Acta Oncologica, 2016; 55: 386–390

Taylor & Francis
Taylor & Francis Group

LETTER TO THE EDITOR

Lymphoma grading with FDG-PET/CT readdressed: Direct and timely histopathological correlation study

HUGO J. A. ADAMS[1], JOHN M. H. DE KLERK[2], ROB FIJNHEER[3], BEN G. F. HEGGELMAN[4], STEFAN V. DUBOIS[5], RUTGER A. J. NIEVELSTEIN[1] & THOMAS C. KWEE[1]

[1]*Department of Radiology and Nuclear Medicine, University Medical Center Utrecht, Utrecht, The Netherlands,* [2]*Department of Nuclear Medicine, Meander Medical Center, Amersfoort, The Netherlands,* [3]*Department of Hematology, Meander Medical Center, Amersfoort, The Netherlands,* [4]*Department of Radiology, Meander Medical Center, Amersfoort, The Netherlands and* [5]*Department of Pathology, Meander Medical Center, Amersfoort, The Netherlands*

To the Editor,

Non-Hodgkin lymphomas encompass a histologically heterogeneous group of cancers, ranging from the more indolent follicular lymphoma, to the more aggressive diffuse large B-cell and Burkitt lymphomas [1]. The availability of a non-invasive whole body imaging modality for lymphoma grading would be of advantage because it may be helpful to steer the diagnosis, guide biopsies (to the lymphomatous site that is presumed to be most aggressive at imaging), assess intra-individual tumor heterogeneity, and to detect high-grade transformation in patients with indolent lymphoma.

^{18}F-fluoro-2-deoxy-D-glucose positron emission tomography/computed tomography (FDG-PET/CT) is a standard method for pre- and post-treatment evaluation of lymphoma [2]. With the exception of extranodal marginal zone lymphoma and small lymphocytic lymphoma, most lymphoma subtypes have high FDG avidity [3]. Previous studies reported FDG uptake of aggressive lymphomas to be higher than that of indolent lymphomas [4–6]. However, although these studies included around 100 or more patients [5–7], in one study the time interval between PET scanning and biopsy was up to 90 days and the FDG uptake measurement was discordant from the location of biopsy in some patients [4], and in two studies the time interval between PET scanning and biopsy was not mentioned and the FDG uptake measurement was not reported to match the site of biopsy [5,6]. As a result of the important methodological shortcomings of these studies [4–6], the true value of FDG-PET/CT for lymphoma grading is still unknown.

The purpose of this study was therefore to assess the diagnostic value of FDG-PET/CT for lymphoma grading, using direct (site-matched) and timely (within one month) histopathological correlation.

Material and methods

Inclusion criteria for this study were: histologically proven lymphoma (either Hodgkin or non-Hodgkin lymphoma) and availability of pretreatment FDG-PET/CT within one month before diagnostic (excisional or needle) biopsy of at least one extramedullary (i.e. non-bone marrow) lymphomatous site. Exclusion criteria for this study were: FDG-PET/CT not performed within one month before diagnostic biopsy, inability to determine the lymphoma subtype on the basis of the biopsy, lack of sufficient information to determine the exact location of the diagnostic biopsy, diagnosis of lymphoma solely based on bone marrow biopsy findings, and usage of chemotherapy, immunotherapy, radiotherapy or systemic corticosteroids within three months before FDG-PET/CT.

FDG-PET/CT was performed using a 40-detector row PET/CT scanner (Biograph 40 TruePoint PET/CT, Siemens Healthcare, Malvern, PA, USA). Patients fasted for at least six hours. Blood glucose

Correspondence: H. J. A. Adams, University Medical Center Utrecht, Department of Radiology and Nuclear Medicine, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands. Tel: + 31 88 7556687. Fax: + 31 30 2581098. E-mail: h.j.a.adams@gmail.com

(Received 30 March 2015; accepted 09 April 2015)

ISSN 0284-186X print/ISSN 1651-226X online © 2015 Informa Healthcare
DOI: 10.3109/0284186X.2015.1041652

levels were checked to be less than 11 mmol/L before 3 MBq/kg body weight of FDG was administered. Sixty minutes after FDG injection, PET/CT images were acquired from mid-thigh to skull base.

