Role of Imaging in the Staging and Response Assessment of Lymphoma: Consensus of the International Conference on Malignant Lymphomas Imaging Working Group


See accompanying article on page 3059

ABSTRACT

Purpose
Recent advances in imaging, use of prognostic indices, and molecular profiling techniques have the potential to improve disease characterization and outcomes in lymphoma. International trials are under way to test image-based response-adapted treatment guided by early interim positron emission tomography (PET)–computed tomography (CT). Progress in imaging is influencing trial design and affecting clinical practice. In particular, a five-point scale to grade response using PET-CT, which can be adapted to suit requirements for early- and late-response assessment with good interobserver agreement, is becoming widely used both in practice- and response-adapted trials. A workshop held at the 11th International Conference on Malignant Lymphomas (ICML) in 2011 concluded that revision to current staging and response criteria was timely.

Methods
An imaging working group composed of representatives from major international cooperative groups was asked to review the literature, share knowledge about research in progress, and identify key areas for research pertaining to imaging and lymphoma.

Results
A working paper was circulated for comment and presented at the Fourth International Workshop on PET in Lymphoma in Menton, France, and the 12th ICML in Lugano, Switzerland, to update the International Harmonisation Project guidance regarding PET. Recommendations were made to optimize the use of PET-CT in staging and response assessment of lymphoma, including qualitative and quantitative methods.

Conclusion
This article comprises the consensus reached to update guidance on the use of PET-CT for staging and response assessment for [18F]fluorodeoxyglucose-avid lymphomas in clinical practice and late-phase trials.

J Clin Oncol 32:3048-3058. © 2014 by American Society of Clinical Oncology

INTRODUCTION

Advances in staging and response assessment of lymphomas have occurred with the introduction of prognostic indices,1-4 molecular profiling,5 and more accurate imaging,6 with the potential to improve disease characterization and treatment selection. The International Harmonisation Project (IHP) first published guidelines about the application of positron emission tomography (PET) using [18F]fluorodeoxyglucose (FDG) in lymphoma in 2007, and PET was integrated in revised response criteria.7

The field has continued to evolve. PET combined with computed tomography (CT) has replaced PET alone. Mounting evidence supports the central role of PET-CT in staging8-18 and response assessment in Hodgkin (HL)19-27 and non-Hodgkin lymphomas (NHL).28-34 Multiple international studies are under way to investigate whether PET-CT response can be used to guide therapy to improve patient outcomes.35,36 Concerted efforts have been made to standardize PET-CT methods37-41 and interpretation in the context of trials.42 A five-point scale (5-PS), suited to assess differing degrees of response at mid- and end of treatment, has been developed to score images.43 This scale was recommended as the standard reporting tool at the First International Workshop on PET in Lymphoma in Deauville, France, in 2009.
TABLE 1. Summary of Recommendations

