Positron Emission Tomography/Computed Tomography before Treatment as a Predictor of $^{90}$Y-Ibritumomab Tiuxetan Response

Ryan Yudistiro$^{1,2}$, Yusri Dwi Heryanto$^1$, Sayaka Kodaira$^1$, Tetsuya Higuchi$^1$, Yukiko Arisaka$^1$, Azusa Tokue$^1$, Ayako Taketomi-Takahashi$^2$, Norifumi Tsukamoto$^3$, Akihiko Yokohama$^*$, Hiroshi Handa$^5$, Hiromi Koiso$^6$, Takuma Ishizaki$^5$ and Yoshito Tsushima$^1$

1 Department of Diagnostic Radiology and Nuclear Medicine, Gunma University Graduate School of Medicine, 3-39-22 Showa-machi, Maebashi, Gunma 371-8511, Japan
2 Department of Nuclear Medicine, MRCCC Siloam Hospital, School of Medicine of Universitas Pelita Harapan, Jakarta, Indonesia
3 Oncology Center, Gunma University Hospital, 3-39-15 Showa-machi, Maebashi, Gunma 371-8511, Japan
4 Blood Transfusion Services, Gunma University Hospital, 3-39-15 Showa-machi, Maebashi, Gunma 371-8511, Japan
5 Department of Hematology, Gunma University Hospital, 3-39-15 Showa-machi, Maebashi, Gunma 371-8511, Japan
6 Infection Control and Prevention Center, Gunma University Hospital, 3-39-15 Showa-machi, Maebashi, Gunma 371-8511, Japan

Abstract
Background and Aims: Radioimmunotherapy (RIT) appears as one of the most effective treatment options for patients with relapsed or resistant non-Hodgkin’s lymphoma (NHL). Our aim was to evaluate the role of pre-RIT $^{18}$F-fluorodeoxyglucose-positron emission tomography/computed tomography (FDG-PET/CT) as an early predictor of $^{90}$Y-Ibritumomab tiuxetan treatment response.

Methods: We included consecutive 20 patients with relapsed NHL (10 males; mean age, 58.5 ± 8.9 years old) who were treated with $^{90}$Y-Ibritumomab tiuxetan. FDG-PET/CT was performed before and after treatment. Semiquantitative parameters of all measurable FDG-avid lesions were measured and averaged. A measurable FDG-avid lesion was defined as a lesion that showed FDG uptake higher than liver with a diameter more than 1 cm. Treatment response was determined by visual assessment based on a five-point score criteria from FDG-PET/CT after treatment.

Results: Fourteen patients (70%) were classified as responders and the other six patients (30%) as non-responders. All semiquantitative parameters except for MTV demonstrated significantly lower values in the responders compared with the non-responders ($p<0.05$).

Conclusions: Semiquantitative evaluation by SUVmax, SUVpeak, and TLG before treatment were useful as early predictors of $^{90}$Y-Ibritumomab tiuxetan treatment response.

Introduction

Non-Hodgkin’s lymphoma (NHL) is the most common hematopoietic neoplasm with about 386,000 new cases in 2012 worldwide. The majority of NHL arises from B-cell that overexpress CD20 antigen. Although lymphomas may be classified or grouped in many ways, based on the growth rate, they may be grouped as indolent (slow growth rate) or aggressive (fast growth rate).

Patients with indolent NHL, including follicular B-cell NHL, tend to suffer relapses and become more resistant to therapy. Currently, chemotherapy, molecular targeting therapy, and radiotherapy are the standards of NHL treatment. However, unlike in aggressive NHL, treatment options for patients with relapsed or resistant indolent NHL remain complicated.

Radioimmunotherapy (RIT) has appeared as one of the most effective treatment options for patients with relapsed or resistant indolent NHL. RIT is a combination treatment of radionuclide and immunotherapy and may have better therapeutic outcomes through crossfire and bystander effects. $^{90}$Y-Ibritumomab tiuxetan and $^{131}$I-tositumomab are the first RITs, which have been approved by the United States Food and Drug Administration (FDA) for the treatment of relapsed and refractory low-grade, follicular, or transformed B-cell NHL.
Clinical trials of RIT in relapsed or resistant indolent NHL demonstrated high response rates and durability of remission with well tolerated side effects and toxicities.\textsuperscript{5,9,10} Although the high efficacy of RIT in patients with relapsed or resistant indolent NHL has already been well established, long-term toxicity, high cost, and the availability of RIT are still the potential drawbacks in the clinical setting.\textsuperscript{11} A pre-treatment predictor of favorable outcomes for RIT would have high clinical utility.

