Segmentation and characterization of tumors in 18F-FDG PET-CT for outcome prediction in cervical cancer radio-chemotherapy

Geoffrey Roman-Jimenez[1,2], Julie Leseur[2,3], Anne Devillers[1,2,3], Juan David Ospina[1,2], Guillaume Louvel[2,3], Pascal Haigron[1,2], Antoine Simon[1,2], Renaud de Crevoisier[1,2,3], and Oscar Acosta[1,2]

[1] INSERM, U 1099, Rennes, F-35000, France
[2] Université de Rennes 1, LTSI, F-35000, France
Geoffrey.Roman-Jimenez@univ.rennes1.fr
[3] Département de Radiothérapie, Centre Eugène Marquis, Rennes, F-35000, France

Abstract. Cervical cancer is one of the most common cancer to affect women worldwide. Despite the efficiency of radiotherapy treatment, some patients present recurrency. Early unfavorable outcomes prediction could help oncologist to adapt the treatment. Several studies suggest that tumor characteristics visible with ^{18}F-FDG PET imaging before and during the treatment could be used to predict post-treatment recurrency. We present a framework for segmentation and characterization of metabolic tumor activity aimed at exploring the predictive value of pre-treatment and per-treatment ^{18}F-FDG PET images. Thirty-five patients with locally advanced cervix cancer treated by chemoradiotherapy were considered in our study. For each patient, a coregistered PET/CT scan was acquired before and during the treatment and was segmented and characterized with our semi-automated framework. A segmentation process was applied on the baseline acquisition in order to find the metabolic tumor region (MTR). This MTR was propagated to the follow-up acquisition using a rigid registration step. For every patient, 40 features from the two MTRs were extracted to characterize the tumor changes between the two observation points. We identified explanatory characteristics by exploring the threshold which minimizes the p-value computed from the Kaplan-Meier free-disease survival curves. Seven features were identified as potentially correlated with cancer recurrency (p-value<0.05). Results suggest that our method can compute early meaningful features that are related with tumor recurrence.

Keywords:PET/CT, cervical cancer, radiotherapy, tumor recurrence prediction, image characterization

1 Introduction

Nowadays, cervical cancer is the third most common cancer and the fourth cause of cancer death in females worlwide with 529,800 cases diagnosed and 275,000 cancer deaths in 2008 [1]. The standard treatment of locally advanced cervical cancer is based on chemoradiotherapy and brachytherapy, inducing a risk of acute and late irreversible toxicity[2]. Because recurrence in cervical cancer significantly increase the risk of death [3], prediction of such events is crucial. Medical imaging can provide different

markers not only to plan the therapy but also to help oncologist to adapt the incourse treatment thereby increasing the chance of patients survival. During the last decade, severals studies have shown that coregistered PET/CT improve the diagnostic accuracy in several cancers [4][5]. As opposed to anatomic imaging such as CT or MRI, PET provides metabolic information of the tumor. Since metabolic changes of the tumor precede the morphologic modifications, PET-based studies could provide early prediction outcomes. Several studies have demonstrated that metabolic tumor changes occuring between PET scans allow to quantitatively assess tumor response [6][7]. Further, some studies have shown the predictive value of the informations extracted in PET [8][9].

In this study, we present a semi-automated framework for segmentation and characterization of metabolic tumor activity in cervical cancer from pre and per-treatment PET/CT acquisitions. We demonstrate the utility of our framework in predicting recurrency on a cohort of 35 patients treated for locally advanced cervix cancer. In the proposed framework the pre-treatment PET/CT fused information is exploited to first isolate the tumoral region from the rest of the image, then several characteristics are extracted from the PET at two time points during the treatment to find features that may explain recurrency. The identification of the recurrency correlated features was performed using Kaplan-Meier survival curves and the log-rank test.