An experienced reader (H.J.A.A.) used the Region of interest visualization, evaluation, and image registration (ROVER) software (ABX advanced biochemical compounds GmbH, Radeberg, Germany) for FDG uptake measurements of all biopsied sites. The location of FDG uptake measurement was determined based on radiological and surgical reports. However, the reader was blinded to other clinical, laboratory, histopathological (including lymphoma subtype) and follow-up information at the time of FDG uptake measurements. Using a threshold setting of 40% of the maximum SUV (SUV_{max}) a delineated tumor volume of interest (VOI) was created at the site that was biopsied after the FDG-PET/CT acquisition. The software then automatically calculated 3D partial volume corrected mean SUV ($cSUV_{mean}$), SUV_{max}, and peak SUV (SUV_{peak}) of this VOI.

All tissue specimens were examined by experienced hematopathologists, subtyped and subsequently classified into major groups as either aggressive non-Hodgkin lymphoma, indolent non-Hodgkin lymphoma or Hodgkin lymphoma, according to the 2008 WHO classification of tumours of haematopoietic and lymphoid tissues [7]. Mantle cell lymphoma was regarded as an intermediate form of non-Hodgkin lymphoma and was not classified as indolent or aggressive [7].

Kolmogorov-Smirnov tests were used to check whether $cSUV_{mean}$, SUV_{max}, and SUV_{peak} of the major lymphoma classification groups (i.e. aggressive non-Hodgkin lymphoma, indolent non-Hodgkin lymphoma, and Hodgkin lymphoma) were normally distributed. Differences in mean $cSUV_{mean}$, SUV_{max}, and SUV_{peak} among the major lymphoma classification groups were assessed by using one-way analysis of variance (ANOVA) with Student-Newman-Keuls post-hoc testing. In case of a significant difference in any of the FDG uptake metrics between two of the aforementioned groups, additional receiver operating characteristic (ROC) analysis was performed to determine the area under the ROC curve (AUC) and optimal cut-off value with corresponding sensitivity and specificity. The level of statistically significant difference was set at $p < 0.05$. Statistical analyses were executed using MedCalc version 10.4.5.0 software (MedCalc, Ostend, Belgium).

Results

A total of 121 patients (68 men and 53 women, mean age: 60.9 years, age range: 16–90 years) were finally included. The mean time interval between FDG-PET/CT and diagnostic biopsy was 7.7 days [standard deviation (SD): 7.4 days, range 0–31 days]. Number of patients included with specific lymphoma subtypes are displayed in Table I.

Kolmogorov-Smirnov tests confirmed that $cSUV_{mean}$, SUV_{max}, and SUV_{peak} in each of the major lymphoma classification group (i.e. aggressive non-Hodgkin lymphoma, indolent non-Hodgkin lymphoma, and Hodgkin lymphoma) were normally distributed, justifying the use of the one-way ANOVA. The one-way ANOVA revealed a significant overall effect of the major lymphoma classification group on the measured mean $cSUV_{mean}$ ($p < 0.001$), SUV_{max} ($p < 0.001$), and SUV_{peak} ($p < 0.001$). Pairwise comparisons revealed that aggressive non-Hodgkin lymphomas had significantly higher ($p < 0.05$) $cSUV_{mean}$, SUV_{max}, and SUV_{peak} than indolent non-Hodgkin lymphomas, and that aggressive non-Hodgkin lymphomas had significantly higher ($p < 0.05$) $cSUV_{mean}$, SUV_{max}, and SUV_{peak} than Hodgkin lymphomas (Figure 1). However, there were no significant differences ($p > 0.05$) in $cSUV_{mean}$, SUV_{max}, and SUV_{peak} between indolent non-Hodgkin lymphomas and Hodgkin lymphomas. AUCs of $cSUV_{mean}$, SUV_{max}, and SUV_{peak} for the discrimination between aggressive non-Hodgkin lymphoma and indolent non-Hodgkin lymphoma were 0.892 (95% CI 0.814–0.946), 0.889 (95% CI 0.810– 0.944), and 0.867 (95% CI 0.784–0.927), respectively (Figure 2). Optimal cut-off values for $cSUV_{mean}$, SUV_{max}, and SUV_{peak} for the discrimination between aggressive non-Hodgkin lymphoma and indolent non-Hodgkin lymphoma were 9.4, 10.8, and 9.9, and yielded sensitivity and specificity combinations of 80.4% and 89.4%, 82.4% and 89.4%, and 70.6% and 93.6%, respectively. AUCs of $cSUV_{mean}$, SUV_{max}, and SUV_{peak} for the discrimination between aggressive non-Hodgkin lymphoma and Hodgkin lymphoma were 0.753 (95% CI 0.637–0.847), 0.777 (95% CI 0.663–0.866), and 0.763 (95% CI 0.648–0.855), respectively (Figure 2). Optimal cut-off values for $cSUV_{mean}$, SUV_{max}, and SUV_{peak} for the discrimination between aggressive lymphoma and Hodgkin lymphoma were 9.2, 12.3, and 8.8, which yielded sensitivity and specificity combinations of 80.4% and 66.7%, 70.6% and 76.2%, and 78.4% and 66.7%, respectively.

