Original article

18F-FDG-PET-based tumor delineation in cervical cancer: Threshold contouring and lesion volumes

F. Erlich a, C. Camisão b, A. Nogueira-Rodrigues c, S. Altino d, C.G. Ferreira c, M. Mamede c,e,*

a Radiation Oncology Service, National Cancer Institute of Brazil, Rio de Janeiro, Brazil
b Radiology Service, National Cancer Institute of Brazil, Rio de Janeiro, Brazil
c Clinical Research Coordination, National Cancer Institute of Brazil, Rio de Janeiro, Brazil
d Nuclear Medicine Section, Samaritano Hospital, Rio de Janeiro, Brazil
e Molecular Imaging Center, Universidade Federal Minas Gerais, Belo Horizonte, Brazil

A R T I C L E   I N F O

Article history:
Received 13 March 2012
Accepted 11 June 2012

Keywords:
18F-FDG-PET
Cervical cancer
Gross tumor volume
Tumor delineation

A B S T R A C T

Objective: To evaluate a semi-automated PET-image tumor segmentation algorithm for gross tumor volume (GTV) delineation in patients with locally advanced cervical cancer.

Material and methods: Thirty-two patients with locally advanced cervical cancer were retrospectively evaluated. Semi-automated PET-image-based GTV delineation was applied using a previous established algorithm (GTV 2SD) and 2 fixed threshold-based methods (GTV 40% and GTV 50%). GTV 2SD was determined as the pixel with the mean value plus 2-standard deviation of the liver intensity, and GTV 40% and GTV 50% with 40% and 50% of the maximum tumor intensity (Tmax), respectively. The derived volumes were then compared with the GTVs generated manually using MR (GTV Sem).

Results: The mean value of GTV 2SD, GTV 40%, and GTV 50% was 85.3 cc, 16.2 cc and 24.1 cc, respectively. Good agreement was noticed between GTV 2SD and GTV Sem (ρ = 0.88). GTV 40% and GTV 50% showed weaker correlation with GTV Sem (ρ = 0.68 and ρ = 0.71, respectively).

Conclusions: This study provides preliminary evidence that metabolic tumor volume delineation is feasible using computer-generated measurements in 18F-FDG PET images. Generation of PET-based tumor volumes is affected by the choice of threshold level used. Metabolic tumor bulk calculated using the pixel with the mean value plus 2-standard deviations of the liver intensity (GTV 2SD) correlates better with the MR-derived tumor volumes. The method is a simple and clinically applicable approach to generate PET-derived GTV for radiation therapy planning of cervical cancer.

© 2012 Elsevier España, S.L. and SEMNIM. All rights reserved.

Palabras clave:
18F-FDG-PET
Cáncer cervical
Volumen tumoral grueso
Delimitación tumoral

Delineación tumoral basada en 18F-FDG-PET en el cáncer cervical: delimitación del contorno umbral y volúmenes de la lesión

R E S U M E N

Objetivo: Evaluar el algoritmo de segmentación tumoral de la imagen PET semiautomatizada para delinear el volumen tumoral grueso (GTV) en pacientes con cáncer cervical localmente avanzado.

Material y métodos: Se evaluó retrospectivamente a 32 pacientes con cáncer cervical localmente avanzado. Se utilizó la delineación GTV basada en imagen PET semiautomatizada, utilizando un algoritmo previamente establecido (GTV 2SD) y 2 métodos basados en umbrales fijos (GTV 40% y GTV 50%). Se calculó GTV 2SD como el pixel con el valor medio más 2 desviaciones estándar de la intensidad del hígado, y GTV 40% y GTV 50% con el 40% y el 50% de intensidad tumoral máxima (Tmax), respectivamente. A continuación se compararon los volúmenes derivados con los GTV generados manualmente, utilizando RM (GTV Sem).

Resultados: El valor medio de GTV 2SD, GTV 40% y GTV 50% fue de 85.3 cc, 16.2 cc y 24.1 cc, respectivamente. Se halló una buena concordancia entre GTV 2SD y GTV Sem (ρ = 0.88). GTV 40% y GTV 50% mostraron una menor correlación con GTV Sem (ρ = 0.68 y ρ = 0.71, respectivamente).