2 Materials & Methods

2.1 Data and Clinical Protocol

Thirty-five patients (median age 52.44 years [32.15 - 84.62]) with locally advanced cervix cancer treated at the Centre Eugene Marquis (CEM), Rennes, France, were considered. The patients were treated with external beam radiation therapy (EBRT) with concurrent chemotherapy (CDDPcc) followed by brachytherapy (BT). With median follow-up of 29.21 months [7.44 - 52.64], eight patients developed tumor recurrence and three patients died. As shown in Fig. 1, each patient underwent three ^{18}F-FDG PET/CT scans performed with a DISCOVER ST scanner each observing the same time between injection and acquisition. In our study, we will only consider the pre and per-treatment acquisitions. The two pre-treatment images are denoted PET1 and CT1. At 40 Gy of the EBRT, the acquisitions were denoted PET2 and CT2. In order to have the images comparables, each PET image was converted into standardized uptake value (SUV) which is a standardized decay-corrected value of ^{18}F-FDG activity per unit volume of body weigth (MBq/kg)[11].

2.2 Segmentation & Characterization of the metabolic tumor region

The framework of the proposed method is illustrated in Fig. 2. The first step aims to isolate the tumoral area in PET1, avoiding the bladder hyperfixation. Because the tumor uptake in PET2 is less visible in responder patients due to the treatment, a rigid registration is carried out to map the region found in PET1 on PET2 thereby accessing to the same tumoral region. In step three, the two regions found in the previous stages are used to characterize the tumor before and during the treatment through features computation.

Fig. 1: Acquisitions during the treatment protocol: A baseline exam, PET/CT1, is acquired before the treatment. A second acquisition, PET/CT2, is performed at 40 Gy of the chemoradiotherapy. The last acquisition, PET/CT3, is performed 4-6 weeks after the brachytherapy

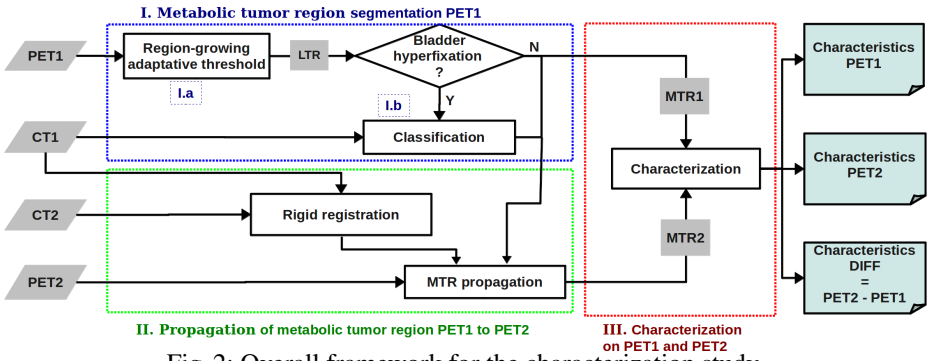

Fig. 2: Overall framework for the characterization study.

I. Metabolic tumor region segmentation in PET1: The metabolic tumor region (MTR) is denoted as the region where the acquired ^{18}F-FDG uptake corresponds to the tumor activity. The purpose of this step is to determine the metabolic tumor region in PET1 (MTR1). We considered here that the ^{18}F-FDG hyperfixation of the tumor before the treatment is topologically compact (one connected object). In that case, a region-growing threshold allows to select the voxels belonging to the MTR1. Region-growing was used to limited the inclusion of adjacent intense structures, such as the bladder, lymph nodes, and bowel [20].

I.a - *Region-growing adaptative threshold (RGAT)*: Let's denote likely tumoral region (LTR) the region wherein the ^{18}F-FDG hyperfixation is high enough to be considered as tumor metabolism. In order to extract the LTR, an adaptative threshold as the one proposed by Daisne [10] was implemented. This method aimed to adapt the threshold needed to segment the LTR according to the signal-to-background (S/B) measured in the image. The relationship between S/B and the threshold required was set up through a physical phantom study. Thus, because of the compactness hypothesis, the threshold was applied using a region-growing in order to keep a compact LTR.