Discussion

The results of this study, in which a direct (site-matched) and timely (within one month) histopathological correlation was used, show that FDG uptake measurements have a high diagnostic value in differentiating aggressive from indolent non-Hodgkin lymphoma, and a moderately high diagnostic value in differentiating aggressive non-Hodgkin lymphoma

Table I. Mean, SD, and range of $cSUV_{mean}$, SUV_{max}, and SUV_{peak} for each individual lymphoma subtype.

Lymphoma subtype	Mean $cSUV_{mean}$±SD	Range $cSUV_{mean}$	Mean SUV_{max}±SD	Range SUV_{max}	Mean SUV_{peak}±SD	Range SUV_{peak}
Aggressive non-Hodgkin lymphoma (n = 51)	14.5 ± 7.1	2.8–36.6	17.1 ± 8.1	4.0–38.7	14.2 ± 7.3	3.6–36.4
Diffuse large B-cell lymphoma (n = 38)	15.5 ± 7.0	3.7–36.6	18.5 ± 7.8	4.4–38.7	15.4 ± 7.5	3.7–36.4
High-grade (3b) follicular lymphoma (n = 5)	14.4 ± 8.9	5.7–29.0	14.4 ± 7.0	7.3–24.6	12.2 ± 6.2	6.3–21.4
Anaplastic large cell lymphoma (n = 2)	8.2 ± 4.6	4.9–11.4	9.8 ± 5.3	6.0–13.5	7.9 ± 5.0	4.3–11.4
Burkitt lymphoma (n = 2)	14.9 ± 7.0	9.9–19.8	22.8 ± 12.2	14.1–31.4	17.2 ± 5.9	13.0–21.3
Peripheral T-cell lymphoma not otherwise specified (n = 2)	9.9 ± 2.0	8.5–11.3	9.3 ± 2.5	7.5–11.1	7.7 ± 2.5	5.9–9.5
Angioimmunoblastic T-cell lymphoma (n = 1)	9.1	-	11.2	-	10.1	-
T-cell lymphoblastic lymphoma (n = 1)	2.8	-	4.0	-	3.6	-
Indolent non-Hodgkin lymphoma (n = 47)	6.2 ± 2.7	1.2–12.3	7.4 ± 3.3	1.8–16.4	6.3 ± 2.8	1.3– 13.9
Low-grade (1–3a) follicular lymphoma (n = 33)	7.2 ± 2.6	2.5–12.3	8.5 ± 3.1	3.1–16.4	7.1 ± 2.7	2.2–13.9
Small lymphocytic lymphoma (n = 5)	4.5 ± 1.8	2.5–7.1	6.0 ± 1.9	4.3–8.8	5.2 ± 1.8	3.6–7.8
Lymphoplasmacytic lymphoma (n = 4)	2.8 ± 1.7	1.2–4.3	3.7 ± 2.3	1.8–6.5	3.4 ± 2.5	1.3–6.6
Extranodal marginal zone lymphoma (n = 3)	4.2 ± 0.4	3.8–4.6	5.0 ± 0.4	4.7–5.4	4.1 ± 0.4	3.8–4.5
Nodal marginal zone lymphoma (n = 2)	4.6 ± 3.5	2.1–7.0	5.0 ± 3.8	2.3–7.7	4.4 ± 3.7	1.8–7.0
Mantle cell lymphoma (n = 2)	8.6 ± 1.1	7.8–9.4	9.9 ± 1.8	8.6–11.1	8.9 ± 2.0	7.5–10.3
Hodgkin lymphoma (n = 21)	8.9 ± 3.9	3.8–17.5	10.1 ± 3.9	5.3–18.8	8.4 ± 3.6	4.1–16.3

