Radiobiological Modeling Based on \(^{18}\text{F-Fluorodeoxyglucose}\) Positron Emission Tomography Data for Esophageal Cancer

Mariana Guerrero*, Shan Tan\(^1\)\(^2\) and Wei Lu\(^1\)

\(^1\)Department of Radiation Oncology, University of Maryland School of Medicine, Baltimore, 21201, USA
\(^2\)Department of Intelligent Science and Technology, School of Automation, Huazhong University of Science and Technology, Wuhan, China

**Abstract**

**Background:** We investigated the relationship of standardized uptake values (SUVs) to radiobiological parameters, such as a control probability (TCP), to allow for quantitative prediction of tumor response based on SUVs from \(^{18}\text{F-fluorodeoxyglucose}\) \((^{18}\text{F-FDG})\) PET imaging before and after treatment for esophageal cancer.

**Methods:** We analyzed data from 20 esophageal cancer patients treated with chemoradiotherapy (CRT) followed by surgery. Tumor pathologic response to CRT was assessed in surgical specimens. Patients underwent \(^{18}\text{F-FDG}\) PET imaging before and after CRT. Rigid image registration was performed between both images. Because TCP in a heterogeneous tumor is a function of average cell survival, we modeled TCP as a function of \(<\text{SUV}_R>\), a possible surrogate for average cell survival \(<\text{SUV}_R> = <\text{SUV}_{\text{after}} - \text{SUV}_{\text{before}}>\). TCP was represented by a sigmoid function with two parameters: \(\text{SUV}_{\text{after}}\) the \(<\text{SUV}_R>\) at which TCP=0.5, and \(\gamma_1\), the slope of the curve at \(\text{SUV}_{\text{mean}}\).

**Results:** A TCP model as a function of SUV before and after treatment was developed for esophageal cancer patients. The maximum-likelihood estimate of \(\text{SUV}_{\text{max}}\) was 0.47 (90% CI, 0.30-0.61) and for \(\gamma_1\) was 1.62 (90% CI, 0.4-2). High initial SUV and larger metabolic response (larger \(<\text{SUV}_{\text{after}} - \text{SUV}_{\text{before}}>)\) were correlated, and this correlation was stronger among responders.

**Conclusions:** Our TCP model indicates that \(<\text{SUV}_{\text{after}} / \text{SUV}_{\text{before}}\>\) is a possible surrogate for cell survival in esophageal cancer patients. Although CIs are large as a result of the small patient sample, parameters for a TCP curve can be derived and an individualized TCP can be calculated for future patients. Initial SUV does not predict response, whereas a correlation is found between surrogates for initial tumor burden and cell kill during therapy.

**Keywords:** \(^{18}\text{F-FDG}\) PET/CT; Tumor response; Esophageal cancer; Quantitative imaging; Radiation therapy

**Introduction**

\(^{18}\text{F-fluorodeoxyglucose}\) \((^{18}\text{F-FDG})\) positron emission tomography (PET) is routinely used as a tool to diagnose and evaluate response in many cancer sites. A body of published evidence has documented correlation of standardized uptake value (SUV) with tumor clonogenic cell 60 density (or tumor cellularity) and tumor proliferation (e.g., Zhou et al. [1], Fischer et al. [2]). For more than 20 years, reports have specifically documented correlations between \(^{18}\text{F-FDG}\) uptake changes and response to therapy (e.g., Wahl et al. [3]). In esophageal cancer and gastroesophageal junction tumors, Omloo et al. [4] and Wu et al. [5] found mixed results in terms of SUV correlation with survival and/or pathological response (both for pretreatment SUV and changes in SUV before and after chemoradiation). Even when a correlation between SUV and response is established, in most cases no known quantitative relationship between SUVs and tumor characteristics and response patterns can be identified. \(^{18}\text{F-FDG}\) PET imaging, therefore, is typically used in a qualitative or semiquantitative manner. The purpose of this work is to investigate the quantitative relationship between mean patient SUVs and radiobiological parameters (such as cell survival and tumor control probability [TCP]) to facilitate quantitative prediction of tumor response based on SUVs from \(^{18}\text{F-FDG}\) PET before and after treatment. One example in which quantitative knowledge of tumor control probability is essential in determining which patients are suitable candidates for surgery after neoadjuvant chemoradiotherapy (CRT) is that of esophageal cancer. Tan et al. [6] showed that the use of features from the spatial distribution of SUV gives a more accurate prediction of esophageal cancer patients’ pathological response to CRT than the use of a single SUV value, such as the maximum SUV \((\text{SUV}_{\text{max}})\) within the tumor. That study was based on a group of 20 esophageal cancer patients with pretreatment and posttreatment \(^{18}\text{F-FDG}\) PET-CT images that were registered using rigid registration, allowing a voxel-to-voxel investigation of changes in SUVs before and after treatment [6]. In this work, we investigate the same cohort of 20 esophageal cancer patients evaluated in our previous study and use the mean values of the distribution of SUV in each patient to obtain a tumor control probability curve (based on pathological response) as a function of mean ratios of SUV before and after CRT.
Materials and Methods

