FDG PET-CT in Follicular Lymphoma: A Case-Based Evidence Review

Stephen D. Smith, M.D.
Department of Medicine, Division of Medical Oncology
University of Washington School of Medicine
Seattle Cancer Care Alliance
825 Eastlake Ave E, G3-200
Seattle, WA 98109
ssmith50@u.washington.edu

Mary Redman, Ph.D.
Assistant Member, Public Health Sciences Division
Fred Hutchinson Cancer Research Center

Kieron Dunleavy M.D.
Staff Clinician, Lymphoid Malignancies Branch
National Cancer Institute
Bethesda, MD
Case

A 52 year-old female develops right neck lymphadenopathy, in the absence of systemic symptoms. Excisional biopsy shows grade 1-2 follicular lymphoma (FL). Physical exam reveals a cluster of 1-2 cm right cervical nodes, but no other abnormalities. Lactate dehydrogenase is 400 (normal, 140-210 U/L), beta-2 microglobulin 4.1 (normal, 0-2.5 mcg/mL), and a CBC with differential is normal.

18-Fluorodeoxyglucose (FDG) positron emission tomography with computed tomography (PET-CT) is performed, showing FDG-avid lymphadenopathy above and below the diaphragm, ranging in size from 1.2-6 cm. Three FDG-avid lymph nodes measure more than 3 cm. A single 6 cm pelvic node has a maximum standardized uptake value (SUVmax) of 22, and other SUVmax range from 6-8. An absence of FDG marrow uptake is noted.

How should these PET-CT findings be integrated into further evaluation, prognostication, and treatment recommendations for this patient with FL? Should repeat PET-CT be performed upon completion of therapy?

Background

The use of PET-CT in oncology is increasing, and its role in the assessment and management of lymphoma has evolved. Given the variable glucose avidity and heterogeneous behavior of lymphomas, it is unsurprising that histology, timing in relation to therapy, and interpretation methods influence PET-CT findings. Indolent lymphomas, characterized by variable FDG avidity and often a prolonged natural history, comprise a unique context for assessing the merits of PET-CT. Prior consensus guidelines in 2007, reflecting a paucity of data, recommended PET-CT in restricted fashion for indolent lymphomas, such as clinical trials incorporating response rate as a primary endpoint. However, increasing evidence supports its role in FDG-avid indolent NHL subtypes, particularly FL, a histology in which PET-CT is frequently performed in both community and academic settings. Recently, formal guidelines for the use of PET-CT in FL have shifted, recommending its use for initial staging, evaluation for transformation, and response assessment after first-line therapy. PET imaging offers several benefits, including the potential for improved staging accuracy and evaluation for large cell transformation, which may optimize selection of first-line therapy. In the post-treatment setting, accurate identification of patients at highest risk of early relapse and mortality may inform surveillance methods, or the need for additional therapy. Nonetheless, until prospective studies are available, the impact of PET-CT on outcomes in FL remains to be defined, and integration into clinical management will require nuanced judgment based primarily on retrospective data. We present an evidence-based, focused review of the role of PET-CT in follicular lymphoma, identifying relevant references through PubMed searches, existing review articles, and expert sources. A pooled analysis was undertaken to assess the impact of PET-based imaging on initial stage, compared to conventional CT. Recommendations for the use of PET-CT in FL are provided, and rated in terms of strength according to the GRADE system.
PET-CT for Initial Staging of FL

Historically, a number of studies have demonstrated that PET-based imaging is sensitive for staging FL irrespective of grade.\textsuperscript{10-14} PET-based imaging (and more recently, combined PET-CT) identifies a greater extent of nodal and extranodal disease sites than standard staging including CT.\textsuperscript{12-17} A study published in 2008 by Janikova and colleagues found that among 62 newly diagnosed FL patients, PET-based staging identified a different disease distribution (compared to conventional CT) in 29 patients, and changed the stage in 6/62 (10%).\textsuperscript{16} Another retrospective study restricted to patients with early stage FL by conventional staging including CT found that among 42 patients, PET results were projected to alter stage designation in 13 patients (31%), and management in 19 (45%).\textsuperscript{13} To elucidate the impact of PET-based staging on management, Scott and colleagues performed a prospective study in which clinicians devised treatment plans for 74 patients with indolent NHL before and after PET imaging.\textsuperscript{18} The addition of PET to staging led to a revised treatment plan in 25 patients (34%), including a shift to palliative-intent therapy in 7 patients. Patients with stage I-II indolent NHL defined by PET imaging (treated primarily with radiotherapy) had excellent outcomes, superior to patients with stage III-IV disease.

