

Qualitative interpretation of PET scans using a Likert scale to assess neck node response to radiotherapy in head and neck cancer

Johanna Sjövall¹  · Ulrika Bitzén² · Elisabeth Kjellén³ · Per Nilsson³ · Peter Wahlberg¹ · Eva Brun³

Received: 15 June 2015 / Accepted: 4 September 2015 / Published online: 2 October 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract

Purpose The aim of this study was to determine whether PET scans after radiotherapy (RT), visually interpreted as equivocal regarding metabolic neck node response can be used to accurately categorize patients as responders or nonresponders using a Likert scale and/or maximum standardized uptake value (SUVmax). Other aims were to determine the performance of different methods for assessing post-RT PET scans (visual inspection, a Likert scale and SUVmax) and to establish whether any method is superior in predicting regional control (RC) and overall survival (OS).

Methods In 105 patients with neck node-positive head and neck cancer, the neck node response was evaluated by FDG PET/CT 6 weeks after RT. The scans were clinically assessed by visual inspection and, for the purposes of this analysis, re-evaluated using the Deauville criteria, a five-point Likert scale previously used in lymphoma studies. In addition, SUVmax was determined.

Results All assessment methods were able to significantly predict RC but not OS. The methods were also able to significantly predict remission of tumour after completion of RT. Of the 105 PET scans, 19 were judged as equivocal on visual inspection. The Likert scale was preferable to SUVmax for grouping patients as responders or nonresponders.

Conclusion All methods (visual inspection, SUVmax and the Likert scale) identified responders and nonresponders and predicted RC. A Likert scale is a promising tool to reduce to a minimum the problem of PET scans judged as equivocal. Consensus regarding qualitative assessment would facilitate PET reporting in clinical practice.

Keywords Positron emission tomography · Head and neck cancer · Qualitative evaluation · Therapy assessment · Patient outcome assessment · Radiotherapy

Introduction

Curatively intended treatment in locally advanced head and neck squamous cell carcinoma often involves a combination of treatment modalities such as surgery and/or radiotherapy (RT) with or without chemotherapy. With an organ preservation approach, accurate evaluation of the treatment response is crucial, as salvage surgery can be offered when a resectable, residual tumour is detected. Different imaging modalities can be used for response assessment following RT [1, 2]. During recent years, a number of studies have focused on the evaluation of neck node response. FDG PET/CT has been shown to provide more accurate assessment of treatment response than contrast-enhanced CT alone [3, 4]. Studies using PET/CT have demonstrated high negative and positive predictive values (NPV, PPV) with respect to persistent tumour after RT [5, 6].

FDG uptake over time corresponds to glucose uptake in the tissues and FDG PET scans can be evaluated either quantitatively or qualitatively. The most accurate way of analysing the metabolism of a tumour is to calculate the metabolic rate of glucose (MRglu) using either non-linear regression [7] or

✉ Johanna Sjövall
johanna.sjovall@med.lu.se

¹ Department of Otorhinolaryngology (ORL)–Head and Neck Surgery, Skane University Hospital, Lund University, 22185 Lund, Sweden

² Department of Clinical Physiology and Nuclear Medicine, Skane University Hospital, Lund University, Lund, Sweden

³ Department of Oncology and Radiation Physics, Skane University Hospital, Lund University, Lund, Sweden

Patlak analysis [8]. Both methods are able to quantify the metabolic activity of the tumour by measuring the FDG activity in tissue over time, normalized to the integrated activity of the substance in plasma. However, multiple blood samples are required and the methods are laborious and not clinically feasible. Semiquantitative ways of assessing tumour metabolism, though less accurate, are in clinical use. The standardized uptake value (SUV) is a semiquantitative approach to assess metabolism in which FDG activity in tissue is normalized for the injected dose of FDG and the patient's body mass. Several ways of determining SUV are in use; maximum SUV (SUV_{max}) is the most commonly used parameter [9]. SUV shows limited value for assessing treatment response and predicting outcome in the individual patient [10]. The more recently introduced parameter metabolic tumour volume or total lesion glycolysis may be more promising [11].

Visual inspection of the relative difference between the tumour metabolism and the metabolism of the surrounding, normal tissue is usually adequate to identify tumours and determine treatment response [12]. Although this is clinically valid there is a lack of standardized criteria for visual inspection to enable comparisons between patients and between studies. In order to overcome this problem, different Likert scales are used in evaluation of PET scans. The Deauville criteria Likert scale is used for qualitative evaluations of FDG metabolism in Hodgkin lymphoma, and has a reproducible and excellent predictive ability in patients undergoing chemotherapy treatment [13]. In other tumours, such as head and neck cancer, only a few studies have produced data on clinically convenient and reproducible qualitative interpretation criteria for FDG metabolism [14]. The Deauville criteria are based on the relative metabolic activity in the tumour compared with the mediastinal blood pool and liver. The prognostic value of a PET scan is challenged when FDG uptake is reported as equivocal, i.e. when the metabolism of a lesion is close to that of the surrounding normal tissue. Whether this represents responding tumour tissue and a good prognosis or a patient at risk for residual disease is of clinical importance and also a relevant clinical PET problem. In longitudinal PET studies of Hodgkin lymphoma, equivocal scans have proved to represent a good, rather than bad, prognosis [15].

