Comparison of $^{18}$F-FDG-PET/CT and CT Tumour Volume Delineation in Head and Neck Cancer Patients

Besenyi Z$^1$, Hideghety K$^1$, Lengyel Z$^1$, Farkas SUI$^1$, Bakos A$^1$, Sipka G$^1$, Séra T$^1$ and Pávics L$^1$

$^1$University of Szeged, Department of Nuclear Medicine, Szeged, Hungary
$^2$University of Szeged, Department of Oncotherapy, Szeged, Hungary
$^3$Positron Diagnostics Ltd., Budapest, Hungary

Abstract

Introduction: Aim of the current study was to compare gross tumour volume delineation for radiation therapy planning by PET/CT and CT scan in head and neck tumour patients.

Methods: 70 oncological patients with primary head and neck cancer were enrolled in the study. CT and $^{18}$F-FDG-PET/CT scans were performed within 3 weeks of enrolment in the planned irradiation position. For radiation therapy planning delineation of the target volumes was performed manually both in conventional topometopic CT slides and in FDG-PET/CT images. Gross tumour target volume was calculated (GTVcm$^3$) using both modalities. Numerical and geometrical (intersection divided union ratio) comparisons were assessed. Intraobserver, interobserver, and intermodality variation analyses were performed.

Results: Intra- and interobserver agreement (intraclass correlation coefficients: 0.99 and 0.97) of tumour volume delineation with $^{18}$F-FDG-PET/CT was excellent. Radiotherapy target volume as assessed by FDG-PET/CT differed in 98% (69/70) from target volumes calculated by the CT scans. According to the metabolic information, the target volume was larger in 12 cases (18%) and smaller in 58 cases (82%). Significant differences were found between CT and PET/CT based tumour volume (paired t-test, p=0.0001). Intersection divided union ratio showed low overlap (0.32).

Conclusions: Tumour delineation for radiation therapy planning with PET/CT shows good reproducibility in patients with head and neck cancer. CT based morphologic information compared to PET/CT overestimates the tumour volume in larger lesions and may underestimate it in smaller target lesions. Differences are not evident in tumour volume, but in geometry: Morphology of tumour volume and position.

Keywords: FDG-PET/CT; Computer tomography; Head and neck cancer; Tumour contouring; Radiation therapy planning

Introduction

Head and neck tumours are detected mostly in an advanced stage and have poor prognosis. Due to anatomic complexity of the region complete surgical resection is rarely possible [1]. These in most cases histologically squamous cell tumours are sensitive to radiation therapy. Prospective and randomized clinical trials have confirmed the central role of radiation therapy in the treatment of disease [2]. Precise radiation therapy treatment planning of tumours in this region is especially important as several organs of vital importance are located close to each other in a small area [3].

There is a large amount of evidence proving the efficacy of fluorodeoxyglucose positron emission tomography ($^{18}$FDG PET) during detection of unknown primary, staging of disease, restaging after salvage surgery or neck block dissection as well as in assessment of response to therapy in head and neck cancer [4-7]. There is also growing interest in utilizing $^{18}$FDG PET in growth target volume delineation for radiation therapy planning in this sensitive region [8]. $^{18}$FDG PET offers delineation and precise contouring of the biologically active tumour -based on metabolic information- essential for effective treatment. It also allows for potential identification of lymph nodes missed by conventional computer tomography (CT) and magnetic resonance imaging (MRI) imaging [9]. Nevertheless the use of $^{18}$FDG PET in this regard is still under study, due to lack of good spatial and temporal resolution, questionable reproducibility [10]. According to current guidelines target volume delineation for radiation therapy planning is recommended by CT supplemented by MRI certain cases [10].