\textsuperscript{18}F-Fluorodeoxyglucose positron emission tomography/computed tomography (FDG-PET/CT) has been widely used and recommended as a main imaging modality in FDG-avid lymphoma. It has been reported that FDG-PET/CT may improve the accuracy of initial staging, prognostic assessment, treatment planning, and treatment response monitoring of lymphoma.\textsuperscript{12,13} It has also been reported that FDG-PET/CT performed after two cycles of chemotherapy (interim FDG-PET/CT), had prognostic value in NHL patients.\textsuperscript{14,15} However, the role of FDG-PET/CT before treatment as an early RIT response predictor has yet to be confirmed. The purpose of this study was to evaluate pre-RIT FDG-PET/CT as an early predictor of \textsuperscript{90}Y-Ibritumomab tiuxetan treatment.

\section*{Materials and Methods}

\subsection*{Patients}
A retrospective study was performed in twenty patients who underwent \textsuperscript{90}Y-Ibritumomab tiuxetan (Zevalin\textsuperscript{®}; 0.3-0.4 mCi/kg body weight; Fuji Film RI Pharma Co. Ltd, Tokyo, Japan) RIT for relapsed or resistant indolent NHL from 2009 to 2014 in our institution. All patients had FDG-PET/CT at least six months before and after treatment. We excluded patients who had no FDG-PET/CT within six months after treatment, patients with non-FDG-avid lymphoma, and patients who were not followed after treatment. Patients were followed until 1 March 2017. Our database and electronic charts were retrospectively reviewed to obtain follow-up data. The study protocol was approved by our institutional review board, and written informed consent was obtained from each patient before they were enrolled in this study.

\subsection*{Imaging protocol}
\textsuperscript{18}F-FDG was produced in our cyclotron facility. The patients fasted for at least six hours, and were free of alcohol and caffeine for 12 hours. One hour after intravenous injection of \textsuperscript{18}F-FDG (5 MBq/kg), FDG-PET/CT scans were performed with the Discovery STE (GE Healthcare, United States) or Biograph 16 (Siemens Medical Solutions, United States), with a field of view (FOV) of 700-mm and slice thickness of 3.27 mm. The CT was acquired to correct PET transmission using the following parameters: 140 kV, 120-240 mAs, and 128\texttimes 128 matrix. Patients were scanned from head to thigh. Shallow breathing was advised to avoid motion artifacts and to minimize misregistration of CT and PET images. Intravenous iodinated contrast media were not administered for CT scanning, and images were reconstructed by a conventional filtered back projection method. Axial full width at half maximum (FWHM) of 1 cm from the centre of the FOV was 6.3 mm. Intrinsic system sensitivity was 8.5 eps/kBq for 3-dimensional acquisition.

\subsection*{Image analysis}
FDG-PET/CT images before and after treatment were retrospectively reviewed by two experienced nuclear medicine physicians (SK and HT) who were blinded to all other imaging and clinical information. Semiquantitative parameters were obtained from a two-dimensional region of interest (2D-ROI) and three-dimensional volume of interest (3D-VOI). Adaptive-threshold of 41% was used in 3D-VOI (Fig. 1) to measure maximum and mean standardized uptake value (SUVmax and SUVmean) and metabolic tumor volume (MTV). Total lesion glycolysis (TLG) was obtained from the equation \([\text{SUVmean}] \times \text{MTV}\). The SUVmax and SUVpeak were obtained from 2D-ROI FDG-PET/CT images (Fig. 2A). SUVpeak was defined as an average value of SUV, which was measured in a circle ROI (fixed diameter of 1.2 cm), and placed at the hottest lesion area (Fig. 2B). These semiquantitative parameters of all measurable FDG-avid lesions from each patient were averaged and analyzed. A measurable FDG-avid lesion was defined as a lesion that showed FDG uptake higher than liver and a diameter more than 1 cm.