The objective of the present study was to determine whether early post-RT PET scans, visually assessed as equivocal regarding neck node metabolism, can be used to accurately categorize patients as responders or nonresponders using a Likert scale and/or SUV_{max}. Other aims were to determine the performance of different methods for assessing post-RT PET scans (visual inspection, a Likert scale and SUV_{max}) and to establish whether any method is superior in predicting regional control (RC) and overall survival (OS).

Materials and methods

The present study was a retrospective re-evaluation of PET scans from patients who were included in a previously reported prospective study performed at a tertiary referral hospital [16]. The report of that study includes a thorough description of the materials and methods. In the previous study a PET-directed neck policy was evaluated. Patients were enrolled from 2009 to 2012 and provided informed consent, and the study was approved by the regional ethics committee. Patients with node-positive squamous cell carcinoma of the larynx, hypopharynx and oropharynx, scheduled for organ preservation therapy with curative intent, and with no initial neck surgery, were included in the study. Patients were only eligible if the neck nodes demonstrated hypermetabolism on a PET scan performed before treatment. All 105 eligible patients underwent a baseline PET/CT scan (PET1) during the 2 weeks before the start of RT. Another PET/CT scan (PET2) was performed 6 weeks after completion of RT. The median time from completion of RT to PET2 was 43 days (range 34–87 days); 90 % of the patients had their PET2 scan between 37 and 52 days after therapy.

The PET result determined the management of the neck, i.e. the decision regarding neck dissection versus observation. If PET2 was negative, another scan was scheduled 12 weeks later (PET3). The multidisciplinary team ultimately decided whether patients with an equivocal PET2 scan would undergo neck dissection or be scheduled for PET3. The decision was based on a combination of prognostic factors including initial N classification, human papillomavirus (HPV)/p16-status (p16 was sometimes used as a surrogate marker for HPV infection), performance status and also the clinical response to RT. If PET3 was considered positive or equivocal the patient underwent neck dissection and if negative, the patient was scheduled for follow-up according to routine clinical practice. Clinical follow-up examinations were scheduled every 3 months during the first 2 years after either neck dissection or negative PET3, every 4 months in year 3 and every 6 months during the last 2 years of follow-up.

Patients

A total of 105 patients underwent PET1 and PET2, and 86 patients were also scheduled for PET3. For the Likert scale assessment, 104 patients were eligible for analysis of PET2. Adequate scanning including determination of the mediastinal blood pool was missing in one patient. Patient demographics and tumour characteristics are presented in Table 1. All included patients had complete clinical data and follow-up until

Table 1 Characteristics of the 105 patients and 107 tumours. The majority of patients had p16-positive, oropharyngeal cancer, T2N2. Synchronous head and neck primary tumours were found in two patients

Characteristic	Value
Age (years), median (range)	61 (34–89)
Sex	
Female	27 (26 %)
Male	78 (74 %)
Primary site (<i>n</i> =107)	
Tonsil	71 (66 %)
Tongue base	18 (17 %)
Oropharynx (other than above)	6 (6 %)
Supraglottis	7 (7 %)
Hypopharynx	5 (4 %)
HPV or p16 ^{INKa} positivity (<i>n</i> =107)	
HPV-positive	43 (40 %)
p16-positive	36 (34 %)
HPV or p16-negative	25 (23 %)
Unknown	3 (3 %)
T classification (<i>n</i> =105+2)	
T1	19+1 (19 %)
T2	55+1 (52 %)
T3	18 (17 %)
T4	13 (12 %)
N classification	
N1	14 (13 %)
N2a	18 (17 %)
N2b	58 (56 %)
N2c	13 (12 %)
N3	2 (2 %)
Treatment	
Radiotherapy alone	97 (92 %)
Concurrent cisplatin	6 (6 %)
Induction cisplatin and 5-fluorouracil	2 (2 %)

Values are *n* (%) (patients or tumours), except years of age as median (range)

December 2014. The median follow-up time was 42 months (9–67 months).

PET imaging

In accordance with our standard clinical PET protocol, the patients were injected intravenously with FDG 4 MBq/kg body weight to a maximum dose of 400 MBq after a 4-h fasting period. All patients were imaged with an integrated PET/CT system (Philips Gemini TF; Philips Medical Systems, Cleveland, OH). After a 1-h uptake period at rest, images were acquired for 2 min per bed position. At baseline, the treatment planning CT scan was acquired together with the PET scan. For follow-up studies, low-dose CT scans (50 mAs) were used. The CT scan was used for attenuation correction and anatomical localization. The PET1 scan was acquired from the vertex to the upper thighs. The subsequent PET scans were acquired from the vertex to the upper abdomen.