Conclusiones: Este estudio prueba de modo preliminar que la delimitación del volumen tumoral metabólico es posible utilizando las mediciones generadas informáticamente en las imágenes de 18F-FDG PET. La generación de los volúmenes tumulares basados en PET se ve afectada por la elección del nivel de umbral utilizado. El grueso del tumor metabólico calculado utilizando el pixel con el valor medio más 2 desviaciones de la intensidad del hígado (GTV 2SD) guarda una mejor correlación con los volúmenes tumurales derivados de RM. El método constituye un enfoque simple y clínicamente aplicable para generar el GTV derivado del PET, para la planificación de la terapia de radiación del cáncer cervical.

© 2012 Elsevier España, S.L. y SEMNIM. Todos los derechos reservados.

* Corresponding author.
E-mail address: mamede.mm@gmail.com (M. Mamede).

2253-8099/© 2012 Elsevier España, S.L. and SEMNIM. All rights reserved.
Introduction

Cervical cancer is the fifth most deadly cancer in women worldwide, being the cause of 286,451 estimated deaths in 2008.\(^1\) Based on NCI’s SEER Cancer Statistics Review, it was estimated that 11,270 women would be diagnosed with and 4070 women would die of cancer of the cervix uteri in 2009.\(^2\) In Brazil, 18,430 new cases are estimated in 2010, with a 5-year overall survival rate of 49%.\(^3\) Locally advanced disease is still the most prevalent presentation, with the standard of care being concurrent chemotheraphy and radiation therapy.

Historically radiation therapy planning is done using computerized tomography (CT) images. Besides CT, magnetic resonance (MR) is also very useful to show more anatomic detail, especially soft tissues. However, CT and MR only demonstrate anatomic tumor extent, and do not reflect biological tumor behavior.

Fluorodeoxyglucose (FDG) labeled with fluor-18 (\(^{18}\)F-FDG), an analog of glucose, has been used extensively to diagnose, stage, re-stage and follow up patients with cancer.\(^4\) \(^{18}\)F-FDG PET images reflect the glycolytic metabolic tumor behavior, which is important for treatment planning. Recently, \(^{18}\)F-FDG PET has also been applied in several studies to the radiation therapy planning for different tumor types.\(^5\) They have shown significant alterations to the gross tumor volume (GTV) delineation and consequently to the radiation therapy fields.\(^5\)–\(^10\)

Faria et al. showed that PET altered the contour in 18 (56%) of 32 cases of non-small-cell lung cancer compared to CT alone.\(^7\) Van Loon et al. showed that incorporating \(^{18}\)F-FDG PET information in radiotherapy planning for patients with limited-disease small-cell lung cancer changed the treatment plan in 24% of patients compared to CT.\(^8\) Grisby showed that \(^{18}\)F-FDG PET can be useful to target active metabolic disease and aid radiotherapy planning in cervical cancer.\(^9\)

Although these initial studies showed the great advantages of incorporating \(^{18}\)F-FDG PET into radiotherapy treatment planning, there is no consensus yet on the use of these images for target delineation. Fixed-level threshold techniques have been used to delineate GTV. These fixed-level thresholds are not always accurate when compared to CT or MR defined volumes, and depending on the level used, significant variability on GTV values has been reported.\(^11\),\(^12\) Biehl et al. showed that no single threshold delineating the PET-GTV provides accurate volume definition, compared with that provided by the CT-GTV, for the majority of NSCLCs.\(^13\) In another study in patients with head and neck cancer, the authors demonstrated that PET-based tumor volumes are strongly affected by the choice of threshold level, leading to a significant dosimetric impact.\(^14\)

However, PET-defined tumor volumes in radiotherapy treatment planning can significantly affect the contour of the GTV due to variation in setting image signal thresholds; its incorporation can reduce considerably the inter- and intra-observer evaluations. Thus, there is no consensus yet as how the PET-images should be used to generate the GTV volumes for radiation therapy planning. The goal of this study is to evaluate a semi-automated segmentation algorithm for GTV delineation by comparing the metabolic PET-derived volumes with the anatomic MR-based volumes in cervical cancer patients. The originality of this manuscript relies on its strength to correlate well with volumes generated from MR images and its simplicity which makes it suitable for routine clinical use.