<table>
<thead>
<tr>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 1: Interpretation of PET-CT scans</td>
</tr>
<tr>
<td>1. Staging of FDG-avid lymphomas is recommended using visual assessment, with PET-CT images scaled to fixed SUV display and color table; focal uptake in HL and aggressive NHL is sensitive for bone marrow involvement and may obviate need for biopsy; MRI is modality of choice for suspected CNS lymphoma (type 1)</td>
</tr>
<tr>
<td>2. Five-point scale is recommended for reporting PET-CT; results should be interpreted in context of anticipated prognosis, clinical findings, and other markers of response; scores 1 and 2 represent CMR; score 3 also probably represents CMR in patients receiving standard treatment (type 1)</td>
</tr>
<tr>
<td>3. Score 4 or 5 with reduced uptake from baseline likely represents partial metabolic response, but at end of treatment represents residual metabolic disease; increase in FDG uptake to score 5, score 5 with no decrease in uptake, and new FDG-avid foci consistent with lymphoma represent treatment failure and/or progression (type 2)</td>
</tr>
<tr>
<td>Section 2: Role of PET-CT for staging</td>
</tr>
<tr>
<td>1. PET-CT should be used for staging in clinical practice and clinical trials but is not routinely recommended in lymphomas with low FDG avidity; PET-CT may be used to select best site to biopsy (type 1)</td>
</tr>
<tr>
<td>2. Contrast-enhanced CT when used at staging or restaging should ideally occur during single visit combined with PET-CT, if not already performed; baseline findings will determine whether contrast-enhanced PET-CT or lower-dose unenhanced PET-CT will suffice for additional imaging examinations (type 2)</td>
</tr>
<tr>
<td>3. Bulk remains an important prognostic factor in some lymphomas; volumetric measurement of tumor bulk and total tumor burden, including methods combining metabolic activity and anatomical size or volume, should be explored as potential prognosticators (type 3)</td>
</tr>
<tr>
<td>Section 3: Role of interim PET</td>
</tr>
<tr>
<td>1. If midtherapy imaging is performed, PET-CT is superior to CT alone to assess early response; trials are evaluating role of PET response–adapted therapy; currently, it is not recommended to change treatment solely on basis of interim PET-CT unless there is clear evidence of progression (type 1)</td>
</tr>
<tr>
<td>2. Standardization of PET methods is mandatory for use of quantitative approaches and desirable for routine clinical practice (type 1)</td>
</tr>
<tr>
<td>3. Data suggest that quantitative measures (eg, 85SUV/max) could be used to improve on visual analysis for response assessment in DLBCL, but this requires further validation in clinical trials (type 2)</td>
</tr>
<tr>
<td>Section 4: Role of PET at end of treatment</td>
</tr>
<tr>
<td>1. PET-CT is standard of care for remission assessment in FDG-avid lymphoma; in presence of residual metabolically active tissue, where salvage treatment is being considered, biopsy is recommended (type 1)</td>
</tr>
<tr>
<td>2. Investigation of significance of PET-negative residual masses should be collected prospectively in clinical trials; residual mass size and location should be recorded on end-of-treatment PET-CT reports where possible (type 3)</td>
</tr>
<tr>
<td>3. Emerging data support use of PET-CT after rituximab-containing chemotherapy in high–tumor burden FL; studies are warranted to confirm this finding in patients receiving maintenance therapy (type 2)</td>
</tr>
<tr>
<td>4. Assessment with PET-CT could be used to guide decisions before high-dose chemotherapy and ASCT, but additional studies are warranted (type 3)</td>
</tr>
</tbody>
</table>

Abbreviations: ASCT, autologous stem-cell transplantation; CMR, complete metabolic response; CT, computed tomography; DLBCL, diffuse large B-cell lymphoma; FDG, [18F]fluorodeoxyglucose; FL, follicular lymphoma; HL, Hodgkin lymphoma; MRI, magnetic resonance imaging; NHL, non-Hodgkin lymphoma; PET, positron emission tomography; SUV, standardized uptake value; 85SUV/max, change in maximum SUV.

and these so-called Deauville criteria have been widely applied in trials in preference to earlier criteria.44-49 Quantitative applications of FDG-PET are also recognized as objective tools for response monitoring,50 although accurate measurement relies on consistent methods for acquisition and processing and rigorous quality assurance of equipment for widespread application.39,51-54

In response to changing requirements for PET-CT, to accommodate assessments at staging and during and after treatment, especially for response-adapted trials, a workshop was convened at the International Conference on Malignant Lymphoma (ICML) in 2011, attended by representatives from major cooperative groups. ICML working groups were established to update guidelines. The imaging group reported to colleagues at follow-up workshops at the Fourth International Workshop on PET in Lymphoma in Menton, France, in 2012 and the 12th ICML in Lugano, Switzerland, in 2013. This article represents the consensus reached regarding the use of PET-CT in lymphoma in clinical practice and late-phase trials.
which is distinguished from physiologic uptake and other patterns of disease with increased FDG uptake including infection and inflammation, according to distribution and/or CT characteristics.