Patient cohort and imaging techniques

This Institutional Review Board–approved study was based on a cohort of 20 esophageal cancer patients treated at our institution with trimodality therapy (CRT followed by surgery) from 2006 to 2009. All patients underwent both pre-CRT and post-CRT PET/CT imaging. PET images were attenuation corrected, with a resolution of 4.0×4.0×4.0 mm³ and CT images had a resolution of 0.98×0.98×4.0 mm³. Rigid image registration (VersorRigidDTTransform in the Insight Segmentation and Registration Toolkit 4.6.0; National Library of Medicine, Bethesda, MD) was used to register post-CRT CTs to pre-CRT CTs (details of patient characteristics, imaging protocols, and registration technique can be found in Tan et al. [6]). All patients were treated with external-beam radiotherapy (50.4 Gy; 1.8 Gy/day; 5 days/week) with concurrent chemotherapy consisting of cisplatin and 5-fluorouracil. The tumor volume in the pre-CRT PET image was defined as the region with SUV >2.5 (an SUV of 2.5 has been widely used for tumor delineation based on SUV PET values from before and after treatment). If we assume that the SUV value in a voxel j of the ¹⁸F-FDG 140 PET images correlates with the number of clonogenic cells in that voxel, it follows that the ratio of SUV values after and before CRT (i.e., SUV before/j(SUV after)/j) can potentially be a surrogate for the average survival fraction in that voxel [14]. Therefore, the average of SUVR(j) over all the voxels in the tumor delineated on pre-CRT PET can potentially be a surrogate for the average survival fraction in the tumor. We calculated the average <SUVR> for each patient. We also calculated other possible surrogates, such as <SUVR> and <SUVR> and <SUVR> for comparison. The mean values were calculated in the tumor region (defined in the PET images before therapy with SUV >2.5). For a possible surrogate of the average survival fraction to be useful, its values for responders must be significantly different from the values for nonresponders. This was evaluated by comparing the average values of each possible surrogate for responders and nonresponders with a t test at a significance level ≤ 0.05.

Maximum-likelihood estimate of TCP model parameters

Although <SUVR> as defined above is a reasonable surrogate for average survival fraction, the explicit functional dependence of <SUVR> with survival fraction is not known, so we cannot directly use Eq. (2) to relate TCP and <SUVR>. We propose the use of a sigmoid function to relate TCP with <SUVR> as a reasonable starting point. The sigmoid function is widely used to model TCP and normal tissue complication probability (NTCP) as a function of dose and was used previously to model tumor control probability based on PET images [11]. Figure 1 is a graphical representation of a typical sigmoid function.

It is constrained to the interval (0–1) and typically defined by two parameters: \( D_{50} \), the dose at which TCP is 0.5, and \( \gamma_{50} \), the normalized slope of the (sigmoid) curve at D50. Similarly, we define \( \gamma_{50} \), the value of <SUVR> at which TCP is 0.5 and \( \gamma_{50} \), as the slope of the curve at <SUVR>. For convenience we define the tumor recurrence probability
where the error function is the standard definition:

\[ \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt \]  

The log-likelihood (LL) can be calculated as:

\[ LL(SUV_{50}, \gamma_{50}) = \sum_{i} \log(p_i) \hat{R}_i + \log(1 - p_i)^{(1 - \hat{R}_i)} \]  

where \( R_i = 1 \) if the patient is a nonresponder and \( R_i = 0 \) if he or she is a responder [15]. The most likely values for SUVR\(_{50}\) and \( \gamma_{50} \) are obtained by maximizing LL(SUVR\(_{50}\), \( \gamma_{50} \)). The confidence intervals (CI) for SUVR\(_{50}\) and \( \gamma_{50} \) can be estimated using standard statistical methods assuming a Gaussian distribution with two degrees of freedom (e.g., as in Beringer et al. [16]).