Material and methods

Thirty-two patients with locally advanced cervical cancer were retrospectively evaluated. This cohort was part of a phase I/II trial of erlotinib combined to cisplatin and radiation.\(^15\),\(^16\) Staging was performed with clinical examination and all of the patients had CT, MR and \(^{18}\)F-FDG PET/CT scans before treatment, as well as cistoscopy and rectosigmoidoscopy. Locally advanced disease was defined as FIGO stage greater than IB2 and lesser than IVB. Table 1 shows the demographics of our patient cohort. The study protocol was approved by our local ethics committee.

MR examinations were performed using a 1.5 T superconducting magnet (Symphony, Siemens Medical Systems, Erlangen, Germany) covering the abdomen and pelvis with identical protocol. All patients received 20 ml of intravaginal gel applied via syringe. All patients fasted 4 h prior to the examination. No bowel relaxants were routinely administered. The acquisition protocol consisted on the following sequences: (1) spin-echo T1-weighted axial, repetition time/echo time (TR/TE) 5.2/2.6 ms, slice thickness 3–5 mm, interslice gap 0, matrix 512 × 512; 38 cm FOV; (2) fast spin-echo T2-weighted with a phased array coil in sagittal, coronal and axial planes, TR/TE 4500–5000/95–116 ms, slice thickness 3 mm, interslice gap 0.2 mm, matrix 512 × 512; 25 cm FOV; (3) turbo spin echo T2 with fat suppression axial, TR/TE 4500–5000/95–116 ms, slice 5 mm, interslice 0.2 mm, matrix 512 × 512, 30–45 cm FOV; and (4) T1-weighted spin-echo axial sequences before and after intravenous injection of gadolinium (0.1 mmol/kg) TR/TE 5.2/2.6 ms, slice thickness 3 mm, interslice gap 0, matrix 512 × 512; 35 cm FOV. None of the images were acquired using flat table.

The patients fasted for at least 4 h before the intravenous injection of \(^{18}\)F-FDG (783.5 ± 102.3 MBq). Serum glucose levels were measured before \(^{18}\)F-FDG injection. The scanning was performed ~60 min after the injection of \(^{18}\)F-FDG using a Biograph 2 PET/CT scanner (CT: Somatom sensation 2, PET: LSO crystal, Siemens Medical Solution, Malvern, USA) with a 162 mm axial field of view. The images were obtained with the patient in the supine position from the orbital-mental line to the proximal thighs. CT studies for attenuation correction and anatomic co-registration were performed following the administration of diluted barium oral contrast, and with the following imaging parameters: 130 kV, 25 mAs, 0.8 s per CT rotation, and a 5.0-mm slice thickness. The CT data were reconstructed in a 512 × 512 matrix using a filtered backprojection algorithm. Immediately after the CT scan, emission images were obtained in a 3D mode with 3 min acquisitions at each level (5–7 bed positions). The PET images were reconstructed in a 128 × 128 matrix using an iterative-ordered subsets expectation maximization (OSEM) algorithm (4 iterations, 8 subsets), yielding a volume of 47 slices with a voxel size of 3.37 mm × 3.37 mm × 3.37 mm. None of the images were acquired using flat table.