from Hodgkin lymphoma. The reported FDG uptake thresholds may be helpful in clinical practice for separating these entities. However, this method is not useful in distinguishing indolent non-Hodgkin lymphoma from Hodgkin lymphoma. Of interest, relatively high FDG uptake metrics were only found in aggressive non-Hodgkin lymphoma. Unlike in aggressive non-Hodgkin lymphomas, $cSUV_{mean}$, SUV_{max}, and SUV_{peak} in indolent non-Hodgkin lymphomas never exceeded 12.3, 16.4, and 13.9. Similarly, $cSUV_{mean}$, SUV_{max}, and SUV_{peak} in Hodgkin lymphomas never exceeded 17.5, 18.8, and 16.3.

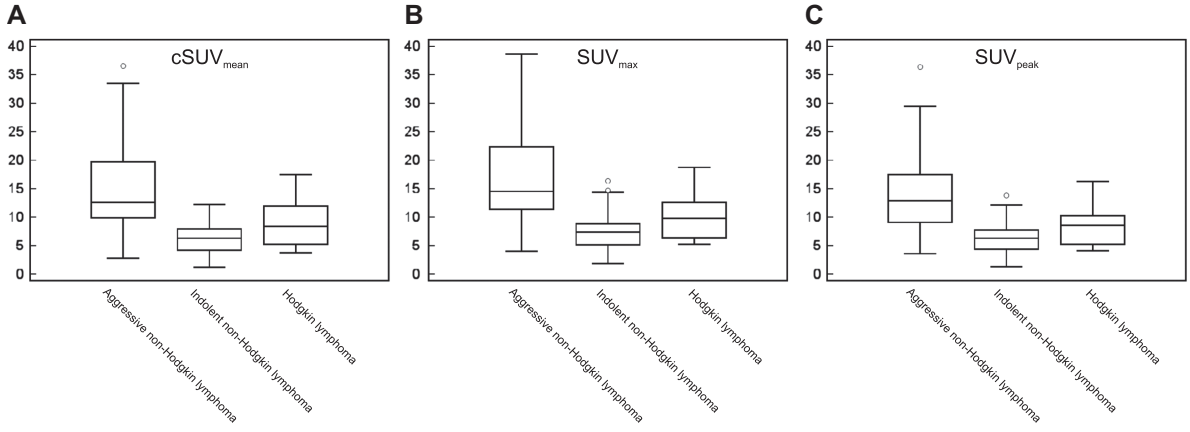

Figure 1. Box-and-whisker plots show median (middle line of box), quartiles (top and bottom lines of box), upper extreme value (upper whisker), lower extreme value (lower whisker), and outliers (circles) for $cSUV_{mean}$ (a), SUV_{max} (b), and SUV_{peak} (c) according to major lymphoma classification group. $cSUV_{mean}$, SUV_{max}, and SUV_{peak} of aggressive non-Hodgkin lymphoma were significantly higher than those of indolent non-Hodgkin lymphoma and Hodgkin lymphoma ($p = 0.05$). However, there were no significant differences ($p > 0.05$) in $cSUV_{mean}$, SUV_{max}, and SUV_{peak} between indolent non-Hodgkin lymphomas and Hodgkin lymphomas.