**Results**

Table 1 presents the average values of each of the possible survival fraction surrogates for responders and nonresponders, as well as standard deviations and their P values. Table 1 shows that <SUVR\(_{after}/SUVR_{before}\)>, <SUVR\(_{after}/SUVR_{before}\)> and <SUVR\(_{before} - SUVR_{after}\)> each have significantly different values for responders and nonresponders (P<0.05). The fact that <SUVR\(_{after}/SUVR_{before}\)>, <SUVR\(_{after}/SUVR_{before}\)> and <SUVR\(_{after} - SUVR_{before}\)> are significantly different for responders and nonresponders validates these quantities as good candidates for surrogates of the effective survival fraction. Here we present our results by calculating TRP as a function of SUVR\(_{50}\), to illustrate the method. Equivalent results can be obtained using <SUVR\(_{after}/SUVR_{before}\)> or <SUVR\(_{after}/SUVR_{before}\)> . Figure 2 is a two-dimensional plot in which the x and y axes represent SUVR\(_{50}\) and \( \gamma_{50} \), respectively, and the color scale shows the LL calculated using Equation 6.

The maximum LL value is at SUVR\(_{50}=0.47\) (90% CI, 0.3-0.6) and \( \gamma_{50}=1.61 \) (90% CI, 0-4.2). To compare the model with our patient data, we divided our patients’ SUVR\(_{50}\) results into three bins (0.2-0.4, 0.4-0.6, and >0.6). Based on the numbers of responders and nonresponders in each group we plotted the histogram with the TRP and compared it with the model prediction (Figure 3). Although the error bars are large (as well as the confidence interval for the parameters), Figure 3 shows that the model describes the data reasonably well and illustrates the way in which a TRP and, equivalently, a TCP as a function of SUV can be derived from clinical data.

Table 1 also shows that <SUVR\(_{before}\)> tends to be higher for responders than for nonresponders. This trend, although not significant, seems to contradict the assumption that SUV is correlated with tumor burden, because we expect nonresponders to have higher tumor burdens than responders. This trend has been observed before in studies of initial SUV\(_{max}\) for lung cancer [17] as well as in esophageal cancer [18]. To address this puzzling issue, we studied the correlation of <SUVR\(_{after}\)> with <SUVR\(_{after}/SUVR_{before}\)> and <SUVR\(_{before} - SUVR_{after}\)> to determine
whether the response surrogates are correlated with the initial SUV values. Figure 4 shows a significant correlation between \(<SUV_{\text{before}}>-<SUV_{\text{after}}><SUV_{\text{after}}\) and \(<SUV_{\text{before}}\) (Pearson correlation coefficient=0.77; \(p=0.0008\)).

This correlation was even stronger among responders (correlation coefficient=0.92; \(p=0.005\)). A correlation trend between \(SUV_r\) and \(<SUV_{\text{before}}\) was also found but did not reach statistical significance.

Discussion

This work presents a proof of principle for a method to quantitatively relate the ratio of mean SUV after and before treatment to the probability of tumor recurrence in patients with esophageal. To the best of our knowledge, this is the first study where a quantitative relationship between average SUV before and after treatment and tumor control probability has been developed.

Although our confidence intervals are large as a result of 245 of the small sample size, our example shows that a patient-specific TRP curve (as depicted in Figure 3) can be derived with this method. This TRP curve could potentially be used to estimate the probability of disease recurrence after CRT given the value of \(SUV_r\) for a given patient, which could help in a personalized medicine approach to determine the need for subsequent surgery. Instead of looking for an arbitrary cut-off in SUVs and determining the sensitivity and specificity of a positive or negative test, our method uses a continuous TRP curve and offers the advantage of easily identifying patients for whom PET imaging response results should be labeled as inconclusive. In our example, patients with \(SUV_r\) between 0.3 and 0.6 (CI for \(SUV_{300}\)) have a 50/50 chance of recurrence. In that case, \(SUV_r\) should not be used as a determining factor for sending the patient to surgery. A limitation of our approach is that the patient number was small and a larger population may be needed to obtain TCP values with reasonably small CIs to clinically validate the model parameters. We used a sigmoid function to represent the TRP because it is a common choice for TCP versus dose and it is restricted to values from 0 to 1. The sigmoid function also has the property that for a steep slope it reproduces a step function, which is commonly used to report \(^{18}\text{F-FDG}\) PET imaging results.

