<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Delineation of Biological Tumor Volume from Positron Emission Tomography Images in Nasopharyngeal Carcinoma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Wong, KW; Kwong, DLW; Khong, PL; Lee, VHF; Ng, SCY; Law, MWM</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Journal of Biomedical Science and Engineering, 2014, v. 7 n. 11, p. 857-865</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2014</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/202758">http://hdl.handle.net/10722/202758</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Delineation of Biological Tumor Volume from Positron Emission Tomography Images in Nasopharyngeal Carcinoma

Ka-Wai Wong1*, Dora Lai-Wan Kwong2*, Pek-Lan Khong3, Victor Ho-Fun Lee2, Sherry Chor-Yi Ng4, Martin Wai-Ming Law5

1The Department of Clinical Oncology, Queen Mary Hospital, Hong Kong, China
2The Department of Clinical Oncology, The University of Hong Kong, Hong Kong, China
3The Department of Diagnostic Radiology, The University of Hong Kong, Hong Kong, China
4The Department of Clinical Oncology, Queen Elizabeth Hospital, Hong Kong, China
5The Department of Diagnostic Radiology, Queen Mary Hospital, Hong Kong, China

Email: *garykwwong@yahoo.com, *dlwk Wong@hkucc.hku.hk

Received 22 June 2014; revised 7 August 2014; accepted 22 August 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Abstract

Objective: To develop an auto-segmentation method for delineating the biological tumor volume of nasopharyngeal carcinoma from positron emission tomography images. Methods: A phantom consisting of a water tank with fixed background fluorodeoxyglucose [18F-FDG] activity and spheres with diameters ranging from 2.1 to 5 cm with varying activities of FDG were used to simulate tumors of different sizes and FDG uptake. The phantom was scanned with a PET/CT scan at different sphere to background intensity ratios. An optimum fixed percentage threshold (FT) approach and a signal-to-background ratio (SBR) approach were developed to estimate the true size of the spheres from the PET images. Both approaches were further evaluated in patient images for validation. Twenty-two patients with NPC from stage T1 to T4 were included. The PET based biological tumor volumes (BTV) were delineated with both FT (BTVFT) and SBR (BTVSBR) approaches and compared with the gross tumor volume localized from MRI (GTVMR). The mean volumes of BTVFT and BTVSBR were compared and the degree of overlap between GTVMR and both BTVs was evaluated. Paired t-tests were used for statistical analysis. Results: The optimal FT value was 36.5% of maximal intensity, and SBR approach was represented by an inverse linear regression model. The estimated volume of spheres segmented by both approaches shows no significant difference from the true volume of spheres (p > 0.05), but the average absolute errors were smaller from SBR approach than FT approach (p = 0.008). GTVMR was larger than both BTVFT (p = 0.003) and BTVSBR (p < 0.009). The overlapping volume of BTVSBR with GTVMR is significantly larger than with BTVFT.

*Corresponding authors.

Conclusions: SBR approach is more feasible than FT approach to localize the BTV of NPC from FDG-PET image, and BTV might help to modify the extent of GTV<sub>MR</sub> for radiotherapy planning purpose.

**Keywords**

Segmentation Approach, SBR, Fixed Threshold, NPC, PET

---

**1. Introduction**

Intensity-modulated radiotherapy (IMRT) is the state of the art treatment for head and neck cancers. The critical first step to implement IMRT requires an accurate delineation of the gross tumor volume (GTV). Anatomical information from computed tomography (CT) and magnetic resonance imaging (MRI) is widely used for tumor localization. The introduction of 18F-fluorodeoxyglucose (FDG) positron emission tomography (PET) illustrates the biological activity of the tumor. This will add another dimension showing the biological activity within the anatomical extent of infiltration of tumor as defined by computed tomography (CT) and magnetic resonance imaging (MRI). However, the resolution of PET scan is poor and the exact boundaries of the tumor on PET images are difficult to define.