More recently, Luminari and colleagues reported on 142 FL patients with available pretreatment PET-CT imaging who were included in the FOLL-05 trial, a prospective 3-arm comparison of first-line chemioimmunotherapy regimens.\textsuperscript{14} Forty-three percent had a different number of nodal sites (including 34% showing more sites) than were visualized with conventional CT. PET-CT upstaged 15 patients, and downstaged 5. Analogous to findings of Wirth and colleagues, most patients (15 of 24) previously classified as having early stage disease by CT were found to have advanced stage disease using PET-CT. Frequent bone, spleen, and GI tract extranodal sites were also visualized. PET was insensitive for detecting marrow involvement (identifying 43% of patients with histologically confirmed disease), a consistent finding in indolent NHL.\textsuperscript{19}

Based on six studies, we calculated a pooled proportion of FL patients in whom stage would be altered if PET-based imaging were employed instead of CT.\textsuperscript{12-17} A weighted average was calculated, based on data from a total of 252 patients, using the Freeman-Tukey transformation (arcsine square root transformation).\textsuperscript{20} This analysis indicates that the estimated proportion of FL patients whose stage is altered by PET-based staging is 19%, with a 95% confidence interval of 14-23%. The increased accuracy of PET-CT staging may hold most clinical relevance in the management of early stage FL. Exclusion of occult, distant disease using PET-CT—as was observed historical FL cohorts staged with laparotomy—may translate to improved disease control and survival rates for such patients.\textsuperscript{18,21} Even when stage designation is unchanged, PET-CT may assist in defining margins of the radiation field. In modern practice and in clinical trials, routine application of PET-CT staging in FL is likely to cause stage migration, introducing bias in survival outcomes. Overall, the recommendation for routine use of PET-CT for FL staging is tempered by limitations of available data, which are generally retrospective, and lack routine histologic confirmation of suspected distant sites or long term follow-up. False positive results with PET are well described, and occur with normal physiologic processes, additional malignancy, and inflammatory and benign lesions.\textsuperscript{22} Thus, single PET imaging findings that may influence management should be confirmed with biopsy.\textsuperscript{23}
PET-CT and Prognosis in FL

Since PET-CT regularly identifies additional disease sites in FL, some impact on the Follicular lymphoma International Prognostic Index (FLIPI)—incorporating number of nodal sites and disease stage based on conventional imaging—would be anticipated.\textsuperscript{24} In a descriptive study of PET-CT staging in FL patients at NCCN centers, no difference in FLIPI distribution was seen between those staged with PET and without.\textsuperscript{6} However, in an analysis of a group of patients enrolled on a prospective trial, who had a FLIPI score calculated using both CT and PET-CT at staging, Luminari and colleagues found that PET-CT resulted in a different FLIPI risk group in 24% of patients.\textsuperscript{14} In 2011, Le Dortz and colleagues showed that bone uptake and presence of 6 or more nodal sites on staging PET imaging predicted poor outcomes following chemoimmunotherapy.\textsuperscript{17} In that study, a PET-based prognostic score was developed, but has yet to be validated in prospective trials. Given the sensitivity of PET-CT, the value of the FLIPI index must be reassessed, and new prognostic models incorporating number, intensity, and location of FDG-avid sites should be explored. Interestingly, Abou-Nassar and colleagues found that patients undergoing PET staging at NCCN centers were treated earlier, and more frequently with an anthracycline, but the significance of this observation, derived from an uncontrolled setting, is unclear.\textsuperscript{6} Alteration of content or timing of first-line therapy for FL based on PET-CT findings alone cannot be recommended based on available data.

PET-CT in Evaluation of Histologic Transformation (HT)

The presence of histologic transformation (or discordant presentation, used synonymously in this review) carries implications for prognosis and first-line therapy. Biopsy evidence of HT requires consideration of anthracycline-based therapy, and is predicted by clinical factors including elevated lactate dehydrogenase, poor performance status, and adverse risk group according to standard prognostic models.\textsuperscript{25,26} Early reports of PET imaging in NHL noted higher SUVs in aggressive NHL than in indolent forms, though with wide variation and overlap.\textsuperscript{27-29} In 2005, Schoder and colleagues confirmed that in 97 NHL patients, SUVs are lower in indolent lymphomas, and that a maximum SUV (SUVmax) > 10 at a given biopsy site was 81% specific for an aggressive histology.\textsuperscript{30} Subsequently, investigators have described SUVmax of biopsy-proven HT sites, as well as the highest SUVmax on a given scan, and its variation between nodal sites, as predictors of HT.\textsuperscript{31-34}