Image interpretation

As described in the previous prospective study, all PET2 scans were visually evaluated as part of the clinical routine by nuclear medicine physicians regarding metabolic response [16]. For the purpose of the present study, the clinical reports were categorized as “low metabolism”, “remaining hypermetabolism” or “equivocal”. In the present study two experienced readers, without knowledge of the clinical outcome, re-evaluated all patients regarding metabolic neck node response. Images were read and SUVmax levels obtained using an Extended Brilliance Workspace workstation (V4.5.3.40140; Philips Healthcare, The Netherlands). Metabolic responses were scored according to the Deauville criteria [17]. The Deauville criteria are given in Table 2 together with the categories used in the present study. Examples of corresponding PET images are shown in Fig. 1. If FDG uptake was found to have remained in the neck nodes, the highest uptake was scored.

Definitions and statistical considerations

A complete response to RT in the neck and RC was defined as no residual or recurrent tumour involving the neck after completion of RT until the last date of follow-up. A residual tumour

Table 2 The Deauville criteria and interpretation into categories regarding the neck node response to treatment

Deauville score	FDG uptake	Category
1	No uptake	Complete metabolic response
2	Uptake ≤ mediastinum	Probably complete metabolic response
3	Uptake > mediastinum but ≤ liver	Probably postradiation inflammation
4	Uptake moderately increased compared with the liver at any site	Probably persistent tumour
5	Uptake markedly increased compared with the liver at any site and/or new sites of disease	Persistent tumour

Baseline		6 weeks after radiotherapy		
Fusion PET/CT	¹⁸ F-FDG-PET	¹⁸ F-FDG-PET	Fusion PET/CT	Deauville score
				2
				3
				4
				5

Fig. 1 PET images and corresponding Deauville scores. Score 1 was not used in this study and is not shown

in the neck was characterized as persistent tumour, according to the pathology report, after neck dissection scheduled as a result of the PET2 or PET3 scan. Neck relapses that occurred after PET3 in patients who achieved a complete remission prior to this were categorized as recurrences. The Deauville criteria Likert scale was dichotomized into responders (scores 2 and 3) and nonresponders (scores 4 and 5). The SUVmax was dichotomized correspondingly, with a cut-off value of 2.0.

A PET scan demonstrating “low metabolism”, SUVmax ≤ 2.0 or a Deauville score of 2 or 3 was considered true-negative if the patient had neither any tumour, according to the pathology report after neck dissection, nor any neck relapse during follow-up. In false-negative PET scans, either residual tumour was found or a relapse occurred during the follow-up period. Scans reported as showing “remaining hypermetabolism”, SUVmax > 2.0 or a Deauville score of 4 or 5 were considered true-positive if neck dissection revealed residual tumour, according to the pathology report, or if progressive neck disease was seen in patients with non-operable disease during the clinical follow-up. In false-positive PET scans, no residual tumour, according to the pathology report, was found in the neck dissection specimen.

Differences between groups were evaluated by Fisher’s exact test or the chi-squared test for categorical variables and the Mann-Whitney *U* test for continuous variables. RC and OS were estimated using the Kaplan-Meier method and groups were compared using the log-rank test. Time was calculated from the date of diagnosis. Receiver-operating characteristic (ROC) curve analysis was performed for the Likert scale, and SUVmax and the ROC data were used to determine

cut-off values. All tests were two-sided and a *p* value of < 0.05 was considered statistically significant. Where applicable, 95 % confidence intervals (CIs) were used. The software package IBM SPSS statistics 20 (SPSS Inc., Chicago, IL, USA) was used for the statistical analysis.

Results

Outcome

The median follow-up time from the date of diagnosis was 42 months (range 9–67 months). The OS at this time-point was 83.7 %. As expected, there was a significant difference in OS between patients with p16-positive and patients with p16-negative tumours (94.7 % and 52.3 %, respectively; $p < 0.00001$). During follow-up, 28 patients had recurrent tumour at the primary site (T site), in the neck (N), at a distant site (M) or at a combination of sites (Table 3). Residual neck tumour after treatment completion or recurrence in the neck was seen in 23 patients.

Metabolic response and regional tumour control

A significant difference in RC was seen for every single category of visual evaluation, i.e. “low metabolism”, “remaining hypermetabolism” and “equivocal” ($p < 0.0005$). A significant difference was also found for each of the Deauville criteria ($p < 0.00001$), and the Deauville criteria dichotomized into responders (score 2 or 3) and nonresponders (score 4 or 5; $p < 0.0005$). Likewise, a significant difference in RC ($p = 0.0005$) between patients with SUVmax ≤ 2.0 and those with SUVmax > 2.0 (Fig. 2).