In the past decade, different segmentation methods such as visual interpretation, thresholding methods, variational approaches, and other software-based approaches were developed to localize the biological tumor volume (BTV) from FDG-PET [1], and many studies had illustrated that the incorporation of this biological information in radiotherapy planning procedure would often lead to a change in the local extent of anatomical based GTV [2]-[6]. However, there is still no consensus on which segmentation method would best localize the primary tumour [1] [7]. Among all, fixed threshold approach and adaptative threshold approach are most commonly applied in most of the tumours in the head and neck region, but as yet, it has not been applied in nasopharyngeal carcinoma (NPC).

In this study, two phantom-based segmentation approaches including the optimum fixed percentage threshold (FT) approach and the signal-to-background ratio (SBR) approach were developed. The accuracy and feasibility of both approaches were evaluated based on the phantom studies and further validated in the patients’ images. Unlike other head and neck cancers, we cannot have actual measurement of the tumor since nasopharyngeal carcinoma is not treated surgically with resection. MRI shows the most accurate anatomical extent of the primary tumor and the gross tumor volume in nasopharynx as delineated on MRI (GTV<sub>MR</sub>) was used as the standard for comparison. The BTVs as delineated by fixed threshold (BTV<sub>FT</sub>) and SBR (BTV<sub>SBR</sub>) were compared with GTV<sub>MR</sub> for their difference in volumes and degree of overlap in anatomical position.

**2. Methods**

**2.1. Phantom Study**

Two in-house produced phantoms (Phantom 1 and Phantom 2), consisting of 2 water tanks with width × length × height = 23 × 20 × 23 cm<sup>3</sup>, together with 7 difference sizes of glass spheres (5.2, 9.3, 16.0, 23.7, 37.4, 52.6 and 68.1 cm<sup>3</sup>) with internal diameter ranged from 2.1 cm to 5 cm were developed. Each sphere was made with a single-end opening, and with wall thickness of 2.5 - 3.0 mm. All the spheres were sealed with 18F-FDG solution at an activity concentration of 108 kBq/ml (±5%). They were immersed in the phantom pool which contained 18F-FDG solution at a constant activity of 3.145 kBq/ml (±5%) ([Figure 1](#)). The spheres simulated tumor with different sizes and FDG uptake. While the activity of 18F-FDG inside the spheres decayed, PET/CT images were acquired at 9 different sets of sphere to background intensity ratios (33, 24, 16, 12, 8.5, 6, 5, 4 and 2). A combined PET/CT helical mode scanner (Discovery VCT; GE Medical Systems) was used to acquire the images. PET/CT images were reconstructed at 0.625 mm intervals. The reconstructed field of view (FOV) is 35 cm, with 512 × 512 matrix size. The images were transferred to MIMcontouring Advanced—MIM™ 5 (MIM Software Inc.) for analysis.
2.2. Fixed Threshold Approach

Threshold of 38% maximal standardized uptake value (SUV) was first selected for estimation of the volume of sphere as derived from the PET images. This value was iteratively tested for less than 6 times with step size equal 0.5% until the resultant volume of sphere has the least square difference with the true volume of sphere. The true volume of each sphere was calculated from the volume as outlined from CT images. After all, the measured values from all spheres were averaged to get the optimum fixed percentage threshold value.

2.3. Signal to Background Approach

A varying threshold value which relies on the relationship between the percentage threshold values and the measured sphere to background intensity ratios were derived to aid lesion segmentation. In this model, the percentage threshold value was determined as a percentage of the maximal activity within the spheres located from PET images. The maximum radioactivity of the sphere was the hottest area (16 mm diameter) drawn on the transverse slice of the PET image. The average background activity was measured from the 18F-FGD activity in each set of phantom pool. Four circles with 20 mm diameter were drawn across each slice with 1 cm apart from the sphere, and without overlapping with each other. The average background activity was then collected from a total of 112 circles drawn at each phantom pool. When both the maximum activity and the mean background activity were collected, the intensity ratios for each couple of sphere to background were obtained. The measured SBR and the resultant threshold are fitted by a regression function to yield the best regression parameters for the equation.