We followed the Lugano classification, a visual assessment based on a five-point score criteria from FDG-PET/CT after treatment,\textsuperscript{12} to evaluate treatment

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{Responder FDG-PET/CT images}
\end{figure}

A 57-years-old female patient showed relapse of follicular lymphoma four years after her last chemotherapy. In her pre-RIT FDG-PET/CT [A], multiple FDG-avid lesions were seen above and below the diaphragm. Some FDG-avid lesions in left axilla, pelvic, and inguinal regions were not analyzed since FDG uptake was lower than the liver. Complete response was seen on FDG-PET/CT three months after treatment (B). This patient was classified as a responder.
response to ⁹⁰Y-Ibritumomab tiuxetan. In this visual assessment, mediastinal blood pool and liver were used as threshold organs. Responders were defined as patients whose measurable FDG-avid lesions disappeared, or whose visual score(s) of the residual FDG-avid lesion(s) (scored ≤3) decreased. Non-responders were defined as patients who showed residual FDG-avid lesions (score ≥4 without significant change in FDG uptake) or appearance of new FDG-avid lesion(s).

Statistical analysis

Mann-Whitney U-test was used to compare the responders and non-responders for each parameter. The area under a receiving-operator-curve (ROC) was used to assess the accuracy and to infer the optimum cut-off values of each parameter for predicting response. Fischer exact test was also used. The survival curve was estimated using the Kaplan-Meier method. Overall survival was defined as the time from RIT until death from any cause. All statistical analyses were performed using GraphPad Prism version 7.00 (GraphPad Software, Inc., CA) and SPSS version 23 (SPSS, Inc., Chicago, IL), and p-value <0.05 was considered significant.

Results

The patients’ characteristics are summarized in Table 1. Interval time between FDG-PET/CT before treatment to RIT for all subjects was 2 ± 1.3 (range: 1 - 6) months. According to FDG-PET/CT after treatment, fourteen patients (70%) were classified as responders, whereas six patients (30%) were non-responders. Before RIT, twelve patients (60%) had early stage (stage I/II) disease and eight patients (40%) had advanced stage (stage III/IV) disease. There was no difference in pathological diagnosis or grade between two groups. A total of 74 FDG-avid lesions were analyzed, and the average number and length diameter of the FDG-avid lesions did not differ between responder and non-re-

<table>
<thead>
<tr>
<th>Table 1 Patient characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Patient gender</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Lesions</td>
</tr>
<tr>
<td>Number per patient</td>
</tr>
<tr>
<td>Length diameter (mm)</td>
</tr>
<tr>
<td>Non-Bulky</td>
</tr>
<tr>
<td>Bulky‡</td>
</tr>
<tr>
<td>Patient age</td>
</tr>
<tr>
<td>&lt;60</td>
</tr>
<tr>
<td>&gt;60</td>
</tr>
<tr>
<td>Histopathology</td>
</tr>
<tr>
<td>Follicular</td>
</tr>
<tr>
<td>Mantle cell</td>
</tr>
<tr>
<td>Grade</td>
</tr>
<tr>
<td>1 or 2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>Not available</td>
</tr>
<tr>
<td>Stage before RIT</td>
</tr>
<tr>
<td>I-II</td>
</tr>
<tr>
<td>III-IV</td>
</tr>
</tbody>
</table>

N.S=Not Significant; RIT: radioimmunotherapy, ‡Mean ± standard deviation (range); Mann-Whitney U-test

Fig. 2 Non-responder FDG-PET/CT images
A 50-years-old female patient showed relapse of follicular lymphoma two years after her last course of chemotherapy. In her pre-RIT FDG-PET/CT [A], two bulky FDG-avid lesions were seen in the pelvic region (white arrow). Residual bulky FDG-avid lesions (white arrowhead) [B], decreased in size, were still seen in FDG-PET/CT at three months after treatment. This patient was classified in a non-responder.
FDG-PET/CT for RIT predictor

sponder groups. All lesion in responders were non-bulky lesions, and only two bulky lesions were observed in one non-responder.