The ROC analysis established that the Likert scale and SUVmax were similar in their ability to correctly identify the level of metabolism that correlated with regional tumour control, with areas under the curve (AUC) of 0.81 (CI 0.70–0.91) and 0.78 (CI 0.67–0.88), respectively. None of the evaluation methods was able to significantly predict OS or disease-specific survival. These results seem to be valid independently of p16 status. However, the small number of patients did not allow statistical analysis.

Sensitivity, specificity, PPV, NPV and accuracy for the dichotomized Likert scale were 47.8 % (CI 27.4–68.9 %),

Table 3 Recurrences during the follow-up period, $n = 28$

Site of relapse	No. of patients
T site	9
N site	5
M site	7
T and N site	5
N and M site	2

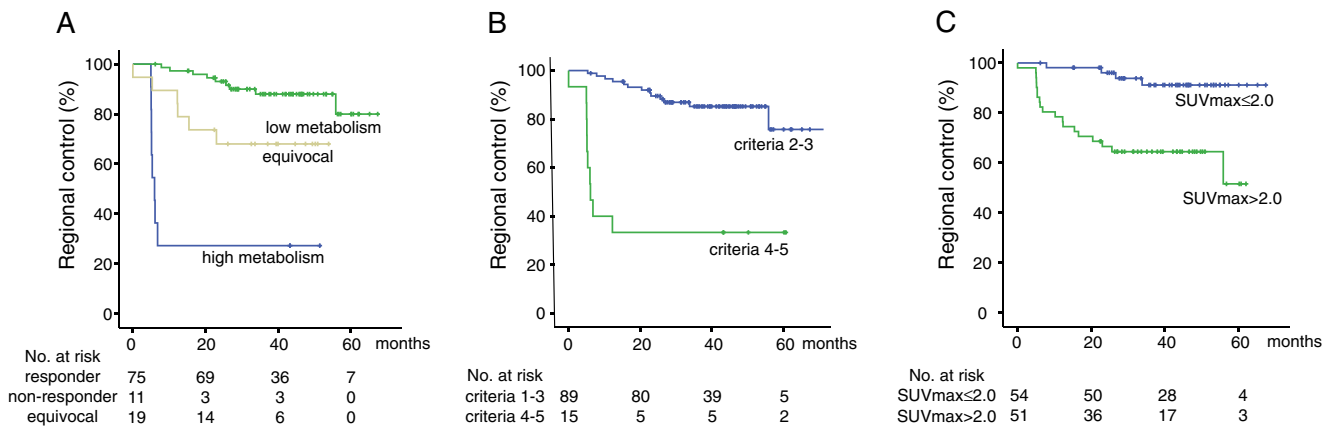


Fig. 2 Regional control in relation to **a** visual inspection, **b** dichotomized Deauville criteria, and **c** maximum standardized uptake value (SUVmax)

93.8 % (CI 85.6–97.7 %), 68.7 % (CI 41.5–87.9 %), 86.4 % (CI 77.0–92.5 %) and 83.6 % in predicting neck node control after RT. Corresponding values for SUVmax were 82.6 % (CI 60.5–94.3 %), 61.0 % (CI 49.5–71.4 %), 37.3 % (CI 24.5–51.9 %), 92.6 % (CI 81.2–97.6 %) and 65.7 %. If the visual inspection was dichotomized by considering the equivocal PET scans together with the visually positive PET scans (sensitive reading), the sensitivity, specificity, PPV, NPV and accuracy for visual inspection were 60.9 % (CI 38.8–79.5 %), 80.5 % (CI 70.0–88.1 %), 46.7 % (CI 28.8–65.4 %), 88.0 % (CI 78.0–94.0 %) and 76.2 %. Corresponding values considering the equivocal PET scans together with the visually negative PET scans (specific reading) were 34.8 % (CI 17.2–57.2 %), 96.4 % (CI 89.0–99.1 %), 72.7 % (CI 39.3–92.7 %), 84.0 (CI 74.7–90.5 %) and 82.9 %.

Metabolic response and remission in the neck after radiotherapy

All methods were able to significantly predict metabolic response and neck node remission: $p=0.00002$ for visual inspection, $p=0.000007$ for the Deauville criteria with dichotomization and $p<0.000002$ without dichotomization, and $p<0.0005$ for SUVmax. The median value of SUVmax was 1.9 (CI 1.7–2.0) in patients with a complete response and 2.6 (CI 2.0–3.9) in patients with remaining or recurrent tumour.

PET scans judged as equivocal

Of the PET2 scans, 19 were reported as equivocal regarding neck nodes and seven patients underwent neck dissection 3–4 weeks after the scan. According to the pathology report only one patient had remaining tumour cells in the neck specimen and was soon diagnosed with distant metastases as well. The remaining 12 patients were scheduled for PET3 18 weeks after RT and were all assessed as complete responders, although in five of these patients a neck node recurrence developed during

the follow-up period. In total, 13 of the patients with equivocal PET scans maintained RC. The tumour was HPV/p16-positive in 14 and HPV/p16-negative in 5 of these 19 patients, a distribution that was representative of the whole study group.