Integrated $^{18}$FDG PET/CT imaging with combined morphological and metabolic information has the potential of improving radiotherapy planning, gross tumour volume (GTV) delineation. However there is currently no consensus on methods of delineation, volume definition or in regards to overall utility of $^{18}$FDG-PET/CT scans in radiotherapy planning of in head and neck cancer patients. This field remains an active area of research [11]. The present work aims to study the reproducibility of tumour volume determination by $^{18}$FDG-PET/CT and to compare it to the current standard contouring by CT scan in head and neck cancer patients.

Patients and Methods

We enrolled 70 patients undergoing radiotherapy therapy planning for head and neck cancer (Group 1) in the current study. The average age of the patients was 58 ± 9 years (19-77 years). 20% of the studied patients (14/70) were women and 80% (56/70) were men. Histologic distribution was the following: (94.3%; 66/70 patients) squamous
cell carcinomas, 3 adenocarcinomas (4.2%), and 1 patient had neurofibrosarcoma (1.5%).

Interobserver and intraobserver variance analyses were performed on the ¹⁸F-FDG-PET/CT scans of 16 randomly selected patients (Group 2). In this group, the average age of patients was 58 years (44-73 years), and the male to female ratio was 16:0. Histology examination confirmed squamous cell carcinoma in all 16 patients. The locations of tumours are listed in Table 1.

Image acquisition protocol was the following. During patient preparation the head and neck region was positioned with four or five-point thermoplastic masks. Treatment planning CT (topo CT) was carried out according to standard protocols (Siemens Emotion 6 CT), 512*512 matrix, and 5-3 mm slice thickness. After CT imaging, ¹⁸F-FDG-PET/CT scan was performed within 3 weeks in the same patient position using thermoplastic fixation (SIEMENS Biograph 6 HD, Siemens Knoxville TN). After 6-hours fasting 3.7 MBq/kg ¹⁸F-FDG radiopharmaceutical was administrated intravenously. Blood glucose levels were checked beforehand (capillary blood via fingerstick, serum blood glucose <8 mmol/L). Imaging of the head and neck region was carried out 60 minutes after the administration of the radioisotope in the irradiation position, which was identical to the patient’s topo CT positioning with thermoplastic mask fixation. PET scan data collection was completed by plain low-dose CT (130 kV and 70 mAs) mapping. The PET scan was performed in three dimensional (3D) modes for 2 to 4 minutes per bed position, and the field of view was 21 cm with 25% overlap. The 3D images were reconstructed by using an iterative process with resolution enhancement technology (point spread function correction) in a 336 × 336 pixel matrix. After imaging the head and neck region, and removing the mask, a routine whole-body mapping was performed extending from the skull base to the upper third of the thighs.

GTV was manually delineated in each patient with both modalities. GTV selection for treatment planning was manually performed on both the conventional CT based topometric slides (GTVCT) and the PET/CT registered images (GTV PET) by a skilled radiologist and by a nuclear medicine specialist independently. XIO 3D planning system (Elekta Ltd, Crawly, UK) was used for evaluation. GTVCT contouring on the radiation planning CT was performed according to the guidelines of the Department of Oncotherapy. Available images of former contrast CT or MRI scans were reviewed under the supervision of a radiologist. Tumour volumes were recorded and determined in units of cm³ (Figure 1).

In each case, the GTV volume based on the performed ¹⁸F-FDG-PET/CT scan was compared to the GTV provided by the CT scan and their difference was described both in cm³ and in percentage. For the geometric comparison of each tumour volume of interest (VOI), the Intersection/Union (I/U) ratio was calculated. The I/U quotient is a ratio which ranges from 0 to 1. In case of full overlap, the I/U is 1, whereas in full diversity, the I/U is 0. For example, I/U=0.33 means a 50% overlap of two equal volumes (Figure 2). The total patient population (Group 1), was divided into subgroups based on differences in GTV of the two modalities (Group A: GTV PET-CT>GTVCT; Group B: GTV PET-CT ≤ GTVCT). The before mentioned volumetric and geometrical comparisons were repeated in both subgroups.