I.b - *Fusion and gaussian mixture based classification (FGMC)*: Due to natural filling and emptying of the bladder, ^{18}F-FDG can be fixed in an extremely variable way leading to part of the bladder having intensities comparable to tumoral metabolism. It was necessary to visually determine if voxels in the bladder were selected in the LTR. In such case, a second step of classification was required to separate bladder and tumor. Firstly, the CT images were downsampled to the PET resolution so that each voxel in

PET correspond to only one voxel in CT. A voxel seed was selected as being likely within the tumor. From each voxel selected in the LTR, three features were considered: the standard uptake value on PET (SUV), the Hounsfield unity value on CT (HU) and a tumor membership probability (TMP). The group of voxels in LTR was thus denoted as $X = \{x_i | i \in I\}$. With x_i, a three feature vector $[SUV\ HU\ TMP]$ at the voxel i and I, the 3D coordinates in the image. We computed the TMP as:

$$TMP(x_i) = 1 - \frac{d(x_i, s)}{\max\limits_i d(x_i, s)}, \tag{1}$$

where $d(x_i, s)$ is the Euclidean distance of the voxel x_i from the seed s, and $\max\limits_i d(x_i, s)$ denote the Euclidean distance of the furthest voxel in the LTR from the seed. It was assumed that the further the voxel is from the reference the lower is the probability to belong to the tumor. SUV and HU were normalized between 0 and 1. The three normalized features were then projected in the tri-parametric space represented in Fig. 3. Voxels belonging to the tumor were identified by fitting a gaussian mixture model (GMM). In general terms, the GMM expresses that the distribution of points in this tri-parametric space is a sum of gaussian functions. Two clusters were considered: a cluster T representing voxels in the MTR and a cluster B for voxels in the bladder(eq 2).

$$g(x_i, \Theta) = p_T.f_T(x_i, \mu_T, \Sigma_T) + p_B.f_B(x_i, \mu_B, \Sigma_B), \tag{2}$$

where g denotes the mixture density of each LTR-voxel x_i. The constants p_T and p_B are the mixing proportions of the two gaussian distributions f_T and f_B respectively characterized by the means μ_T and μ_B and variance matrix Σ_T and Σ_B. Θ represents the model parameters of the gaussians mix $\Theta = [p_T, p_B, \mu_T, \mu_B, \Sigma_T, \Sigma_B]$ need to be estimated.

An expectation-maximization algorithm (EM) was used to calculate the maximum likehood estimates of Θ. Given these estimates, each voxel was assigned a label stating if it belongs to the cluster T or B. Fig. 3 gives an example of clustering in the classification space.

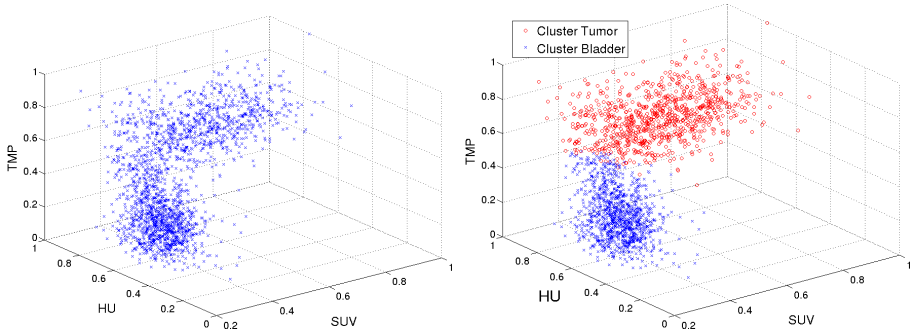

Fig. 3: Fusion and gaussian mixture based classification. (left) Tri-parametric classification space : SUV value, Hounsfield value and Tumor Membership Probability. (right) resulting EM estimation.

The largest connected component in the cluster "Tumor" was selected to be the MTR. Fig. 4 summarizes an example of metabolic tumor region segmentation in the pre-treatment PET.

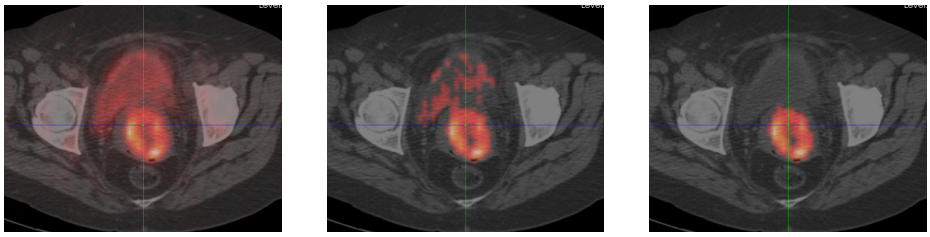

Fig. 4: Example of metabolic tumor region segmentation in PET1. (left) PET/CT1 with a high bladder uptake. (middle) LTR after RGAT. (right) MTR1 after the FGMC.