After both approaches were developed, they were first used to re-segment the spheres. The estimated volumes derived from both approaches were validated with the true volume of spheres. Paired t-test using SPSS® (version 17.0) was applied for statistical analysis.

2.4. Validation in Patients’ PET Images

After the phantom studies, both approaches were validated with patients’ images. A cohort of 22 nasopharyngeal carcinoma patients, 16 male and 6 female, who had PET/CT and MRI scan performed in treatment position for IMRT planning during 2008-2009 were included. The group included 8 patients each with stage T1 and T3, and 3 patients each with T2 and T4 respectively. All patients had their planning PET/CT, and MRI done not more than one week apart. The scanning parameters for PET/CT were 140 kV, auto mA (150 - 350 mA), with slice interval at 2.5 mm. Reconstructed FOV was 35 cm, with 512 × 512 matrix sizes. For PET scan, the FOV was 70 cm, with 128 × 128 matrix sizes. The slice was reconstructed at 3.25 mm interval. The coverage of both PET and CT scans begins from 1 cm above the vertex to 1/3 of upper thigh. MRI was performed using 3T MRI (Philips Achieva 3.0 system). Images were acquired from vertex to clavicle region. FOV was 50 cm, with 560 × 560 matrix sizes. The slice thickness was reconstructed at 3 mm interval.

Patient’s position during PET/CT and MRI were the same. Each patient was lying on a headrest, and cast was used to immobilize the patient only in PET/CT. All the images were fused manually based on some prominent
anatomical position, such as the divergence of basilar artery, lens and dens etc., before radiotherapy planning. In this study, a clinical oncologist was asked to localize the GTV\textsubscript{MR} without looking at the tumour from the PET/CT. Meanwhile, the BTVs were outlined by an experienced radiation therapist using the FT and SBR segmentation approaches as derived from the phantom study. The mean volume of GTV\textsubscript{MR} and both BTVs, BTV\textsubscript{FT} and BTV\textsubscript{SBR}, as well as the overlapping and non-overlapping fraction between GTV\textsubscript{MR} and BTVs were compared as shown in Figure 2. This includes the fraction of GTV\textsubscript{MR} enclosed within the BTV relative to GTV\textsubscript{MR} (OF\textsubscript{MR}), the fraction of BTV enclosed within the GTV\textsubscript{MR} relative to BTV (OF\textsubscript{PET}), the fraction of GTV\textsubscript{MR} excluded outside the BTV relative to GTV\textsubscript{MR} (1-OF\textsubscript{MR}), and also the fraction of BTV excluded outside the GTV\textsubscript{MR} relative to BTV (1-OF\textsubscript{PET}). Paired t-test and confidence interval was used for statistical analysis.

3. Results

55 out of 63 couples of sphere to background intensity ratio were measured. When the sphere to background intensity ratio was less than 4, the lesions could not be differentiated from its background and the percentage threshold value could not be measured. Thus, 8 measurements could not be obtained and excluded from study. From the FT approach, the trend of the percentage threshold values appeared quite constant (Figure 3) between the spherical volume from 5.2 cm\textsuperscript{3} to 68.1 cm\textsuperscript{3} and within the sphere to background intensity ratio from 4 to 33. The threshold values varied from 31.8% to 41.5% (SD = 2.4%) with an average value equal to 36.5% of maximal intensity uptake.

For SBR approach, an inverse function was found between the threshold values and the SBR. The regression model was defined as

\[ \text{Threshold } (\%) = \frac{a}{\text{SBR}} + b, \]

with the regression parameters \( a = 31.79 \) and \( b = 36.54 \) \((R^2 = 0.735)\) best fitted the experimental results (Figure 4). The plots were valid with spherical volume ranges from 5.2 cm\textsuperscript{3} to 68.1 cm\textsuperscript{3}, and with the measured sphere to background intensity ratio from 2.6 to 31.3.