Scans judged as equivocal, and the Likert scale

When applying the Likert scale to scans judged as equivocal the separation of responders from nonresponders was enhanced but no significant difference regarding persistent or recurrent tumour was found between responders with a score of 2 or 3 and nonresponders with a score of 4 or 5 ($p=0.07$; Fig. 3).

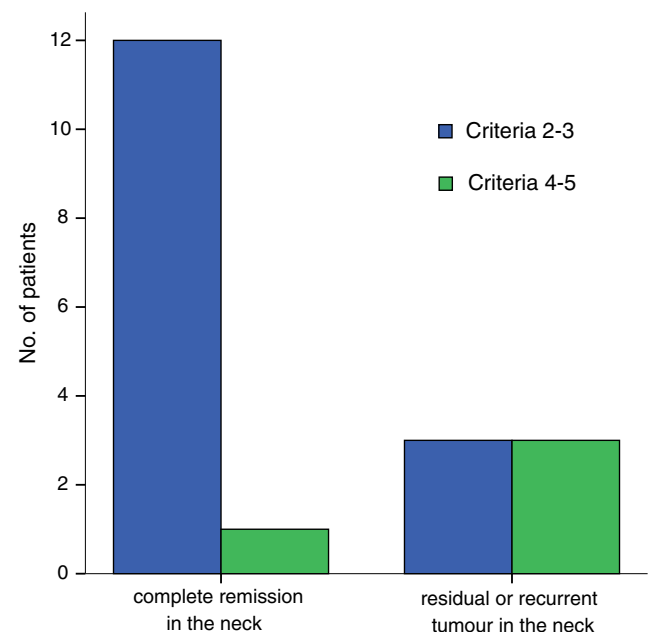


Fig. 3 PET scans judged as equivocal by visual inspection in 19 patients categorized as responders with complete remission in the neck (Deauville score 2 or 3), or as nonresponders with residual or recurrent tumour in the neck (Deauville score 4 or 5)

In 12 of the 13 patients with a lasting complete response, a score of 2 or 3 was recorded. In the remaining patient a score of 5 was recorded but no malignancy was found in the neck node specimen. The SUVmax in this patient was also high (7.1). Among the six patients who failed to achieve regional remission, a score of 4 or 5 was recorded in three and a score of 3 in the other three. The latter three patients included two with a tonsillar tumour in whom locoregional recurrences were diagnosed 9 and 19 months, respectively, after completion of therapy and one with a supraglottic tumour with a neck recurrence who underwent successful salvage surgery 12 months after therapy.

In the group of PET scans judged as equivocal there was no correlation between SUVmax and the presence or recurrence of tumour ($p=0.28$). The ROC analysis established that the Likert scale was superior to SUVmax in correctly identifying the level of metabolism that correlated to regional tumour control, with an AUC of 0.82 (CI 0.65–0.99) for the Likert scale, compared with 0.67 (CI 0.41–0.93) for SUVmax (Fig. 4).

Discussion

In the present study we found a good concordance between visual inspection, the Deauville criteria Likert scale and SUVmax. All methods of assessment predicted RC with high significance and almost equally well. The Likert scale also showed encouraging results in discriminating responders from nonresponders on PET scans judged as equivocal. SUVmax is an estimate of tumour metabolic activity and the most commonly used SUV parameter in PET/CT [9]. It is a single-voxel value adversely affected by noise [10] and does not reflect the metabolic activity of the tumour as a whole. We were able to demonstrate a significant difference in SUVmax between patients with and patients without remaining/recurrent tumour after completion of RT with or without chemotherapy.

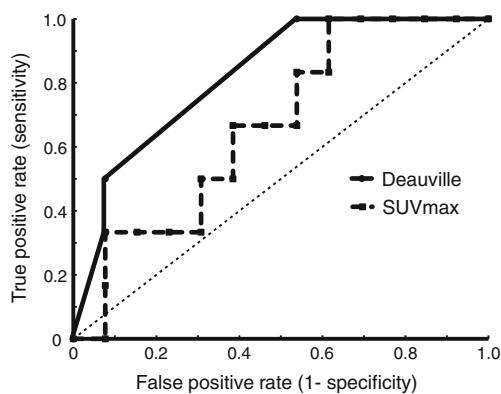


Fig. 4 ROC analysis of the 19 PET scans judged as equivocal by visual inspection. The Deauville criteria Likert scale is superior to SUVmax in identifying the level of metabolism that correlates with regional tumour control

However, measurement of SUV is affected by technical, biological and physical factors [18], and despite an attempt to establish common criteria there are still many different ways of calculating and presenting SUV. Published cut-off values are specific to the method and to the institution [9]. In this study, as also shown previously, SUV did not add any extra value compared with visual inspection in the clinical setting [12, 19].