II. MTR1 to MTR2 registration: This step aims to propagate the MTR extracted in the PET1 to PET2. The propagated MTR was denoted MTR2. Since PET are coregistered with CT, these CT images were used to determine the transformation between PET images. A two-step registration using a block-matching algorithm [12] was adopted. Firstly, CT1 to CT2 were registered in order to match the whole body. Secondly, a rigid registration was applied locally to align the cervix. In practice, this is obtained by registering a VOI around the cervix. Note that local rigid registration was assumed because of the restricted VOI considered. This allows to keep the same shape between MTR in PET1 and PET2 and thus keep the same tumor area.

III. Characterization: Fourteen features were extracted from the two PET MTR to characterize patient's tumor before and during the treatment. In addition to SUVmax, metabolic tumor volume (MTV) and total lesion glycolysis (TLG) often reported in literature [13][14][8], we explored intensity and texture features[15][16]. The features are summarized in tables 1 and 2.

Note that, in PET1, the total activity by voxel volume (TAVV), is equivalent to the total lesion glycolysis (TLG) calculated as $SUV_{mean} \times MTV$ when MTV is well defined. Since we estimated the MTR2 by registration, the TAVV can not be denoted as TLG because MTV in PET2 is roughly delineated. However, for a good responder, it was expected that if the activity outside the true MTV was neglected, the TAVV will give a good approximation of the TLG.

Texture features, originally proposed by Haralick[18], are based on the co-occurence matrix Φ, reflecting spatial grey-level dependencies. The intensity range into the MTR was quantified with 16 bins. The texture features calculated shown in Tab. 2 were computed for every direction covering the 26-connected neighborhood and were averaged to keep a restricted number of characteristics. Finally, to characterize the changes between PET1 and PET2, the difference (DIFF) was calculated between features evaluated in MTR1 and MTR2.

2.3 Tumor recurrence prediction

A total of 40 features were extracted from 35 patients. Fourteen features were extracted from PET1, 13 features from PET2 and 13 from DIFF. The Kaplan-Meier method was

Features	Formula	Definition
SUVmax	$max(X)$	Maximal uptake in the MTR
SUVpeak	$\frac{1}{K}\sum_{k}^{K} x_k, k \in max_{neighbor}$	SUVmax averaged by its 26 neighbors
SUVmean	$\mu = \mathbb{E}[X]$	Average of SUV in the MTR
SUVvariance	$\Sigma = \mathbb{E}[(X-\mu)^2]$	Variance of SUV in the MTR
SUVskewness	$\gamma_1 = \mathbb{E}\left[\left(\frac{X-\mu}{\Sigma}\right)^3\right]$	Asymmetry measure of the MTR activity distribution
SUVkurtosis	$\gamma_2 = \mathbb{E}\left[\left(\frac{X-\mu}{\Sigma}\right)^4\right]$	Peakedness measure of the MTR activity distribution
MTV	$Volume_{voxel} \times Nb_{Voxel}$	Metabolic Tumor Volume of the MTR
TAVV	$Volume_{voxel} \times \sum_{n}^{N} x_n$	Total Activity by Voxel Volume of the MTR

Table 1: Evaluated intensity features

Features	Formula
Energy	$\sum_{i,j} \Phi(i,j)^2$
Entropy	$\sum_{i,j} \Phi(i,j) \times log_2(\Phi(i,j))$
Inertia	$\sum_{i,j} (i-j)^2 \times \Phi(i,j)$
InverseDifferentMoment (IDM)	$\sum_{i,j} (i-j)^2 \times \Phi(i,j)$
Cluster Shade	$\sum_{i,j} (i+j-\mu_j-\mu_i)^3 \times \Phi(i,j)$
Cluster Prominence	$\sum_{i,j} (i-\mu_i+j-\mu_j)^4 \times \Phi(i,j)$

Table 2: evaluated texture features

used to evaluate the disease-free survival curves of the group splitted by a threshold. The feature assessment was perfomed following this method: for each feature, a threshold was found in order to minimize the p-value of the log-rank test from the comparison of the two survival curves generated when splitting the group using this threshold. The level of significance for the log-rank test was fixed to p-value<0.05.