In 2008, Bodet-Millien identified 38 indolent NHL patients with clinical or laboratory signs of HT, and performed a prospective study using PET-CT imaging to guide biopsies which were performed at sites with highest SUVmax.\textsuperscript{31} Seventeen patients were diagnosed with HT by biopsy (45%), with median SUVmax of 18.5 (range 11.7-41.2) compared to 8.6 (range 1.7-17.0) in non-transformed cases. All patients with SUVmax > 17 showed HT on biopsy of that site. Using a cutoff SUVmax of 14— in this group with clinical risk factors for transformation—the positive predictive value for HT was 94%.

Noy and colleagues reported a group of patients with indolent NHL who developed biopsy-proven HT, and underwent PET imaging at time of transformation.\textsuperscript{32} Of 33 patients with evaluable data, the mean SUVmax at the site of HT was 14, and ranged from 3-38 (standard deviation 8.7, calculated for this
review). Among 12 patients with available, paired PET scans (from both diagnosis of indolent NHL and HT), 8 showed more than a 50% increase in highest scan SUVmax at time of HT.

Karam and colleagues reported PET-CT findings in 29 patients with HT, compared to 40 patients with indolent NHL.\(^3\)\(^3\) PET-CT and biopsy were not performed in a standard manner; the study reflected heterogeneous clinical practice. The mean highest SUVmax was 20.4 (standard deviation 9.5) in HT, but 6.5 (standard deviation 4.4) in indolent histologies. The authors describe significantly higher SUV in areas of HT (11.8) than indolent NHL (SUV 2.3), in a subset of patients in whom 2 biopsies were performed. Based on a small sample of patients and without adjustment for clinical risk factors, the authors conclude that a 3-fold higher SUV (in a given scan, or increasing over time) should warrant suspicion for HT.

Blasé and colleagues report staging PET findings of 88 indolent lymphoma patients, 5 of whom were diagnosed with HT at a median of 8 months later.\(^3\)\(^4\) The odds ratio for developing HT was 1.25 (95% CI 1.024-1.513) for each unit of SUV max, and remained elevated when corrected for LDH. Baseline SUV max ranged from 4.2-19.6 but PET imaging was not repeated at time of HT. In 4 of the 5 cases, the site of highest SUVmax on staging PET-CT was used to direct biopsy and successfully identified HT.

Limitations of these data are several-fold and warrant caution in applying these findings to routine clinical practice. Standard PET-CT acquisition and interpretation criteria were not applied, and SUV measurements are known for variability and limited reproducibility.\(^3\)\(^5\) With the exception of the study by Bodet-Millin\(^3\)\(^1\) data are retrospective, include varying indolent NHL subtypes, and do not report or adjust for known clinical risk factors for HT. While SUV cutoffs of 10, 14, and 17 have been proposed to signify high likelihood of HT, the standard deviation of highest SUV values observed in biopsy-proven HT is wide—and a significant proportion of HT (45% in the study by Noy\(^3\)\(^2\)) is associated with SUV 10 or under.

Finally, in initial staging of FL and in absence of risk factors for HT\(^2\)\(^5\)\(^,\)\(^2\)\(^6\) the overall prevalence of HT (or discordant, aggressive NHL) is likely to be low. Even with reasonable specificity at a given SUV cutoff (such as 10), the positive predictive value of PET-CT for detecting true HT will be limited. Therefore, over-reliance on SUV values in asymptomatic or low-risk FL patients at initial staging may expose patients to unnecessary biopsies and excess risk. Until more data is available, clinical factors should drive suspicion for HT, which may subsequently then be confirmed by biopsy.

**PET-CT for Response Evaluation in FL**

While survival rates in FL have improved in the last two decades, a continual pattern of relapse is observed.\(^3\)\(^6\)\(^,\)\(^3\)\(^7\) Consolidation of first remission using high-dose chemotherapy and autologous stem cell transplant (HDT), and maintenance using scheduled rituximab, improve disease control but not overall survival.\(^3\)\(^8\)\(^,\)\(^3\)\(^9\) Nonetheless, certain high-risk subgroups may achieve greater benefit with post-induction therapy, including transplant, rituximab maintenance, or incorporation of novel agents, and comprise a high priority group for inclusion in prospective clinical trials.