A recently published study focused on a triphasic delayed measurement of SUVmax with the main aim of reducing the number of false-positive and equivocal scans after therapy [20]. The authors presented different FDG kinetics in inflammatory lesions compared with residual tumour 12 weeks after therapy completion and were able to improve the PPV. One benefit of the dynamic PET study would be that each patient served as his or her own control, which would reduce inter-institutional variations.

There is so far a lack of consensus on how to qualitatively assess and report PET scan results. In Hodgkin lymphoma the Deauville criteria Likert scale has been validated and adopted in multicentre trials [13, 17, 21]. In 48 patients with head and neck cancer, Krabbe et al. used a five-point scale in a serial PET evaluation, 3, 6, 9 and 12 months after treatment and demonstrated an overall PPV of 51 % and a NPV of 100 % using routine clinical follow-up as the reference [22]. Marcus et al. recently introduced and validated the Hopkins criteria, a five-point scale very similar to the Deauville criteria [14]. They achieved substantial interreader reliability and the criteria corresponded to OS and showed a PPV of 71.1 % and a high NPV of 91.1 %. The corresponding figures in our study were 68.7 % and 86.4 %. In a prospective PET study, Porceddu et al. based qualitative interpretation on focal uptake assessed in relation to uptake in adjacent tissue and the liver [3]. The PET scans were grouped into three different categories: “positive” (for remaining tumour), “negative” and “equivocal”. In the long-term follow-up, NPV was as high as 97.1 % even though all recurrences, no matter when they occurred, were included in the “false-negative” group [5]. As mentioned above, the NPV of the Likert scale used in the present study was lower (86.4 %). This may, at least partly, have been due to the early scheduled PET scans (6 weeks after therapy) as compared with 12 weeks or the study by Marcus et al. in which more than 50 % of the patients had their PET scan scheduled 13–24 weeks after treatment. Recent studies and meta-analyses have demonstrated a higher accuracy of PET scans scheduled later than 7 weeks after treatment [23, 24]. Treatment regimens also differed. In our setting, single modality treatment with RT was used in >90 % of the patients in contrast to similar studies in which chemoradiotherapy is more frequent. Less use of concurrent chemoradiotherapy might have led to more regional recurrences. In the present study the PET results were also categorized as false-negative if recurrent cancer was found at any stage during the follow-up

period in contrast to the studies by Krabbe et al. and Marcus et al. in which a 6-month limit for false-negative scans was applied [22, 14].

As reported in the Results, no posttreatment PET scan scored 1 on the Deauville scale. Score 1 is defined as “no FDG uptake”, which is a category of limited clinical value. A likely cause of no FDG uptake could be complete necrosis of the neck node, but this was not encountered in our patients. Cut-off values can be determined in different ways. We chose to use ROC data to define cut-off values for SUVmax and the Likert scale. Regarding SUVmax a cut-off value of 2.0 was the most favourable, giving a PPV of 37 % and a NPV of 93 %. Cut-off values used in previous studies vary but are often between 2.5 and 3.0; however, in our setting this range was less predictive. The outcome analysis is highly dependent on the chosen cut-off value. From a clinical point of view we consider a high NPV preferable to a high PPV, and the cut-off values were determined based on this preference.

The matter of equivocal PET scans has been debated and the difficulties acknowledged elsewhere [25, 26]. The Likert scale and visual inspection with specific reading (i.e. equivocal PET scans regarded as negative) showed similar accuracy. However, the benefit of adding a Likert scale to the PET report is apparent as responses can be categorized, and are distinct and easily interpreted by the oncologist or head and neck surgeon who have to act upon the PET result. In the present study, 18 % of scans (19 of 105 PET2 scans) were clinically reported as equivocal by visual inspection. The number of equivocal scans is similar to that found in previous studies [27]. By considering the equivocal PET scans together with those scored as 2 or 3 (responders) or those scored as 4 or 5 (nonresponders), we were able to correctly categorize 15 of the 19 patients. In three of the four remaining patients, categorized as responders, we diagnosed recurrences >9 months after therapy completion. In the fourth patient, we cannot explain why a visually clearly positive PET scan, a Deauville score of 5 and SUVmax >7 did not correspond to a positive pathology report. The group of PET scans assessed as equivocal was small and statistics should be interpreted with caution. However, the Likert scale adequately categorized 79 % of the equivocal PET scans, which is encouraging. In the group of PET scans judged as equivocal, SUVmax provided no additional predictive value.

In this study we chose to focus on the neck nodes with the highest FDG uptake in relation to treatment response. We did not evaluate the primary site response but it would be of interest to investigate whether using the Likert scale could minimize the number of equivocal scans here as well [28].