<table>
<thead>
<tr>
<th>Tumor localization</th>
<th>Patient Group 1 n=70</th>
<th>Patient Group 2 n=16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral cavity and larynx</td>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td>Nasopharynx</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Oropharynx, hypopharynx</td>
<td>33</td>
<td>6</td>
</tr>
<tr>
<td>Nasal cavity and nasal sinuses</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1: Primary localization of head and neck tumours. Patient Group 1: The total population; Patient Group 2: Subgroup of the total population randomly selected for intra- and interobserver analyses.

Figure 1: Interobserver analysis of ¹⁸F-FDG PET/CT gross tumour volume delineation. Axial (A), sagittal (B) and coronal slices (C) of fused ¹⁸F-FDG PET/CT images and 3D Gaussian filtered contours (D) of a head and neck cancer patient. Each observer delineated the primary tumour. Contours from the three observers are shown in different colors. Red: 2.08 cm³, yellow: 1.38 cm³, blue: 0.87 cm³, I/U ratio: 0.41.
For 18F-FDG-PET/CT intraobserver and interobserver variation analyses of manual contouring 20 solid lesions with different size were evaluated in the scan images of 16 randomly selected patients (Group 2). Three experienced observers (A, B, C) determined the macroscopic tumour volume independently, with the use of the Medical Imaging Interaction Toolkit (MITK, German Cancer Research Centre) software. Primary tumour contouring was performed in the axial slides of 18F-FDG PET scans. The contours were finalized with the help of CT images taking patient anatomy into consideration, and the GTV was also determined in cm³. Observer A repeated the manual GTV delineation described above on two more occasions in random order within a week. Evaluating physicians were blinded to the patient’s clinical data and results of other imaging tests. To determine geometric interobserver and intraobserver differences, the I/U ratio was also calculated in both in the case of observers A, B and C (Figure 3), and in settings A1, A2 and A3.

Data analyses was performed with R Statistical Software version 3.4.3 (The R foundation for statistical computing, Vienna, Austria). Data were reported as mean ± standard deviation and range in parenthesis wherever appropriate. P<0.05 was considered to be statistically significant. Values for I/U were interpreted in the following way: 0.8-1: excellent, 0.5-0.8: good, 0.33-0.50: moderate and 0.0-0.33: low. Applied statistical methods: Paired t-test and unpaired two-sample Wilcoxon Rank test for comparisons of GTV CT and GTV PET-CT as applicable and intraclass correlation for inter and intraobserver differences.

The scheme of experiments complied with the ethics of research. It agreed with the declaration of the Medical World Federation proclaimed in Helsinki in 1964, therefore, the Human Investigation review Board did not raise any objection to it from ethical point of view and supported it. Reference number: 3647, name of the ethics committee: Human Investigation Review Board, University of Szeged, Hungary.

Results

Comparison of radiation therapy target volumes delineated by 18F-FDG PET/CT and CT

Tumour volumes were identical in 1 out of 70 cases (1%), decreased in size in 57 cases (81%) of the cases, and increased in 12 cases (18%) of the cases, when the results acquired by the 18F-FDG-PET/CT scan were compared to the results defined by topo-CT scan. Tumour volumes defined by the two different imaging modalities were non-identical in 99% of the cases. In 11 out of 70 cases, the difference in the percentage of volumes was lower than 10%. The difference in the percentage of volumes was higher than 10% in 59 out of 70 patients (84% of all patients). Tumour volume defined by the 18F-FDG-PET/CT was significantly smaller (paired T test, p<0.0001) than tumour size defined...
± 0.12 in overall comparison of the members in the three series. For intraobserver comparison difference in volumes (A1, A2, and A3) was 1.12 ± 1.1 cm³ that is 12.31 ± 7.4%. The intraclass correlation coefficient was 0.9903. The average I/U ratio concerning A1-A2-A3 combined was 0.76 ± 0.08.