3 Results

3.1 FGMC evaluation

Among the 35 PET1 segmented in step I.a, 12 visually presented high bladder uptake and were clustered in step I.b. In order to evaluate the deletion of bladder voxels performed by the FGMC, we used expert manual cervix and bladder CT-segmentation available for six patients. After I.b, the number of voxels remained in the cervix mask was considered as true positive (TP) whereas false positive (FP) was the number of those in the bladder mask. For the six clusterization evaluated, the averaged results

were 0.80 ± 0.17 sensitivity, 0.97 ± 0.08 specificity and 0.85 ± 0.11 accuracy. We can observe a lower sensitivity due to the worst patient which presented a classification with 0.53 sensitivity, 1.00 specificity and 0.63 accuracy. In this case, the patient presented a small tumor and as the number of voxels in the cervix mask in LTR was limited, any suppressed voxel during the FGMC had an important impact on the resulting sensitivity.

3.2 Explanatory features identification

Tab. 3 summarizes the significant features extracted in step III. Results show that the explanatory features found are mainly based on SUV intensity. The four characteristics found from PET1 might suggest that tumor recurrence could be predicted before the treatment. Moreover, the TAVV evolution between the two exams(DIFF) might be a recurrency correlated feature that could be used by oncologists to adapt the ongoing treatment. An example of a Kaplan-Meier curve with the TAVV feature in PET1 and DIFF is diplayed in Fig. 5.

exam	feature	p-value	threshold
PET1	SUVmax	0.0335	10.52
	SUVpeak	0.0478	7.27
	SUVmean	0.0335	6.30
	TAVV or TLG	0.0114	155.62
PET2	SUVmax	0.0307	6.26
	IDM	0.0419	0.16
DIFF (PET1-PET2)	TAVV	0.0171	-266.15

Table 3: Assessed features by the Kaplan-Meier method with p-value<0.05 and the corresponding threshold.

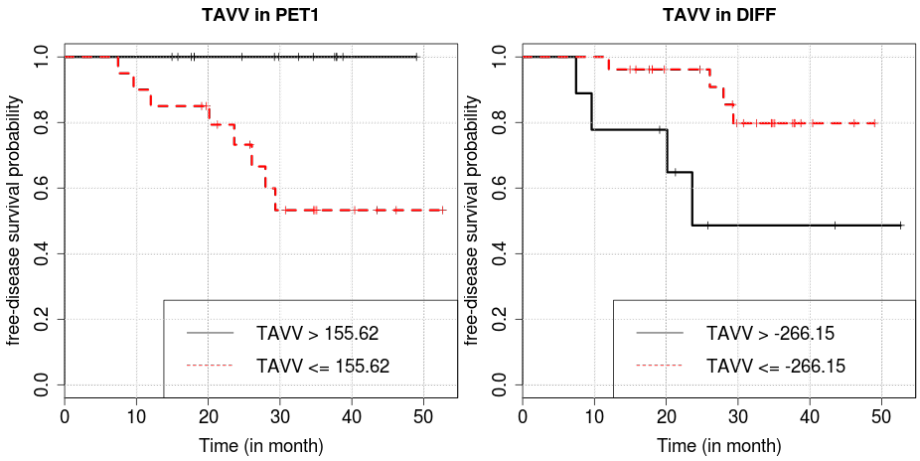

Fig. 5: Example of feature evaluation with Kaplan-Meier method. (left) TAVV in PET1 (p-value=0.0114). (right) TAVV in DIFF (p-value=0.0171).

4 Discussion & Conclusion

In this study, we presented a whole framework to characterize cervical cancer tumor from ^{18}F-FDG PET imaging and to predict its response to radiotherapy. One of the challenges was to isolate the tumor region from the bladder in PET1. With the use of CT information, we proposed a semi-automated method to segment and characterize PET imaging based on a RGAT followed by a FGMC approach. Since metabolic informations can be hidden in lower uptake than the applied threshold, an intensity-based segmentation can lead to an underestimation of the true LTR extracted. Also, an adaptative threshold suffers from poor reproducibility[19] and a more evolved segmentation method could enhance this. The FGMC has demonstrated the ability to classifying voxels between tumor and bladder with good specificity, sensitivity and accuracy (respectively 0.80, 0.97 and 0.85). Nonetheless, for one given patient, we visually decided if the classification step was necessary. An integration of the FGMC into the segmentation step is in progress in order to reduce user dependency. The characterization step was perfomed by evaluating intensity and texture features on MTR1 and MTR2. Also, we characterized the tumor changes by computing the difference between the features in MTR1 and MTR2. In future work, shape metrics could be extracted to describe the tumor thereby providing complimentary information. We identified seven features that might be correlated with cervical cancer recurrency by assessing free-disease survival curves. This is an exploratory study so, in future work, the predictive capabilities of the found thresholds will be assessed through a larger cohort of patients.