Conclusion

A Likert scale, in this case the Deauville criteria, for assessment of FDG metabolism in neck nodes following organ

preservation therapy in head and neck squamous cell carcinoma is a promising tool to reduce to a minimum the problem of PET scans judged as equivocal. All patients with a posttreatment PET scan scored as 4 or 5 on the Likert scale are at great risk of harbouring remaining tumour cells and should be scheduled for neck dissection. There was an association between visual inspection, SUVmax and the Deauville criteria in the evaluation of the neck node response after RT. All methods adequately identified responders and nonresponders and predicted RC. A Likert scale seems to satisfy the requirements for a common qualitative way of assessing PET scans and also the reporting of PET results in the evaluation of head and neck cancer treatment. Consensus regarding qualitative assessment would facilitate PET reporting in clinical practice and enable comparative studies between patients and institutions.

Compliance with ethical standards

Funding This study was supported by grants from the Swedish foundation Acta Oto-Laryngologica, governmental funding of clinical research within Lund University Hospital and Region of Scania R&D funding.

Conflicts of interest None.

Ethical approval All procedures were performed in accordance with the ethical standards of the institutional and regional research committee and with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Pellini R, Manciooco V, Turri-Zanoni M, Vidiri A, Sanguineti G, Marucci L, et al. Planned neck dissection after chemoradiotherapy in advanced oropharyngeal squamous cell cancer: the role of US, MRI and FDG-PET/TC scans to assess residual neck disease. *J Craniomaxillofac Surg.* 2014;42(8):1834–9. doi:10.1016/j.jcms.2014.06.023.
- Schouten CS, Hoekstra OS, Leemans CR, Castelijns JA, de Bree R. Response evaluation after chemoradiotherapy for advanced staged oropharyngeal squamous cell carcinoma: a nationwide survey in the Netherlands. *Eur Arch Otorhinolaryngol.* 2015;272(11):3507–13. doi:10.1007/s00405-014-3389-4.
- Porceddu SV, Pryor DI, Burmeister E, Burmeister BH, Poulsen MG, Foote MC, et al. Results of a prospective study of positron emission tomography-directed management of residual nodal abnormalities in node-positive head and neck cancer after definitive radiotherapy with or without systemic therapy. *Head Neck.* 2011;33(12):1675–82. doi:10.1002/hed.21655.
- Passero VA, Branstetter BF, Shuai Y, Heron DE, Gibson MK, Lai SY, et al. Response assessment by combined PET-CT scan versus CT scan alone using RECIST in patients with locally advanced head and neck cancer treated with chemoradiotherapy. *Ann Oncol.* 2010;21(11):2278–83. doi:10.1093/annonc/mdq226.