From the phantom’s images analysis, the estimated BTV derived from FT approach (BTV\textsubscript{FT}) and SBR approach (BTV\textsubscript{SBR}) were of no significant difference from the true volume of spheres (FT: \( p = 0.088; \) SBR: \( p = 0.714 \)). The true volume of spheres, which are the actual volume contain in each sphere, has less than 1% deviation from the volume as outlined from CT images. However, the absolute differences between the estimated volume and the true volume of spheres derived from SBR approach was significantly less than that derived by FT approach \( (p = 0.008).\) A larger absolute error was commonly found in the smallest size (5.2 cm\textsuperscript{3}) of sphere (Figure 5), the difference was 13.3% and 8.5% respectively for FT and SBR approach.

From the patients’ images analysis, 18 out of 22 patients were available for validation. Four patients with stage T1 disease were excluded because both segmentation approaches failed to differentiate the edge of tumour from its background tissues. Their GTV\textsubscript{MR} ranged from 0.52 cm\textsuperscript{3} to 5.11 cm\textsuperscript{3} with an average volume equal to 1.71 cm\textsuperscript{3}. Two patients with T4 disease has the GTV\textsubscript{MR} 1.4 and 2.5 times greater than the largest size of our

<table>
<thead>
<tr>
<th>Item</th>
<th>Volume of the compartments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. OF\textsubscript{MR}</td>
<td>B / A</td>
</tr>
<tr>
<td>2. OF\textsubscript{PET}</td>
<td>B / C</td>
</tr>
<tr>
<td>3. 1 - OF\textsubscript{MR}</td>
<td>A - B / A</td>
</tr>
<tr>
<td>4. 1 - OF\textsubscript{PET}</td>
<td>C - B / C</td>
</tr>
</tbody>
</table>

Figure 2. A schematic diagram describing the volume analysis between the overlapping fraction and non-overlapping fraction of GTV\textsubscript{MR} and BTV.
Figure 3. A schematic diagram describing the volume analysis between the overlapping fraction and non-overlapping fraction of GTV_{MR} and BTV.

Figure 4. Relationship between the sphere to background ratio and the threshold activity to apply to adequately segment the spheres of volumes ranging from 5.2 to 68.1 cm$^3$.

glass sphere used, however, it was expected that threshold measurement is insensitive to large volume, thus they were also included in the study group. The average volume of GTV_{MR}, BTV_{FT} and BTV_{SBR} were 38.22 cm$^3$, 17.61 cm$^3$, and 21.23 cm$^3$ respectively. GTV_{MR} was significantly larger than BTV_{FT} ($p = 0.003$) and BTV_{SBR} ($p < 0.009$). As the average BTV_{SBR} was larger than BTV_{FT}, the average volume of overlapping between GTV_{MR} with BTV_{SBR} was also larger than with BTV_{FT}, they were 18.65 cm$^3$ and 16.1 cm$^3$, respectively ($p < 0.0005$).

The result of overlap analysis (Table 1) indicated that the percentage of volume overlapping between BTV and GTV_{MR} relative to GTV_{MR} increased significantly when the segmentation approach was changed from FT to SBR, the increment was from 42% to 52% ($p = 0.018$). Meanwhile, the percentage of BTV overlapping with GTV_{MR} only reduced slightly from 89% to 84%. On the other hand, there were almost half the volume of GTV_{MR} excluded outside the BTV, whereas, there were 11% - 16% of BTV excluded outside GTV_{MR}. 


Table 1. A summary of mean overlapping fraction of volume between GTV<sub>MR</sub> and BTV<sub>FT</sub> or BTV<sub>SBR</sub> with 95% confidence interval.

