippincott Company, 1994.
2. Bronstein AD, Nyberg DA, Schwartz AN, Schuman WP, Griffin BR. Soft-tissue changes after head and neck radiation: CT findings. *AJNR* 1989, **10**, 171–175.
3. Glazer HS, Niemeyer JH, Balfe NM, *et al*. Neck neoplasms: MR imaging. Part II. Posttreatment evaluation. *Radiology* 1986, **160**, 349–354.
4. Som PM, Biller HF. Computed tomography of the neck in the postoperative patient: radical neck dissection and the myocutaneous flap. *Radiology* 1983, **148**, 157–160.
5. DiSantis DJ, Balfe DM, Hayden RE, Sagel SS, Sessions D, Lee JKT. The neck after total laryngectomy: CT study. *Radiology* 1984, **153**, 713–717.
6. Wahl RL, Hutchins GD, Buchsbaum DJ, Liebert M, Grossman HB, Fisher S. 18F-2-deoxy-D-glucose uptake into human tumour xenografts: feasibility studies for cancer imaging with positron emission tomography. *Cancer* 1991, **67**, 1544–1550.
7. Haberkorn U, Strauss LG, Reisser Ch, *et al*. Glucose uptake, perfusion, and cell proliferation in head and neck tumours: relation of positron emission tomography to flow cytometry. *J Nucl Med* 1991, **32**, 1548–1555.
8. Lindholm P, Leskinen-Kallio S, Minn H, *et al*. Comparison of fluorine-18-fluorodeoxyglucose and carbon-11-methionine in head and neck cancer. *J Nucl Med* 1993, **34**, 1711–1716.
9. Meyers LL, Wax MK, Nabi H, Simpson GT, Lamonica D. Positron emission tomography in the evaluation of the N0 neck. *Laryngoscope* 1998, **108**, 232–236.
10. Lowe VJ, Dunphy FR, Varvares M, *et al*. Evaluation of chemotherapy response in patients with advanced head and neck cancer using [F-18]fluorodeoxyglucose positron emission tomography. *Head Neck* 1997, **19**, 666–674.
11. Minn H, Paul R, Ahonen A. Evaluation of treatment response to radiotherapy in head and neck cancer with fluorine-18 fluorodeoxyglucose. *J Nucl Med* 1988, **29**, 1521–1525.
12. Chaiken L, Rege S, Hoh C, *et al*. Positron emission tomography with fluorodeoxyglucose to evaluate tumour response and control after radiotherapy. *Int J Radiation Oncology Biol Phys* 1993, **27**, 455–464.
13. Greven KM, Williams III DW, Keyes Jr JW, *et al*. Positron emission tomography of patients with head and neck carcinoma before and after high dose irradiation. *Cancer* 1994, **74**, 1355–1359.
14. Brun S, Ohlsson T, Erlandsson K, *et al*. Early prediction of treatment outcome in head and neck cancer with 2-18FDG PET. *Acta Oncol* 1997, **37**, 741–747.
15. Rege S, Maas A, Chaiken L, *et al*. Use of positron emission tomography with fluorodeoxyglucose in patients with extracranial head and neck cancer. *Cancer* 1994, **73**, 3047–3058.
16. Greven KM, Williams 3rd DW, Keyes Jr JW, McGuirt WF, Watson Jr NE, Case LD. Can positron emission tomography distinguish tumour recurrence from irradiation sequelae in patients treated for larynx cancer. *Cancer J Sci Am* 1997, **3**, 353–357.
17. Lapela M, Grénman R, Kurki T, *et al*. Head and neck cancer: detection of recurrence with PET and 2-[F-18]fluoro-deoxy-D-glucose. *Radiology* 1995, **197**, 205–211.
18. Keyes Jr JW. SUV: standard uptake or silly useless value? *J Nucl Med* 1995, **36**, 1836–1839.
19. Hamacher K, Coenen HH, Stöcklin G. Efficient stereo specific synthesis of no-carrier-added 2-[18F]fluoro-2-deoxy-D-glucose using aminopolyether supported nucleophilic substitution. *J Nucl Med* 1986, **27**, 235–238.
20. Solin O, Bergman J, Haaparanta M, Reissell A. Production of 18F from water targets, specific radioactivity and anionic contaminants. *Int Radiat Isotopes* 1988, **39**, 1065–1071.
21. Zasadny K, Wahl RL. Standardised uptake values of normal tissues at PET with 2-[fluorine-18]-fluoro-2-deoxy-D-glucose: variations with body weight and a method for correction. *Radiology* 1993, **189**, 847–850.
22. Minn H, Aitasalo K, Happonen RP. Detection of cancer recurrence in irradiated mandible using positron emission tomography. *Eur Arch Otorhinolaryngol* 1993, **250**, 312–315.
23. McGuirt WF, Williams III DW, Keyes Jr JW, *et al*. A comparative diagnosis study of head and neck nodal metastasis. *Laryngoscope* 1995, **105**, 373–375.
24. Wong WL, Chevretton EB, McGurk M, *et al*. A prospective study of PET-FDG imaging for the assessment of head and neck squamous cell carcinoma. *Clin Otolaryngol* 1997, **22**, 209–214.
25. Anzai Y, Carroll WR, Quint DJ, *et al*. Recurrence of head and neck cancer after surgery or irradiation: prospective comparison

of 2-deoxy-2-[F-18]fluoro-D-glucose PET and MR imaging diagnoses. *Radiology* 1996, **200**, 135–141.
26. Keyes Jr JW, Watson Jr NE, Williams 3rd DW, Greven KM, McGuirt WF. FDG PET in head and neck cancer. *Am J Roentgenol* 1997, **169**, 1663–1669.
27. Minn H, Leskinen-Kallio S, Lindholm P, *et al*. [^{18}F]Fluorodeoxyglucose uptake in tumours: kinetic vs steady-state methods with reference to plasma insulin. *J Comput Assist Tomogr* 1993, **17**, 115–123.
28. Jabour BA, Choi Y, Hoh CK, *et al*. Extracranial head and neck: PET imaging with 2-[F-18]fluoro-deoxy-D-glucose and MR imaging correlation. *Radiology* 1993, **186**, 27–35.
29. Engel H, Steinert H, Buck A, Wholebody PET. Physiological and artifactual fluorodeoxyglucose accumulations. *J Nucl Med* 1996, **37**, 441–446.
30. Guirt WF, Greven K, Williams D, *et al*. PET scanning in head and neck oncology: a review. *Head Neck* 1998, **20**, 208–215.