Focal FDG uptake within the bone or bone marrow, liver, and spleen is highly sensitive for involvement in HL and aggressive NHL and may obviate the need for bone marrow biopsy. Diffuse increased uptake may occur with abnormal focal uptake, but in HL, diffuse uptake without focal activity often represents reactive hyperplasia and should not be confused with lymphomatous involvement. PET-CT can miss low-volume involvement, typically 20% of the marrow, and coexistent low-grade lymphoma in NHL and may obviate the need for bone marrow biopsy. PET scans are best reported using a fixed display and color table. Focal uptake in HL and aggressive NHL is sensitive for bone marrow involvement and may obviate the need for biopsy. MRI is the modality of choice for suspected CNS lymphoma (type 1).

Resolution of uptake at sites of initial disease indicates metabolic response. Reduction of uptake may also indicate satisfactory response, but the degree of uptake that is indicative of response is dependent on the timing of the scan during treatment and the clinical context, including prognosis, lymphoma subtype, and treatment regimen. The availability of a baseline scan is considered optimal for the accuracy of subsequent response assessment. The IHP criteria specified that uptake should be ≤ the mediastinal blood pool for lesions ≥ 2 cm or the adjacent background for smaller lesions to define metabolic response at the end of treatment. In early-response assessment, treatment is incomplete, so the emphasis is on the degree of response and a continuous or close-to-continuous scale is desirable rather than positive or negative response categories. Early attempts to address this used three response groups (ie, negative, minimal residual uptake, and positive). Further refinement led to the development of the 5-PS, which better represents different grades of uptake.

The 5-PS was intended as a simple, reproducible scoring method, with the flexibility to change the threshold between good or poor response according to the clinical context and/or treatment strategy. For example, a lower level of FDG uptake might be preferred to define a so-called negative result in a clinical trial exploring de-escalation to avoid undertreatment. A higher level of uptake might be preferred to define a so-called positive result in a trial exploring escalation to avoid overtreatment. The 5-PS has been validated for use at interim and the end of treatment and was adopted as the preferred reporting method at the First International Workshop on PET in Lymphoma in Deauville, France (ie, Deauville criteria), and in several international trials.

The 5-PS scores the most intense uptake in a site of initial disease, if present, as follows:

- 1. No uptake
- 2. Uptake ≤ mediastinum
- 3. Uptake > mediastinum but ≤ liver
- 4. Uptake moderately higher than liver
- 5. Uptake markedly higher than liver and/or new lesions
- X. New areas of uptake unlikely to be related to lymphoma

Good interobserver agreement has been reported in HL using scores 1, 2, and 3 to define CMR after two ABVD cycles reported using scores 1, 2, and 3 for 3-year PFS. The UK RAPID (Response Adapted Therapy Using Positron Emission Tomography in Early-Stage Hodgkin Lymphoma) study used the 5-PS in patients with early HL. iPET remained an independent predictor of 3-year progression-free survival (PFS) on multivariable analysis, despite use of a response-adapted design. Conservative scoring was used, with a score of 1 or 2 regarded as complete metabolic response (CMR); patients with CMR after three cycles of ABVD (doxorubicin, bleomycin, vinblastine, and dacarbazine) were randomly assigned to radiotherapy (RT) or no further treatment. Retrospective analysis of an international cohort of 260 patients with advanced HL using scores 1, 2, and 3 to define CMR after two ABVD cycles reported a negative predictive value (NPV) of 94% and positive predictive value (PPV) of 73% for 3-year PFS. iPET and end-of-treatment PET using scores 1, 2, and 3 for CMR were both independent predictors of 2-year