Table 2 shows the results of semiquantitative parameters of pre-therapy FDG-PET/CT in predicting RIT treatment response. All these semiquantitative parameters except for MTV demonstrated significantly lower values in the responders compared to the non-responders ($p < 0.05$). The optimal cut-off values identified by ROC analysis for both SUVmax using 3D-VOI and 2D-ROI were 5.35 and 5.15, respectively, with the same values for sensitivity and specificity of 83.3% and 71.3%. Meanwhile, the optimal cut-off values for SUVpeak and TLG was 4.35 and 19.5 gram, respectively, with sensitivity and specificity for SUVpeak of 83.3% and 92.9% and both 100% for TLG (Fig. 3). The cut-off of TLG was not only useful for predicting RIT response, but also prognosis. Patients with TLG less than 19.5 gram showed a significantly higher percentage of 7-year overall survival (50% versus 100%, $p < 0.01$) and longer median survival time (66 months vs undefined, $p < 0.01$) (Fig. 4). The other semiquantitative parameters did not show a significant difference in survival analysis.

**Discussion**

Our data demonstrated that all patients who responded to RIT had significantly lower values of all semiquantitative parameters on pre-therapy FDG-PET/CT except for MTV.

It has been reported that higher FDG-avidity was correlated with more aggressive lesions, and this may due to overexpression of glucose transporter-3 (GLUT-3). We suspect that our finding might reflect a transformation of lymphoma in non-responder patients to more aggressive subtypes, although we do not have

---

**Table 2** Semiquantitative parameters of pretherapy FDG-PET/CT in predicting RIT treatment response

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Responder*</th>
<th>Non-responder*</th>
<th>$p$-value$^b$</th>
<th>AUC</th>
<th>Optimum cut-off</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUVmax-VOI</td>
<td>4.46 ± 1.13</td>
<td>8.04 ± 3.43</td>
<td>0.008</td>
<td>0.867</td>
<td>5.35</td>
<td>83.33</td>
<td>73.33</td>
</tr>
<tr>
<td></td>
<td>(2.45 - 6.15)</td>
<td>(4.35 - 13.60)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUVmax-ROI</td>
<td>4.24 ± 1.09</td>
<td>7.67 ± 3.44</td>
<td>0.010</td>
<td>0.856</td>
<td>5.15</td>
<td>83.33</td>
<td>73.33</td>
</tr>
<tr>
<td></td>
<td>(2.70 - 5.95)</td>
<td>(3.90 - 13.30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUVpeak</td>
<td>3.33 ± 0.87</td>
<td>6.76 ± 3.45</td>
<td>0.004</td>
<td>0.894</td>
<td>4.05</td>
<td>83.33</td>
<td>93.33</td>
</tr>
<tr>
<td></td>
<td>(2.10 - 5.10)</td>
<td>(3.20 - 12.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUVmean</td>
<td>2.71 ± 0.75</td>
<td>5.24 ± 2.69</td>
<td>0.014</td>
<td>0.844</td>
<td>3.35</td>
<td>83.33</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>(1.30 - 3.80)</td>
<td>(2.30 - 9.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTV (cm$^3$)</td>
<td>11.60 ± 23</td>
<td>18.67 ± 17.50</td>
<td>0.178</td>
<td>0.700</td>
<td>9.1</td>
<td>66.67</td>
<td>73.33</td>
</tr>
<tr>
<td></td>
<td>(1.08 - 93.61)</td>
<td>(3.35 - 43.64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLG</td>
<td>29.37 ± 59.93</td>
<td>111.06 ± 155.61</td>
<td>0.006</td>
<td>0.878</td>
<td>14.85</td>
<td>100</td>
<td>73.33</td>
</tr>
<tr>
<td></td>
<td>(3.20 - 243.37)</td>
<td>(15.53 - 423.31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

AUC = area under the curve; SUVmax = maximum Standardized Uptake Value; VOI = Volume of Interest; ROI = Region of Interest; MTV = Metabolic Tumor Volume; TLG = Total Lesion Glycolysis; $a$ = mean ± standard deviation (range); $b$ = Mann-Whitney test
Our average SUVmax was slightly lower than other studies that showed SUV of >10 as a threshold in differentiating between indolent and aggressive of NHL.17,19 Lymphoma cells in mass lesions may not be homogeneous, possibly containing more aggressive cells generated by transformation. The presence of these aggressive cells is likely to determine the potential tumor aggressiveness, rather than the total number of lymphoma cells or the total volume of tumors. In the current study, SUVs and TLG also reflect the viable cell density of lesions which are useful for predicting treatment response. Meanwhile, MTV was not useful for predicting treatment response. This difference may partially support our suggestion concerning transformation, and our results are consistent with previous reports.20,21