For volume delineation and measurement software developed at the National Institutes of Health (NIH) was used. MIAPA (Center

<table>
<thead>
<tr>
<th>Patient</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>54</td>
</tr>
<tr>
<td>Range</td>
<td>28–72</td>
</tr>
<tr>
<td>Tumor</td>
<td></td>
</tr>
<tr>
<td>Differentiation (WHO)</td>
<td></td>
</tr>
<tr>
<td>Grade 1</td>
<td>1 (3%)</td>
</tr>
<tr>
<td>Grade 2</td>
<td>21 (65%)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>10 (32%)</td>
</tr>
<tr>
<td>FIGO stage group</td>
<td></td>
</tr>
<tr>
<td>IIA</td>
<td>2 (6%)</td>
</tr>
<tr>
<td>IIB</td>
<td>11 (34%)</td>
</tr>
<tr>
<td>IIIA</td>
<td>2 (6%)</td>
</tr>
<tr>
<td>IIIB</td>
<td>16 (51%)</td>
</tr>
<tr>
<td>IVA</td>
<td>1 (3%)</td>
</tr>
</tbody>
</table>

Note: WHO: World Health Association; FIGO: International Federation of Gynecology and Obstetrics.
of Information and Technology-CIT, NIH, Bethesda, USA) software is a medical imaging tool used for image processing, segmentation, registration, fusion, and visualization. MIPAV is user-friendly freeware, available on the Internet and runs on several platforms (Windows, MAC, UNIX, and LINUX).

Semi-automated PET-image-based main tumor GTV delineation was applied using a previous established algorithm (GTV\textsubscript{2SD}) and 2 fixed threshold-based methods (GTV\textsubscript{40x} and GTV\textsubscript{50x}). Briefly, the user defined the voxel with the highest intensity value \( I_{\text{top}} \) within the lesion on the PET image. This intensity was set up as the maximum voxel on the threshold scale. The liver was chosen as the background region because of known adequate performance. Several liver regions of interest (ROI) were drawn in various transaxial slices, and the mean intensity and SD over all voxels in those ROIs were used for further calculation. The lower level of the threshold scale \( I_{\text{bottom}} \) was defined as a result of the difference between \( I_{\text{top}} \) and the sum of the mean intensity in the liver \( I_{\text{mean}} \) plus 2-SD \( I_{\text{2SD}} \): 
\[
I_{\text{bottom}} = I_{\text{top}} - (I_{\text{mean}} + I_{\text{2SD}}) 
\]
After defining \( I_{\text{top}} \) and \( I_{\text{bottom}} \), the user selects a point in the lesion (the “seed point”) and a 3D threshold-based region-growing algorithm segments the lesion of interest to include only pixels between \( I_{\text{top}} \) and \( I_{\text{bottom}} \). If a pixel’s intensity falls between \( I_{\text{top}} \) and \( I_{\text{bottom}} \), it is included in the region. Iterations stop when no more neighbors fall in the acceptable range. Afterwards, MIPAV generates the total metabolic volume of the tumor (GTV\textsubscript{2SD}), GTV\textsubscript{40x} and GTV\textsubscript{50x} were determined with 40% and 50% of \( I_{\text{top}} \), respectively. Then, volumes were compared with the GTVs generated manually using MR (GTV\textsubscript{MR}). For this purpose, T2 weighted axial images were used and the drawing was performed by a radiologist specialized in gynecological MR.

For comparison between the volumes, Spearman’s \( \rho \) correlation test was used in a SPSS 12.0 software (SPSS, Chicago, IL) for statistical analysis.

### Results

Between 2005 and 2008, 32 patients were eligible for analysis. There were predominantly moderately differentiated cancers, with 65% of the patients classified as WHO grade II. More than a half of the patients had IIIB FIGO stage (51%). Of these, 14 patients had disease extending to the pelvic wall and 3 had hydronephrosis (Table 1).

Tumor volume analysis is shown in Table 2. The GTV\textsubscript{MR} ranged from 3.3 cc to 246.7 cc, reflecting the heterogeneity of the patients and very different tumor sizes. Moreover, also the range of SUV\textsubscript{max} showed an important variation, with values ranging between 5.9 and 32.2. The mean volumes obtained with 2SD-algorithm threshold, fixed 40% and fixed 50% were 85.3 cc, 16.2 cc and 24.1 cc, respectively.