References

1. Jemal *et al.*, Global cancer statistics. CA: A Cancer Journal for Clinicians **61** (2011) 69–90
2. Kirwan *et al.*, A systematic review of acute and late toxicity of concomitant chemoradiation for cervical cancer. Radiother Oncol **68** (2003) 217–226
3. Sommers *et al.*, Outcome of recurrent cervical carcinoma following definitive irradiation. Gynecologic Oncology **35** (1989) 150 – 155
4. Bar-Shalom *et al.*, Clinical performance of pet/ct in evaluation of cancer: additional value for diagnostic imaging and patient management. J Nucl Med **44** (2003) 1200–1209
5. Dhingra *et al.*, Impact of fdg-pet and -pet/ct imaging in the clinical decision-making of ovarian carcinoma: an evidence-based approach. Womens Health (Lond Engl) **8** (2012) 191–203
6. Necib *et al.*, Detection and characterization of tumor changes in 18f-fdg pet patient monitoring using parametric imaging. J Nucl Med **52** (2011) 354–361
7. David *et al.*, Multi-observation pet image analysis for patient follow-up quantitation and therapy assessment. Phys Med Biol **56** (2011) 5771–5788
8. Leseur *et al.*, Metabolic monitoring by 18f-fdg pet during radio-chemotherapy for locally advanced cervical cancer: Predicting outcome. International Journal of Radiation Oncology*Biology*Physics **81** (2011) S47–S48 53rd Annual Meeting ASTRO's.
9. Chung *et al.*, Clinical impact of integrated pet/ct on the management of suspected cervical cancer recurrence. Gynecol Oncol **104** (2007) 529–534
10. Daisne *et al.*, Tri-dimensional automatic segmentation of pet volumes based on measured source-to-background ratios: influence of reconstruction algorithms. Radiother Oncol **69** (2003) 247–250
11. Strauss *et al.*, The applications of pet in clinical oncology. J Nucl Med **32** (1991) 623–48; discussion 649–50

12. Ourselin *et al.*, Reconstructing a 3d structure from serial histological sections. Image and Vision Computing **19** (2001) 25 – 31
13. Wahl *et al.*, From recist to percist: Evolving considerations for pet response criteria in solid tumors. J Nucl Med **50 Suppl 1** (2009) 122S–150S
14. Hatt *et al.*, Impact of partial-volume effect correction on the predictive and prognostic value of baseline 18f-fdg pet images in esophageal cancer. J Nucl Med **53** (2012) 12–20
15. Yu *et al.*, Coregistered fdg pet/ct-based textural characterization of head and neck cancer for radiation treatment planning. IEEE Trans Med Imaging **28** (2009) 374–383
16. Naqa *et al.*, Chaudhari, S., Yang, D., Schmitt, M., Laforest, R., Thorstad, W., Deasy, J.O.: Exploring feature-based approaches in pet images for predicting cancer treatment outcomes. Pattern Recognit **42** (2009) 1162–1171
17. Hatt *et al.*, Baseline ^{18}f-fdg pet image-derived parameters for therapy response prediction in oesophageal cancer. Eur J Nucl Med Mol Imaging **38** (2011) 1595–1606
18. Haralick *et al.*, Textural features for image classification. **3** (1973) 610–621
19. Hatt *et al.*, metabolically active volumes automatic delineation methodologies in pet imaging: review and perspectives. Cancer Radiother **16** (2012) 70–81; quiz 82, 84
20. Miller *et al.*, Measurement of tumor volume by pet to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy. International Journal of Radiation Oncology*Biology*Physics **53** (2002) 353 – 359