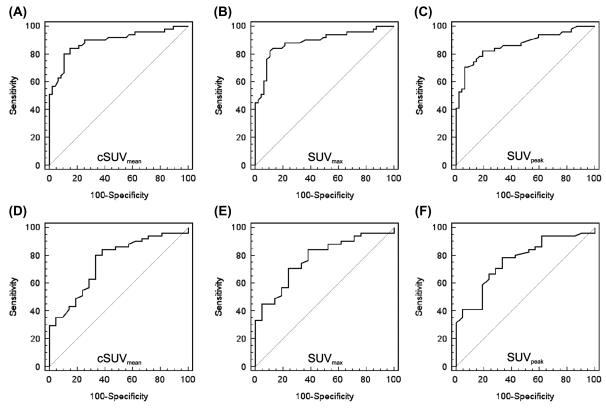

Figure 2. ROC curves for cSUV$_{mean}$ (a), SUV$_{max}$ (b), and SUV$_{peak}$ (c) measurements for the discrimination between aggressive and non-Hodgkin indolent lymphoma. AUCs of cSUV$_{mean}$, SUV$_{max}$, and SUV$_{peak}$ for the discrimination between aggressive and indolent non-Hodgkin lymphoma were 0.892 (95% CI 0.814–0.946), 0.889 (95% CI 0.810– 0.944), and 0.867 (95% CI 0.784–0.927), respectively. ROC curves for cSUV$_{mean}$ (d), SUV$_{max}$ (e), and SUV$_{peak}$ (f) measurements for the discrimination between aggressive non-Hodgkin lymphoma and Hodgkin lymphoma. AUCs of cSUV$_{mean}$, SUV$_{max}$, and SUV$_{peak}$ for the discrimination between aggressive non-Hodgkin lymphoma and Hodgkin lymphoma were 0.753 (95% CI 0.637–0.847), 0.777 (95% CI 0.663–0.866), and 0.763 (95% CI 0.648–0.855), respectively.

Awareness that relatively high FDG uptake is specific for aggressive disease is crucial when interpreting FDG-PET/CT scans because it can suggest the diagnosis of aggressive non-Hodgkin lymphoma and high-grade transformation in indolent lymphoma. Although histological confirmation of high-grade transformation remains obligatory, FDG-PET/CT may guide biopsies to the lymphomatous site that is likely most aggressive.

Three large studies including around 100 or more patients have previously been published on this topic [4–6]. Schöder et al. [4] performed FDG-PET SUV$_{max}$ measurements in 97 patients with non-Hodgkin lymphoma. FDG uptake was reported to be lower in indolent than in aggressive non-Hodgkin lymphoma for patients with new (SUV$_{max}$, 7.0 ± 3.1 vs. 19.6 ± 9.3; $p < 0.01$) and relapsed (SUV$_{max}$, 6.3 ± 2.7 vs. 18.1 ± 10.9; $p = 0.04$) disease. ROC analysis for discriminating between aggressive and indolent non-Hodgkin lymphoma was performed in 69/97 patients for whom the site of biopsy was the same as the site at which SUV$_{max}$ was measured and resulted in an AUC of 0.847. Using a cut-off SUV$_{max}$ of 10 provided a sensitivity of 71% and a specificity of 81% to detect aggressive disease. Drawbacks of the study by Schöder et al. [4] are the large (up to 90 days) time interval between FDG-PET and histological examination, that it was not reported whether the study included patients who underwent FDG-PET/CT after diagnostic biopsy (which could have affected local FDG uptake at the biopsy site), that Hodgkin lymphomas were not included and that outdated stand-alone PET systems were used [4]. In another study, Tsukamoto et al. [5] included 255 patients with various subtypes that were classified into four groups: Hodgkin lymphoma (mean SUV$_{max}$ = 6.6), indolent B-cell lymphoma (mean SUV$_{max}$ = 3.3), aggressive B-cell lymphoma (mean SUV$_{max}$ (mean SUV$_{max}$ = 9.9), and natural killer cell/T-cell lymphomas (mean SUV$_{max}$ = 9.4). Tsukamoto concluded that the SUV$_{max}$ of Hodgkin lymphoma was significantly higher than that of indolent non-Hodgkin lymphoma (note that this was not the case in the present study), but lower than that in aggressive non-Hodgkin lymphoma. In addition, the SUV$_{max}$ in indolent non-Hodgkin lymphoma was significantly lower than that in aggressive and natural killer cell/T-cell non-Hodgkin lymphomas. Unfortunately, Tsukamoto et al. [5] did not perform any ROC analysis. The most important drawbacks of the study by Tsukamoto et al. [5] are that the location of the SUV$_{max}$ measurement was not reported (and thus unlikely to be performed at the site of the biopsy) and that the interval between FDG-PET and diagnostic biopsy was not mentioned. Furthermore, an outdated stand-alone PET system was used [5]. Finally, Ngeow et al. [6] included 63 patients with aggressive and 21 with indolent B-cell non-Hodgkin lymphomas for ROC analysis. The AUC of the FDG-PET SUV$_{max}$ for discriminating between aggressive and indolent non-Hodgkin lymphoma was 0.81, and an optimal cut-off value of 10 yielded a sensitivity of 91% and a specificity of 62%. However, Ngeow et al. [6] did not report the location of the FDG uptake measurement and the time interval between FDG-PET and diagnostic biopsy, which introduces an important potential bias. Unlike the previous three larger studies on this topic that suffered from important methodological drawbacks [4–6], the present study is essentially different in that it provided a direct (site-matched) and timely (within one month) correlation between the FDG uptake measurement and the histopathological specimen.