Spearman’s correlation analysis is shown in Fig. 1. A good correlation was achieved between the volume obtained in MR (GTV\textsubscript{MR}) and the volumes determined as the pixel with the mean value plus 2-standard deviation of the liver intensity (GTV\textsubscript{2SD}) in the \( ^{18} \text{F}-\text{FDG} \) PET images \((\rho = 0.88; p < 0.001)\), being closer to the 45° line. Both, the mean value and the volume range were well correlated, the GTV\textsubscript{2SD} being slightly over-estimated when compared to the GTV\textsubscript{MR}. However, weaker correlation was observed with the fixed threshold volumes \((\rho = 0.68 \text{ and } 0.71 \text{ for } 40\%- \text{ and } 50\%- \text{threshold, respectively})\). Fig. 2 shows typical relationships between PET-GTVs (GTV\textsubscript{2SD}, GTV\textsubscript{40x} and GTV\textsubscript{50x}) with that from MR (GTV\textsubscript{MR}).

As it can be observed in Table 2 and Fig. 1, the fixed level thresholds volumes tended to underestimate the MR-volume. Observing Fig. 3, we can infer that for the mean MR-volume achieved in our patients, the best fit threshold volume would be between 60% and 80% \((n = 17)\). Therefore, in order to generate better PET-image volumes for radiation planning, there is a need to adjust the threshold for each patient.

### Discussion

The recent approach to cervical cancer treatment has been the use of highly conformal Intensity Modulated Radiation Therapy (IMRT) and Image-Guided Radiation Therapy (IGRT) with concomitant dose escalation to the GTVs. These highly conformal treatments are very volume dependent, and non-accurate volume delineation could lead to geographical misses, increasing the chance of treatment failure, or overestimation of GTV, leading to more acute and chronic toxicities. Furthermore, molecular imaging with \( ^{18} \text{F}-\text{FDG} \) PET has been extensively used to delineate the GTV for radiation therapy planning. However, there is not a consensus on how the \( ^{18} \text{F}-\text{FDG} \) PET should be used for tumor-volume delineation, because of variations due to the threshold applied.

In the present study, we evaluated volumes of the main tumor generated by using a previously published algorithm (GTV\textsubscript{2SD}) in 32 patients with locally advanced cervical cancer and compared these results with those derived from MR. MR-derived

### Table 2

<table>
<thead>
<tr>
<th>Factor</th>
<th>Value (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MR</td>
<td>(cc)</td>
</tr>
<tr>
<td>Volume</td>
<td>59.6 (3.3–246.7)</td>
</tr>
<tr>
<td>( ^{18} \text{F}-\text{FDG} ) PET/CT</td>
<td></td>
</tr>
<tr>
<td>Liver mean</td>
<td>(counts)</td>
</tr>
<tr>
<td>(counts)</td>
<td>10.975 (7419–17337)</td>
</tr>
<tr>
<td>Tumor max</td>
<td>(counts)</td>
</tr>
<tr>
<td>SUV\textsubscript{max}</td>
<td>81.824 (34.458–143604)</td>
</tr>
<tr>
<td>Volume</td>
<td>(cc)</td>
</tr>
<tr>
<td>2SD</td>
<td>14.7 (5.9–32.2)</td>
</tr>
<tr>
<td>Fixed 40%</td>
<td>85.3 (5.2–299.7)</td>
</tr>
<tr>
<td>Fixed 50%</td>
<td>16.2 (2.1–82.2)</td>
</tr>
<tr>
<td>2SD</td>
<td>24.1 (3.0–94.5)</td>
</tr>
</tbody>
</table>

Note: SUV\textsubscript{max}: maximum standardized uptake value; 2SD: 2 standard deviation algorithm.

![Fig. 1.](image.png) Spearman’s correlation analyses. Note: MR: magnetic resonance.
tumor volumes were also compared with those obtained by using fixed level metabolic thresholds (40% and 50%). The study showed that the GTV_{2SD} worked better than the fixed level thresholds. Thus, the method provides preliminary evidence that tumor delineation might be feasible using computer-generated metabolic measurements and that it could be used for radiotherapy planning in cervical cancer patients.

\(^{18}\)F-FDG PET is being widely used to delineate GTV in radiotherapy planning. Currently, MR images are considered the best imaging modality to delineate lesion volumes in cervical cancer. However, MR relies only on tumor morphological patterns and does not provide metabolic information, which is known to be highly desirable to define and estimate tumor burden. This is especially important when using state-of-the-art radiotherapy techniques such as IMRT and IGRT.