All FDG uptake parameters, including SUVmax (from both ROI and VOI) and SUVpeak, showed significant differences between responders and non-responders. These parameters reflect viable cell density and aggressiveness of lesions, although each parameter has its limitations. The most widely used parameter in the clinical setting is SUVmax, i.e. the SUV value of the maximum intensity voxel within a defined subspace of a PET image. SUVmax is simple and highly reproducible, but strongly affected by noise and partial volume effect that will underestimate values in small tumors.22 SUVpeak, which was not used in previous studies, was significantly lower in responders in our study. SUVpeak has been recently proposed as a more robust alternative of SUVmax, since SUVpeak is the most reliable parameter for FDG-avid tumor quantification regardless of acquisition time.23 Additional quantitative information from SUVmax and SUVpeak values may strengthen the position of pre-RIT FDG-PET/CT as an imaging biomarker for early treatment response prediction.

Some studies have suggested that metabolic tumor burden such as MTV and TLG may be better predictors of survival than FDG uptake parameters such as SUV in malignancy including NHL.20,21 However, in our findings, only TLG was an early RIT response predictor with excellent sensitivity and specificity. Moreover, low TLG values indicated a significantly higher 7-year overall survival and longer median survival time, indicating TLG is not only a short-term but also a long-term prognostic predictor. MTV reflects the disease burden: meanwhile, TLG is a volumetric measure that combines not only metabolic tumor burden but also the metabolic intensity and reflects average SUV volume. This finding of TLG being more reliable than MTV may support our suggestion that lesion metabolic intensity and aggressiveness contributed more to RIT response than lesion burden.

It has been reported that 111In-Ibritumomab tiuxetan accumulation also correlated with overall response rate and disease progression. The heterogenous intratumoral distribution of 111In-Ibritumomab tiuxetan using single photon emission computed tomography (SPECT) correlated with tumor treatment response. Although SUVmax showed a positive correlation with 111In-Ibritumomab tiuxetan uptake, the heterogenous intratumoral distribution of FDG-PET/CT before treatment was not significantly different between responders and non-responders.21 In another recent study, the heterogeneity of intratumoral measurement using texture parameters and histogram indices in FDG-PET/CT were useful as treatment response predictors for several types of cancer,24 although we need to clarify their usefulness in NHL.

Currently, there are many treatment options available for advanced stage or relapsed NHL patients. Although Rituximab has changed the clinical pathway of NHL patients, RIT has been also demonstrated higher response rates either when used as single or consolidation therapy for the initial treatment.27,28 Therefore, an early predictive biomarker is needed to select patients who are most likely to benefit for RIT. From our study, we suspect that patients with SUVmax, SUVpeak, and TLG above the cut-off values of pre-RIT FDG-PET/CT may not be appropriate candidates for RIT, and other treatment options should be considered.

There were several limitations in our study. First, our study was retrospectively performed and included a small number of inhomogeneous patients with skewed distribution. Our results should be confirmed by further prospective studies with a large number of patients. Second, we did not compare our method with other clinical risk stratification methods. Thirdly, we used FDG-PET/CT imaging protocol commonly used in clinical practice. Quantification of the semiquantitative parameters employed in our study was sensitive to ROI and VOI definition. In our study, however, inter-observer agreement was not tested.

**Conclusion**

We found that the semiquantitative parameters of FDG-PET/CT before RIT treatment had significantly lower values in responders compared to non-responders, supporting the potential of pre-treatment FDG-PET/CT as a valuable imaging biomarker for early prediction of RIT response. In particular, TLG may not only predict short-term response but also long-term prognosis. Further prospective studies with a larger number of patients are encouraged.

**Conflicts of interest**

The authors declare that they have no conflict of financial and non-financial interest.

**References**

3. Fisher RI. Overview of non-Hodgkin’s lymphoma: biology,