This study had several limitations. First, although a total of 121 patients was included in this study, the number of patients with some specific lymphoma subtypes was relatively low. This is due to the fact that many of the potentially eligible patients had to be excluded because a direct and timely correlation between FDG uptake measurements and histopathology could not be achieved, and due to the relatively rare incidence of some lymphoma subtypes. Second, since only one observer performed the FDG uptake measurements once, intra- and interobserver agreement were not determined. Third, although a

standard FDG-PET/CT protocol was applied and an EANM approved PET/CT system was used for image acquisition, it should be realized that the reported FDG uptake metrics may vary from those obtained in other institutions because of differences in technical and patient-related factors.

In conclusion, this direct and timely histopathological correlation study provides evidence to confirm that FDG-PET/CT is reasonably accurate in differentiating aggressive non-Hodgkin lymphoma from indolent non-Hodgkin lymphoma and Hodgkin lymphoma.

Acknowledgments

This project was financially supported by an Alpe d'HuZes/Dutch Cancer Society Bas Mulder Award for T.C.K. (grant number 5409). Data collection, data analysis, and interpretation of data, writing of the paper, and decision to submit were left to the authors' discretion and were not influenced by Alpe d'HuZes/Dutch Cancer Society. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

[1] Shankland KR, Armitage JO, Hancock BW. Non-Hodgkin lymphoma. Lancet 2012;380:848–57.

[2] Barrington SF, Mikhaeel NG, Kostakoglu L, Meignan M, Hutchings M, Mueller SP, et al. Role of imaging in the staging and response assessment of lymphoma: Consensus of the International Conference on Malignant Lymphomas Imaging Working Group. J Clin Oncol Epub 2014 Aug 11.

[3] Weiler-Sagie M, Bushelev O, Epelbaum R, Dann EJ, Haim N, Avivi I, et al. (18)F-FDG avidity in lymphoma readdressed: A study of 766 patients. J Nucl Med 2010;51:25–30.

[4] Schoder H, Noy A, Gonen M, Weng L, Green D, Erdi YE, et al. Intensity of 18fluorodeoxyglucose uptake in positron emission tomography distinguishes between indolent and aggressive non-Hodgkin's lymphoma. J Clin Oncol 2005; 23:4643–51.

[5] Tsukamoto N, Kojima M, Hasegawa M, Oriuchi N, Matsushima T, Yokohama A, et al. The usefulness of (18) F-fluorodeoxyglucose positron emission tomography ((18) F-FDG-PET) and a comparison of (18)F-FDG-pet with (67) gallium scintigraphy in the evaluation of lymphoma: Relation to histologic subtypes based on the World Health Organization classification. Cancer 2007;110:652–9.

[6] Ngeow JY, Quek RH, Ng DC, Hee SW, Tao M, Lim LC, et al. High SUV uptake on FDG-PET/CT predicts for an aggressive B-cell lymphoma in a prospective study of primary FDG-PET/CT staging in lymphoma. Ann Oncol 2009; 20:1543–7.

[7] Swerdlow SH, Campo E, Harris NL, Pileri SA, Stein H, Thiele J, et al. WHO classification of tumours of haematopoietic and lymphoid tissues, 4th ed. Lyon: World Health Organization; 2008.