<table>
<thead>
<tr>
<th>Table 2. FDG Avidity According to WHO Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Histology</strong></td>
</tr>
<tr>
<td>HL</td>
</tr>
<tr>
<td>DLBCL</td>
</tr>
<tr>
<td>FL</td>
</tr>
<tr>
<td>Mantle-cell lymphoma</td>
</tr>
<tr>
<td>Burkitt’s lymphoma</td>
</tr>
<tr>
<td>Marginal zone lymphoma, nodal</td>
</tr>
<tr>
<td>Lymphoblastic lymphoma</td>
</tr>
<tr>
<td>Anaplastic large T-cell lymphoma</td>
</tr>
<tr>
<td>NK/T-cell lymphoma</td>
</tr>
<tr>
<td>Angioimmunoblastic T-cell lymphoma</td>
</tr>
<tr>
<td>Peripheral T-cell lymphoma</td>
</tr>
<tr>
<td>MALT marginal zone lymphoma</td>
</tr>
<tr>
<td>Small lymphocytic lymphoma</td>
</tr>
<tr>
<td>Enteropathy-type T-cell lymphoma</td>
</tr>
<tr>
<td>Marginal zone lymphoma, splenic</td>
</tr>
<tr>
<td>Marginal zone lymphoma, unspecified</td>
</tr>
<tr>
<td>Mycosis fungoides</td>
</tr>
<tr>
<td>Sezary syndrome</td>
</tr>
<tr>
<td>Primary cutaneous anaplastic large T-cell lymphoma</td>
</tr>
<tr>
<td>Lymphomatoid papulosis</td>
</tr>
<tr>
<td>Subcutaneous panniculitis-like T-cell lymphoma</td>
</tr>
<tr>
<td>Cutaneous B-cell lymphoma</td>
</tr>
</tbody>
</table>

NOTE. Data adapted, with additional updates. Abbreviations: DLBCL, diffuse large B-cell lymphoma; FDG, ([18F]fluorodeoxyglucose; FL, follicular lymphoma; HL, Hodgkin lymphoma; MALT, mucosa-associated lymphoid tissue; NK, natural killer. Only 27% of cutaneous sites. Only 62% of cutaneous sites.
PF5 in a recent prospective study in FL34. Other studies in HL and NHL have reported that increasing the threshold to define CMR improved the PPV while maintaining a high NPV.28,34,91,92,94

Scores 1 and 2 are therefore considered to represent CMR. Score 3 also likely represents CMR at intermediate5 and good prognosis at completion of standard treatment.34,94,99 However, in trials where de-escalation is based on PET response, it may be preferable to consider score 3 as inadequate response to avoid undertreatment.42

Recommendation. The 5-PS is recommended for reporting PET-CT. Results should be interpreted in the context of the anticipated prognosis, clinical findings, and other markers of response. Scores 1 and 2 represent CMR. Score 3 also probably represents CMR in patients receiving standard treatment (type 1).

The terms moderately and markedly were not defined initially, because there were insufficient data to define scores quantitatively.43 Meanwhile, it is suggested according to published data25,34,106 that score 4 be applied to uptake > the maximum SUV in a large region of normal liver and score 5 to uptake 2× to 3× > the maximum SUV in the liver. It is acknowledged that mean liver SUV may be less influenced by image noise than maximum SUV, but reproducibility is more dependent on standardizing the location and size of the region of interest.101 Work is ongoing to assess optimal tumor and liver metrics.102 The liver is also affected by insulin levels, and patient preparation is important with respect to fasting and timing of insulin administration in diabetics.103 It is recognized that in Waldeyer’s ring or extranodal sites with high physiologic uptake or with activation within spleen or marrow (eg, with chemotherapy or granulocyte colony-stimulating factor [GCSF]), FDG uptake may be > normal mediastinum and/or liver. In this circumstance, CMR may be inferred if uptake at sites of initial involvement is < surrounding normal tissue, even if the tissue has high physiologic uptake.

Recommendation. Scores 4 and 5 with reduced uptake from baseline likely represent partial metabolic response, but at the end of treatment, they represent residual metabolic disease. An increase in FDG uptake to a score of 5, score 5 with no decrease in uptake, and new FDG-avid foci consistent with lymphoma represent treatment failure and/or progression (type 2).