<table>
<thead>
<tr>
<th>Stage</th>
<th>Mean OF&lt;sub&gt;MR&lt;/sub&gt;</th>
<th>Mean OF&lt;sub&gt;PET&lt;/sub&gt;</th>
<th>Mean 1 - OF&lt;sub&gt;MR&lt;/sub&gt;</th>
<th>Mean 1 - OF&lt;sub&gt;PET&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>BTV&lt;sub&gt;FT&lt;/sub&gt; 95% CI</td>
<td>0.42 (0.36 - 0.48)</td>
<td>0.89 (0.83 - 0.95)</td>
<td>0.58 (0.523 - 0.64)</td>
</tr>
<tr>
<td></td>
<td>BTV&lt;sub&gt;SBR&lt;/sub&gt; 95% CI</td>
<td>0.52 (0.45 - 0.58)</td>
<td>0.84 (0.78 - 0.9)</td>
<td>0.48 (0.42 - 0.55)</td>
</tr>
</tbody>
</table>

Figure 5. The mean absolute relative difference of the estimated volume of spheres. The above T-bar is for (+) 95% CI. ● by FT approach; Δ by SBR approach.

4. Discussion

The results of volume analysis on phantom’s images reflected the SBR approach is more accurate than the FT approach. The SBR approach shows a significantly smaller average absolute error in volume estimation ($p = 0.008$). There is an inverse correlation between the average relative error in volume estimation versus both the size and signal to background intensity ratio. The results were accounted by the limitation of intrinsic spatial resolution of PET, the associated partial volume effect and also edge-blurring effect. Previous studies had indicated that partial volume effect often existed when the lesion was less than 5 cm$^3$ [8] [9], and the edge blurring effect was dominant when the background contrast level was less than 5 [10]. These errors may account for the failure of lesion segmentation when the sphere to background intensity ratio was below 4.

Previous researchers had stated that fixed threshold approach is easy to apply for delineation, but it should be validated or otherwise it might lead to significant errors in volume estimation [4] [11] [12]. Our study demonstrated that SBR approach is more accurate than a calibrated fixed percentage threshold for lesion segmentation because SBR approach is relatively more flexible to include the FDG avid voxels in the segmentation process. The potential limitation in this phantom study was mainly the use of glass sphere which may affect the threshold findings by shifting of gradient peak towards the interior of the sphere, thus causing the measured threshold values to be underestimated [13].

From patients’ images study, the mean GTV<sub>MR</sub> was larger than the mean BTV<sub>FT</sub> by 2.2 folds and BTV<sub>SBR</sub> by 1.8 folds. This was consistent with previous findings, in which the CT or MRI based GTV in most of the head and neck tumours were usually larger than the BTV by not more than 2 folds [2]-[5] [14] [15]. Danise et al. (2004) [2] reported that the manually localized GTV was 1.6 times larger than the BTV, but the BTV had the closest tumour size when compared with the standard pathological specimen. The authors concluded that BTV, which segmented by its institutional-based SBR model, was more accurate than anatomical based GTV to reflect the tumour extension of laryngeal cancer. Although pathological specimens of NPC were not available for analysis, our findings using MR detected GTV were similar to findings for other head and neck cancers. The almost double in size of GTV<sub>MR</sub> over BTV accounted for an average of 58% and 48% of GTV<sub>MR</sub> excluded from BTV<sub>FT</sub> and BTV<sub>SBR</sub> respectively. The difference in tumour extent between GTV<sub>MR</sub> and BTV<sub>FT</sub> or BTV<sub>SBR</sub> was clearly demonstrated in Figure 6 & Figure 7. In Figure 6, there is a tail-like structure of GTV<sub>MR</sub> outlined on the left side of nasopharynx, but no radioactive glucose uptake was reflected on the same area from the FDG-PET image. In Figure 7, the sphenoid and ethmoid sinus involvement outlined from MRI was obviously more than it was detected from PET/CT image.
Figure 6. Comparison of GTV\textsubscript{MR} (red), BTV\textsubscript{FT} (green) and BTV\textsubscript{SBR} (blue) for a stage T2 NPC. MRI and PET/CT was first registered, and then MR based GTV (left) was pasted to the fused PET image (right). Part of the GTV\textsubscript{MR} did not show any FDG uptake with a tail-like structure extended outside the BTVs.