5. Sjøvall J, Chua B, Pryor D, Burmeister E, Foote MC, Panizza B, et al. Long-term results of positron emission tomography-directed management of the neck in node-positive head and neck cancer after organ preservation therapy. *Oral Oncol.* 2015;51(3):260–6. doi:10.1016/j.oraloncology.2014.12.009.
6. Kim SY, Lee SW, Nam SY, Im KC, Kim JS, Oh SJ, et al. The feasibility of 18F-FDG PET scans 1 month after completing radiotherapy of squamous cell carcinoma of the head and neck. *J Nucl Med.* 2007;48(3):373–8.
7. Phelps ME, Huang SC, Hoffman EJ, Selin C, Sokoloff L, Kuhl DE. Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18)2-fluoro-2-deoxy-D-glucose: validation of method. *Ann Neurol.* 1979;6(5):371–88. doi:10.1002/ana.410060502.
8. Patlak CS, Blasberg RG, Fenstermacher JD. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J Cereb Blood Flow Metab.* 1983;3(1):1–7. doi:10.1038/jcbfm.1983.1.
9. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med.* 2009;50 Suppl 1:122S–50. doi:10.2967/jnumed.108.057307.
10. Vanderhoek M, Perlman SB, Jeraj R. Impact of different standardized uptake value measures on PET-based quantification of treatment response. *J Nucl Med.* 2013;54(8):1188–94. doi:10.2967/jnumed.112.113332.
11. Roh JL, Kim JS, Kang BC, Cho KJ, Lee SW, Kim SB, et al. Clinical significance of pretreatment metabolic tumor volume and total lesion glycolysis in hypopharyngeal squamous cell carcinomas. *J Surg Oncol.* 2014;110(7):869–75. doi:10.1002/jso.23729.
12. Ong SC, Schoder H, Lee NY, Patel SG, Carlson D, Fury M, et al. Clinical utility of 18F-FDG PET/CT in assessing the neck after concurrent chemoradiotherapy for locoregional advanced head and neck cancer. *J Nucl Med.* 2008;49(4):532–40. doi:10.2967/jnumed.107.044792.
13. Meignan M, Gallamini A, Haioun C, Polliack A. Report on the Second International Workshop on interim positron emission tomography in lymphoma held in Menton, France, 8–9 April 2010. *Leuk Lymphoma.* 2010;51(12):2171–80. doi:10.3109/10428194.2010.529208.
14. Marcus C, Ciarallo A, Tahari AK, Mena E, Koch W, Wahl RL, et al. Head and neck PET/CT: therapy response interpretation criteria (Hopkins criteria) – interreader reliability, accuracy, and survival outcomes. *J Nucl Med.* 2014;55(9):1411–6. doi:10.2967/jnumed.113.136796.
15. Hutchings M, Loft A, Hansen M, Pedersen LM, Buhl T, Jurlander J, et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. *Blood.* 2006;107(1):52–9. doi:10.1182/blood-2005-06-2252.
16. Sjøvall J, Wahlberg P, Almquist H, Kjellen E, Brun E. A prospective study of positron emission tomography for evaluation of neck node response 6 weeks after radiotherapy in patients with head and neck squamous cell carcinoma. *Head Neck.* 2015. doi:10.1002/hed.24021.
17. Gallamini A, Barrington SF, Biggi A, Chauvie S, Kostakoglu L, Gregianin M, et al. The predictive role of interim positron emission tomography for Hodgkin lymphoma treatment outcome is confirmed using the interpretation criteria of the Deauville five-point scale. *Haematologica.* 2014;99(6):1107–13. doi:10.3324/haematol.2013.103218.
18. Boellaard R. Standards for PET image acquisition and quantitative data analysis. *J Nucl Med.* 2009;50 Suppl 1:11S–20. doi:10.2967/jnumed.108.057182.
19. Fakhry N, Lussato D, Jacob T, Giorgi R, Giovanni A, Zanaret M. Comparison between PET and PET/CT in recurrent head and neck cancer and clinical implications. *Eur Arch Otorhinolaryngol.* 2007;264(5):531–8. doi:10.1007/s00405-006-0225-5.
20. Anderson CM, Chang T, Graham MM, Marquardt MD, Button A, Smith BJ, et al. Change of maximum standardized uptake value slope in dynamic triphasic [18F]-fluorodeoxyglucose positron emission tomography/computed tomography distinguishes malignancy from postirradiation inflammation in head-and-neck squamous cell carcinoma: a prospective trial. *Int J Radiat Oncol Biol Phys.* 2015;91(3):472–9. doi:10.1016/j.ijrobp.2014.11.002.
21. Barrington SF, Qian W, Somer EJ, Franceschetto A, Bagni B, Brun E, et al. Concordance between four European centres of PET reporting criteria designed for use in multicentre trials in Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging.* 2010;37(10):1824–33. doi:10.1007/s00259-010-1490-5.
22. Krabbe CA, Pruijm J, Dijkstra PU, Balink H, van der Laan BF, de Visscher JG, et al. 18F-FDG PET as a routine posttreatment surveillance tool in oral and oropharyngeal squamous cell carcinoma: a prospective study. *J Nucl Med.* 2009;50(12):1940–7. doi:10.2967/jnumed.109.065300.
23. Leung AS, Rath TJ, Hughes MA, Kim S, Branstetter BF. Optimal timing of first posttreatment FDG PET/CT in head and neck squamous cell carcinoma. *Head Neck.* 2015. doi:10.1002/hed.24112.
24. Gupta T, Master Z, Kannan S, Agarwal JP, Ghosh-Laskar S, Rangarajan V, et al. Diagnostic performance of post-treatment FDG PET or FDG PET/CT imaging in head and neck cancer: a systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging.* 2011;38(11):2083–95. doi:10.1007/s00259-011-1893-y.
25. Prestwich R, Sen M, Scarsbrook A. Qualitative 18F-FDG PET/CT response evaluation after chemotherapy or radiotherapy for head and neck squamous cell carcinoma: is there an equivocal group? *J Nucl Med.* 2014;55(12):2081. doi:10.2967/jnumed.114.148114.
26. Subramaniam RM. Reply: Qualitative 18F-FDG PET/CT response evaluation after chemotherapy or radiotherapy for head and neck squamous cell carcinoma: is there an equivocal group? *J Nucl Med.* 2014;55(12):2081–2. doi:10.2967/jnumed.114.148379.
27. van der Putten L, Hoekstra OS, de Bree R, Kuik DJ, Comans EF, Langendijk JA, et al. 2-Deoxy-2-[F-18]FDG-PET for detection of recurrent laryngeal carcinoma after radiotherapy: interobserver variability in reporting. *Mol Imaging Biol.* 2008;10(5):294–303. doi:10.1007/s11307-008-0154-3.
28. Sjøvall J, Brun E, Almquist H, Kjellen E, Wahlberg P. Radiotherapy response in head and neck cancer – evaluation of the primary tumour site. *Acta Otolaryngol.* 2014;134(6):646–51. doi:10.3109/00016489.2014.895038.