The first report describing the use of \(^{18}\)F-FDG PET in radiotherapy planning was published by Kiffer et al. in 1998.\(^{20}\) The authors reported a 26.7% difference in the treatment fields when PET was compared to CT only.\(^{20}\) Since then, many other papers were published showing significant alterations in GTV and treatment fields with PET. In 2004, Bradley et al. proposed the term biological tumor volume (BTV) for PET-derived metabolic bulk.\(^{21}\) However, it still remained unclear what the real accuracy was between the traditional GTV obtained with CT and MR and the new BTV obtained with molecular imaging with PET. Only few studies dealt with this question, and most of them used fixed level thresholds, usually 40% or 50% of the maximum tumor intensity. However, as shown in Figs. 1–3, resulting GTVs could have important volume differences according to the window used for contour delineation.

Hong et al., in a trial published in 2008 in esophageal cancer patients, identified that different PET-based techniques can produce significantly different tumor volumes in a large percentage of them. They also concluded that using fixed level thresholds do not produce accurate correlations between CT and PET findings and, furthermore, proposed a semi-automated algorithm that showed the best correlation, defined by the volumes determined as the pixel with the mean value plus 2-standard deviation of the liver intensity.\(^{5}\)

In the present study, PET-image volumes generated using GTV_{2SD} algorithm correlated very well with MR-image volumes; while weaker correlations were observed when PET-image volumes were generated applying 40%- and 50%-fixed threshold levels and compared with the ones from MR. One reason might be the SUV heterogeneity within the tumor, leading to erroneous volume estimates when fixed-threshold levels are used (either 40% or 50% of T_{max}). These T_{max} cutoff values often underestimate the MR-derived tumor volume, as they only represent the most active areas of the tumor and not the metabolic activity over the entire tumor volume. On the other hand, the proposed algorithm, applies an individual variable to reach the cutoff value: the tumor/liver uptake ratio. By doing this, we showed that the volumes obtained by the PET images are more reliable and well correlated to those obtained by using fixed arbitrary threshold levels.

Tumor volumes were not estimated from the pathological tissue samples, and thus, we were unable to study the degree of variability, between metabolic and anatomic tumor volumes in vivo. This variability could be explained in part by the inherent limitations of PET, and/or by the biological properties of the tumor itself.

It is important to point out that the measurements derived from PET data using this method suffer from certain shortcomings: the main problem is related to limiting the effect of partial volume on the lesion-to-liver ratios. These ratios are typically higher for low- and high-contrast objects.\(^{2}\) There were no histological measurements for comparison, which are considered to be the gold standard.\(^{3}\) Although not observed in the present study, another limitation could be related to the evaluation of lesions that are smaller than the spatial resolution of the PET system. Partial volume effect can also contribute, in part, to the less-than-perfect correlation on this scenario.\(^{4}\) We used MR-volume as our reference for comparisons. However, MR has reasonable resolution; we assumed
errors that might have interfered in our correlations.\textsuperscript{5} PET–MR image were not acquired using flat table; however, it may not have impacted negatively on the GTV measurements.

Conclusion

Our study results provide preliminary evidence that metabolic tumor volume delineation is feasible using computer-generated measurements in \textsuperscript{18}F-FDG PET images. Generation of \textsuperscript{18}F-FDG PET based tumor volumes is affected by the choice of threshold level used. Metabolic tumor bulk calculated using the pixel with the mean value plus 2-standard deviations of the liver intensity (GTV\textsubscript{2SD}) correlates better with the MR-derived tumor volumes than fixed threshold-based algorithms. The method is a simple and clinically applicable approach to generate PET-derived GTV for radiation therapy planning of cervical cancer.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgment

The authors thank Dr. Victor H. Gerbaudo (Divisions of Nuclear Medicine and Molecular Imaging, Brigham and Women’s Hospital, Harvard Medical School) for his tremendous editorial support and suggestions on this project.

References