Figure 7. Comparison of GTV\textsubscript{MR} (red), BTV\textsubscript{FT} (green) and BTV\textsubscript{SBR} (blue) for a stage T4 NPC. After MRI was registered with PET/CT, GTV\textsubscript{MR} was pasted to PET/CT. Manually localized GTV\textsubscript{MR} may over-estimated the ethmoid sinus involvement, while under-estimated the intracranial involvement on the right side.

On the other hand, our results showed that there was up to an average of 84% BTV\textsubscript{SBR} and 89% BTV\textsubscript{FT} mapped within GTV\textsubscript{MR}, while there were 11% - 16% of BTV excluded outside the GTV\textsubscript{MR}. This reflected that when the thresholding approaches, particularly the SBR approach, was used to localized BTV from FDG-PET, it does not only have a high concordance with GTV\textsubscript{MR}, but also detect a significant portion of tumour involvement that might be underestimated from anatomical based information. An example was shown in Figure 7. BTV showed a larger extent of right temporal lobe involvement than it was localized in MRI based GTV. Although, Ng et al. (2009) [12] concluded that PET/CT is less sensitive than MRI to access the local extent of NPC when 50% fixed threshold was used for BTV delineation. Our results suggested that different segmentation approaches would alter the tumour extent from FDG-PET image, and PET/CT has its specificity and sensitivity over MRI to detect the local tumour extent of NPC, provided that an appropriate segmentation approach was applied for the delineation of BTV.

There are 4 out of 8 cases of stage T1 disease excluded in this study. These cases were commonly found to have a poor FDG uptake in the primary lesion but with a relatively higher FDG uptake in the tissues nearby, such as in soft palate, tongue, and pharyngeal constrictor muscle. The measured tumour to background intensity ratio was ranged from 2.7 to 4.5. This low contrast situation limits both the FT and SBR approaches to differentiate the metabolically active tumour cells from their backgrounds, and this phenomenon were also reported in
some previous findings [16]-[18].

While there were no pathological specimen available to correlate with the image findings, it would be difficult to conclude whether GTV_MRI or BTV_SBR could more truly reflect the local tumour extent. It is not sure if the GTV_MRI that is not included in BTV were due to microscopic extension of tumor with lesser proliferative activity or if the MR changes could be due to tissues inflammation, edema or even observer-based variability. On the other hand, the portion of BTV extending outside GTV_MRI could be due to infections, inflammation, or physiologically increased uptake in nearby tissue structures such as tonsils, salivary glands and muscles or partial volume effect. These possibilities should be considered on individual basis when interpreting the images. Our findings suggested that the biological information from FDG-PET would help to refine anatomical based GTV, and the change in local tumour extent would consequently lead to a change in radiation beam arrangement or the management approach as well [19]-[21].

As a whole, SBR approach is more feasible than FT approach to outline BTV of NPC from FDG-PET. However, SBR approach has its own intrinsic limitations. The validation was limited to fixed size of spheres and homogenous activity, it does not take into account the irregular shape of tumour, the heterogeneity of tumour, and also the possibility of overlapping between the tumour and the surrounding normal tissues, thus causing under/over estimating the tumour extent [1] [2] [4] [13]. Secondly, this tool needs to be calibrated to specific institutional image acquisition and reconstruction settings. And last but not the least, it is not ideal for low SBR images [2] [5] [13]. The sample size in this study was small and further study with larger cohort of patients with different T stage of disease would be required for validation of the SBR approach.

5. Conclusion
This study has demonstrated that SBR approach works better than FT approach to delineate BTV of NPC from FDG-PET, and this added on biological information would help to improve the tumour localization by minimizing the chance of over/under estimation of anatomical based GTV in radiotherapy planning procedure.

Acknowledgements
Area of Excellence grant AoE/M-06/08 and General Research Grant HKU 768312M.

References


Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.