**Discussion**

In the current study we report significant differences in tumour target volume and geometrical position as assessed by 18F-FDG-PET/CT and CT in patients with head and neck cancer undergoing radiation therapy planning. Defining the “real” volume of malignant tumours precisely in a standard, reproducible manner is still a challenge for diagnostic experts and radiotherapists. The variability in the defined tumour volume is particularly important in case of head and neck tumours, where there is a relatively large irradiation dose (70 Gray) applied in a small area. An inappropriately identified target volume results in an extensive marginal error of irradiation. Applying decreased irradiation to certain parts of the target volume, may lead to worsening prognosis of the disease (marginal recurrence). On the other hand inaccurate protection of tumour free areas, may lead to increasing severity of adverse effects [12]. Morphologic studies (CT and MRI scans) and functional imaging techniques (18F-FDG-PET) provide different information about the studied region, the tumour, and adjacent tissues. While computer tomography provides...
Validation of tumour volumes by pathology specimens using several imaging modalities has been attempted [13]. Compared to CT, MRI and FDG-PET scan results of 29 patients with head and neck tumour to their surgical specimen. The size of GTVs delineated by CT scan ranged between 1.9 and 137.7 cm³, which was similar to our findings. However, in our study, slightly larger tumours were identified. According to Daise results, the surgical specimens were significantly smaller in size compared to the results of all three imaging studies (CT=20.8 cm³, MRI=23.8 cm³, PET=16.3 cm³, and surgical specimen=12.6 cm³). FDG-PET scan revealed significantly smaller volume than the other two modalities, and it showed the best correlation with pathology results. In terms of GTV<sub>PET</sub> and GTV<sub>CT</sub>, it was identical to our results [14], compared CT and 18F-FDG-PET/CT scan images of 16 patients to tumour volume determined by the pathology. According to their study, both CT and 18F-FDG-PET/CT scan overestimated the tumour volume compared to the specimen; however, 18F-FDG-PET/CT scan results were closer to those of the specimens. In the meta-analysis of Huang et al. [15], data of 717 patients with head and neck tumour were processed; both the irradiation target volume and the pathology tumour volume were determined for all patients by at least two imaging modalities (MRI, CT, or 18F-FDG-PET/CT). Their results showed that the CT-based GTV was larger than the MRI and PET scan-based volume, while GTV<sub>PET</sub> significantly correlated with the pathology tumour volume.

The role of 18F-FDG-PET/CT scan in head and neck tumour in the clinical setting of irradiation therapy has also been studied, albeit in smaller number of patients than in our study and with heterogeneous results [16-18]. Our results show significant differences in GTVs defined by functional and morphological methods, similarly to previous publications. We also report that heterogeneity of these differences is in part related to tumour size. In the current patient group larger PET/CT derived GTV-s as compared to CT were less common and only found in small lesions as defined by CT. Whereas larger CT derived GTV-s as compared to PET/CT were more common, larger, however also heterogeneous in size (Table 2). Further examining the effects of tumour size a cut-off point of 37 cm³ CT derived GTV could be established. In all cases where tumour size was larger than the cut off volume, metabolic information of 18F-FDG-PET/CT reduced radiotherapy gross target volume. Possible cause behind this phenomenon may be inflammation and oedema surrounding larger tumours falsely included by CT in the primary tumour volume. Ciernik et al. also raised the question of tumour size effecting heterogeneous results but failed to show significant differences probably due to small sample size [16]. Paulino et al. analyzed 40 patients with head and neck cancer and found similarly to our results that the tumour size reduced in most cases as defined by 18F-FDG-PET/CT scan (in 30 out of 40 patients), whereas an increase in GTV compared to the results of the CT scan was identified only in 7 patients. In accordance with our experience, this study also found that in the case of relatively large tumours, 18F-FDG-PET/CT reduced the GTV<sub>CT</sub> in 75% of the cases [17].