The derivation of the radiobiological model rests in part on the assumption that SUVs are correlated with tumor burden. This assumption has been a topic of investigation in a number of studies for esophageal cancer that have shown mixed results. In an extensive review of \(^{18}\text{F-FDG-PET\) parameters as prognostic factors in esophageal cancer, Omloo et al. [4] found that 12 of 15 studies showed that although pretreatment \(^{18}\text{F-FDG uptake is a predictor for survival in univariate analysis, only 2 studies showed such uptake to be a predictor of survival in multivariate analysis. In our study, we did not find a statistically significant correlation between initial SUV and pathological response; in fact, we identified a small trend showing higher initial SUV for responders. Rizk et al. [18] found that pretreatment SUV was a significant predictor of survival for patients managed with surgery only (low SUV, greater survival). However, in a subsequent report, Rizk et al. [19] found that pretreatment SUV did not predict survival for patients treated with chemoradiation, in part because of the fact that patients with higher pretreatment SUVs responded better to therapy than those with lower SUVs. In a study of 103 patients Brown et al. [20] also found that high initial \(^{18}\text{F-FDG SUV on PET in esophageal cancer patients was a predictor of survival only for those treated with surgery; in patients treated with neoadjuvant therapy this difference disappeared, and a trend toward better survival was seen in patients with higher initial SUV. These results are consistent with our findings, in which we identified a trend toward higher initial SUV in the CR group and a correlation between better response and higher initial mean SUV, in agreement with the results of Rizk et al. [18]. Whether pretreatment SUV is associated with better outcomes, tumor cell density, or tumor proliferation is not a crucial assumption for our current work. The key assumption in our modeling is that the ratio of \(<SUV_{\text{after}}>/<SUV_{\text{before}}\) or the difference \(<SUV_{\text{before}}>-<SUV_{\text{after}}\) are surrogates of mean effective cell survival in the tumor.

The correlation of changes in SUV uptake after chemoradiation has been studied by several groups with mixed results: 4 of 10 studies in the review by Omloo et al. [4] found such correlation. Most of these studies, however, relied on SUVmax rather than the mean SUV, and some focused on survival as an endpoint rather than pathological response. Tan et al. [6] showed that using the average values (and other features of the distribution) can improve the predictive accuracy of \(^{18}\text{F-FDG PET in esophageal cancer. Our study showed a significant correlation of pathological response with changes in average SUV. Other groups have also found that considering the spatial extent properties of SUVs can increase predictive accuracy [21,22]. We used \(<SUV_{\text{before}}\) as our surrogate for cell survival, but other possible surrogates quantifying change in SUV could be used, for example \(<SUV_{\text{mean}}>-<SUV_{\text{initial}}\) or the ratio of the means of \(<SUV_{\text{mean}}>/<SUV_{\text{initial}}\), because both parameters are significantly different for responders than nonresponders. Moreover, this method could potentially be applied in other imaging modalities when a parameter is significantly different for responders and nonresponders. We believe that the development and validation of quantitative models of TCP as a function of molecular imaging markers will advance the understanding of the radiobiology of those markers.

The correlation of \(<SUV_{\text{mean}}>-<SUV_{\text{initial}}\) with \(<SUV_{\text{initial}}\), shown in Figure 4 underscores the complexity interpretation of SUVs. If SUV is representative of tumor burden, responders would be expected to have smaller \(<SUV_{\text{mean}}>-<SUV_{\text{initial}}\); the result (Table 1) shows a trend that is opposite to this reasoning. Although changes in SUV before and after treatment are typically used to characterize response (e.g., as in Aerts et al. [23]), it is believed that tumor regions with higher initial SUVs are at higher risk of recurrence (the basis of dose painting strategies).
However, the fact that in our group of esophageal cancer patients subjects with higher $\text{SUV}_{\text{ave, after}}$ tended to have larger decreases in SUV as represented by $\text{SUV}_{\text{ave, before}} - \text{SUV}_{\text{ave, after}}$ (a stronger response) shows that simple interpretations may not work because of the complex correlations among radiobiological parameters. As discussed above, other groups have also found correlations between initial SUV and response to chemoradiation [19-21] but those studies focused on either pathological response or patient survival as endpoints. To the best of our knowledge, this is the first study to investigate and find a correlation between initial average SUV and change in average SUV before and after treatment for esophageal cancer patients. Our result helps explain the findings from previous investigators as discussed above [18-21] and make the case for the need of systematic studies of these correlations to help understand and improve the interpretation of $^{18}$F-FDG PET images as 325 well as other molecular imaging markers.

**Conclusions**

The TCP model was characterized using SUV in tumor before and after therapy. According to the TCP model, $\frac{\text{SUV}_{\text{ave, after}}}{\text{SUV}_{\text{ave, before}}}$ is a possible surrogate for cell survival in esophageal cancer patients. Despite the fact that CIs are large because of the small patient sample, parameters for a TCP curve can be derived and an individualized TCP can be calculated for future patients. Initial SUV did not predict for response, and a correlation was found between surrogates for tumor burden and cell kill.

**Acknowledgment**

This work was supported in part by the National Cancer Institute Grant R01 CA172638. Shan Tan was supported in part by National Natural Science Foundation of China (NNSFC), under Grant Nos. 60971112 and 81375018, and Fundamental Research Funds for the Central Universities, under Grant No.2012QN086. We would like to thank Dr. Andrew Smith for his help with the software for data analysis and fruitful comments and discussion. We are also grateful to Dr. Nancy Knight for the technical editing of the manuscript.

**References**