The potential for PET-CT to identify FL subsets at high risk of relapse following rituximab-containing chemomunotherapy has been consistently observed in recent reports. Analyses of FL patients
enrolled in the PRIMA\textsuperscript{39} and Foll-05\textsuperscript{40} studies, who had PET performed off-study and interpreted locally within 3 months of completing therapy, found that a positive post-treatment PET was seen in about 25% of patients and predicted poor progression-free survival.\textsuperscript{41,42} In both studies, more than half of patients previously classified as having a partial response by CT-based imaging were reclassified to complete response by PET, and PET-defined CR was a more powerful prognostic indicator than the FLIPI score. A prospective study by Dupuis and colleagues study enrolled 121 high tumor burden FL patients from 2007-2009, performing PET imaging before RCHOP, after 4 cycles, and post-treatment and outcomes, with the primary objective to determine PFS according to PET findings.\textsuperscript{43} Maintenance therapy was not given. Standard PET acquisition parameters and central review according to the 5-point, visual Deauville scale\textsuperscript{44} were undertaken; the best discrimination and inter-observer concordance was achieved using a Deauville cutoff of 3 or lower to define a negative PET scan. At 23 months of follow-up, a negative post-treatment PET was seen in 76% and predicted a superior 2 year PFS (87% vs 51% for PET positive, \textit{p}<.001) and OS (100% vs 88% for PET positive, \textit{p}=.01). The predictive value of interim PET was less powerful, and similar to retrospective studies, post-treatment PET was a more powerful predictor than FLIPI score on multivariate analysis. A recent pooled analysis of these three studies\textsuperscript{41-43} conducted independent analysis of available PET scans using the Deauville scale, showing that patients with a positive PET (score of 4 or 5) following completion of therapy (occurring in 17%) had a poor PFS (23% at 4 years vs. 63% for PET-negative).\textsuperscript{45}

The prognostic role of PET-CT has yet to be formally compared with that of minimal residual disease (MRD) in FL. MRD detected by polymerase chain-reaction (PCR) — and more recently, next-generation sequencing, to identify and monitor tumor-specific DNA — predicts relapse in FL and aggressive NHL.\textsuperscript{46-50} Therefore, while a positive PET-CT predicts poor PFS in high-tumor burden FL after chemoimmunotherapy, how to ameliorate this course—and the value of PET with respect to molecular techniques for MRD monitoring— remain to be defined by prospective trials.

**Case Conclusion, Summary, and General Recommendations**

An additional biopsy of the 6 cm mass with SUV 22 was undertaken, showing diffuse large B cell lymphoma. Bone marrow aspirate and biopsy showed 30% involvement by grade 1-2 follicular lymphoma. Following complete pretreatment evaluation, anthracycline chemoinmunotherapy was initiated.

In light of evolving data, national trends, and recent clinical guidelines, PET-CT is poised for increasing integration into routine FL management.\textsuperscript{3,4,6-8} Nonetheless, prospective data defining its role is scant. Retrospective studies suggest that PET-CT increases accuracy of initial staging, with implications for patients under consideration for localized radiotherapy; routine use in such patients is recommended. (Grade 1C) Evaluation for transformation should not be determined solely by PET-CT results including SUV, but should incorporate known risk factors for transformation, to limit the risk of false discovery and unnecessary biopsies. PET-CT can be used to direct the site of biopsy, in FL patients with existing clinical risk factors for HT. (Grade 1B) The sensitivity of PET-CT is relatively low for bone marrow involvement. While PET-CT does not impact FLIPI prognostic group in most patients, the relevance of the FLIPI requires reassessment, and functional imaging may offer novel prognostic information that is
best defined in the context of prospective trials. After standard chemoimmunotherapy treatment, PET discriminates prognosis among high tumor-burden FL patients when visual interpretation methods are employed. Nonetheless, the role of post-treatment PET-CT alongside emerging MRD assays, and its potential to meaningfully inform surveillance or treatment decisions, remain to be defined. Thus, while PET-CT is useful for specific purposes in FL, clinical judgment, use of standardized acquisition and interpretation methods and judicious use of confirmatory biopsy are required. With a shift toward novel, non-cytotoxic treatments for FL, the role of PET-CT in FL is likely to require continual reassessment.

Acknowledgments

Authorship

Contribution: S.D.S. and K.D. developed the concept for this review. S.D.S. conducted the literature search, and wrote the first draft. M.R. designed and performed the pooled analysis. All authors provided input and critical review of the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Stephen D. Smith, Seattle Cancer Care Alliance, Box G3-200, Seattle, CA 98109; email ssmith50@u.washington.edu.

References


FDG PET-CT in follicular lymphoma: a case-based evidence review

Stephen D. Smith, Mary Redman and Kieron Dunleavy