Along with others our results support the view that morphological tumour volume delineation is significantly different from the metabolic tumour volume, and the difference is even more marked if volume geometry is also considered. The accuracy of tumour delineation by the different imaging techniques is not only determined by tumour volume given in cm³. Two tumours of completely equal volume (e.g., two tumours of 2 cm³) may be different in shape and localization. During comparison measurements, not only the difference in the size (cm³) of the tumour should be considered, but also differences in localization and shape. We used, similarly to several other authors, the I/U ratio to compare the individual tumour volumes, and identified a great variety of the quotient [19-21]. Anderson et al. analyzed the manually delineated tumour volumes based on findings of CT, FDG-PET/CT and MRI (post contrast T1) scans performed for establishing the plain irradiation treatment plan in 14 patients with head and neck tumour [21]. Similarly to their work, we also examined the geometrical match between the results of the different modalities besides and found low overlap.

Contouring which may also influence results is complicated by several factors in case of both imaging technique. In CT scans, the consequences of the postoperative condition may leave asymmetry and scarring. In metabolic FDG-PET scans, the areas with physiological FDG uptake may also show diversity. In the head and neck region, there may be a physiological, moderate increase in glucose metabolism in the adenoid and lymphoid tissues, in the nasopharynx, soft palate, and tonsils (lingual and palatine), and in the salivary glands. An increased uptake of the radiopharmaceuticals in the vocal cords bilaterally and in the tongue might refer to speaking (muscle activity) during tracer accumulation time, and symmetrical uptake in the eye muscles may refer to eye movements, which are both considered being normal. For these reasons, the assessment of 18F-FDG-PET and CT scans, and the recognition of abnormalities in the head and neck region require skills and practice in hybrid imaging. Riegel et al. found tumour contouring based on 18F-FDG-PET/CT scan to be more reproducible, as opposed to the target volume defined solely by CT scan [22]. Several further studies have shown in case of head and neck tumours co-registered anatomic and metabolic images provide significantly more accurate tissue contouring than CT scan alone [23]. This is also true for imaging of specified organs at risk [24]. Vinod et al. processed the data of 56 studies in their met a-analysis, and found that observer-dependent differences can be reduced by learning, teaching, and using uniform standards [20]. Semiautomatic contouring based on various techniques seems also to reduce the differences [25]. Our intraobserver variance analysis showed that 18F-FDG-PET/CT manual contouring was well-reproducible if the assessment was performed by the same person. Similarly to data in literature, interobserver analyses reproducibility was somewhat lower but still excellent [18,20]. Furthermore, intra and interobserver comparisons were also made taking into account not only volumetric but geometric results (tumour overlap) as well. Intra and interobserver agreement remained good by 18F-FDG-PET/CT manual contouring even when geometric differences were considered.

The current study shares all the inherent limitations of retrospective studies. This study was aimed at the determination of the primary tumour volume. The effect of the PET/CT scan was not evaluated separately on the changes in staging, and whether previously undetected potential lymph node metastases were revealed. Due to the heterogeneity of tumour stages in the relative low number of patients the study was underpowered for clinical follow-up. Although several steps were taken to ensure uniform patient preparation and image acquisition, small differences caused by patient movement/position alternation cannot be ruled out.
In conclusion manual tumour contouring of 18F-FDG-PET/CT images provides excellent intra- and interobserver reliability in head and neck cancer patients. CT based morphologic information compared to PET/CT overestimates the tumour volume in larger lesions and may underestimate it in smaller target lesions during radiation therapy planning. The differences in the results of functional imaging studies and the CT scans are not merely due to the differences in the size, but to the geometrical differences in the morphology of tumour volume and position. Prospective randomized large studies are needed to evaluate if these imaging modality based differences in radiation therapy planning influence clinical outcomes in head and neck cancer patients.

**Declarations**

The Authors declare that they have no conflicts of interest.

**References**