Non-specific FDG uptake may occur with treatment-related inflammation. Patients should be scanned as long after the previous chemotherapy administration as possible for interim assessment. A minimum of 3 weeks, but preferably 6 to 8 weeks, after completion of the last chemotherapy cycle,7 2 weeks after GCSF treatment, or 3 months after RT is recommended.39

Role of PET-CT for Staging

Previous clinical trials have used the Ann Arbor staging system to select patients and report outcomes.104 Currently, prognostic indices are mostly used to risk stratify patients at diagnosis to inform therapy, but most include stage as a factor,1-3,105 so imaging-determined stage remains relevant.

PET-CT using FDG is more accurate than CT for staging in HL9,10,106-111 and NHL11-13,18,112,113 with increased sensitivity, particularly for extranodal disease.6 Upstaging occurs more often than downstaging, with management alterations in some patients (Table 3). Management change after upstaging is more common in FL14,15 than other lymphomas, especially for patients with limited disease on CT.16,17

The intensity of FDG uptake is higher in aggressive than indolent lymphomas, and FDG PET-CT may be used to target biopsy in patients with suspected transformation.65,114,115

Recommendation. PET-CT should be used for staging in clinical practice and clinical trials, but it is not routinely recommended in

### Table 3. Studies Comparing PET or PET-CT With CT Alone for Staging of Lymphomas

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>PET or PET-CT</th>
<th>No. of Patients</th>
<th>Disease</th>
<th>Upstaging (%)</th>
<th>Downstaging (%)</th>
<th>Management Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangerter et al106</td>
<td>1998</td>
<td>PET</td>
<td>44</td>
<td>HL</td>
<td>12</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>Partridge et al107</td>
<td>2000</td>
<td>PET</td>
<td>44</td>
<td>HL</td>
<td>41</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>Jerusalem et al108</td>
<td>2001</td>
<td>PET</td>
<td>33</td>
<td>HL</td>
<td>10</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Weihrauch et al109</td>
<td>2002</td>
<td>PET</td>
<td>22</td>
<td>HL</td>
<td>18</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Munker et al110</td>
<td>2004</td>
<td>PET</td>
<td>73</td>
<td>HL</td>
<td>29</td>
<td>3</td>
<td>NS</td>
</tr>
<tr>
<td>Naumann et al111</td>
<td>2004</td>
<td>PET</td>
<td>88</td>
<td>HL</td>
<td>13</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Hutchings et al10</td>
<td>2006</td>
<td>Mostly PET-CT</td>
<td>99</td>
<td>HL</td>
<td>19</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Rigacci et al10</td>
<td>2007</td>
<td>Mostly PET</td>
<td>186</td>
<td>HL</td>
<td>14</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Buchmann et al112</td>
<td>2001</td>
<td>PET</td>
<td>52</td>
<td>HL (n = 27), NHL (n = 25)</td>
<td>8</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Wirth et al113</td>
<td>2002</td>
<td>PET</td>
<td>50</td>
<td>HL (n = 19), NHL (n = 31)</td>
<td>14</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Rasani et al114</td>
<td>2006</td>
<td>PET-CT</td>
<td>103</td>
<td>HL (n = 32), NHL (n = 68)</td>
<td>31</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>Elstrom et al115</td>
<td>2008</td>
<td>PET-CT</td>
<td>61</td>
<td>HL and NHL</td>
<td>18</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Pelosi et al116</td>
<td>2008</td>
<td>PET</td>
<td>65</td>
<td>HL (n = 30), NHL (n = 35)</td>
<td>11</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Karm et al117</td>
<td>2006</td>
<td>PET</td>
<td>17</td>
<td>FL</td>
<td>41</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>Janikova et al118</td>
<td>2008</td>
<td>Mostly PET</td>
<td>82</td>
<td>FL</td>
<td>NS</td>
<td>NS</td>
<td>18</td>
</tr>
<tr>
<td>Wirth et al119</td>
<td>2008</td>
<td>PET</td>
<td>42</td>
<td>FL stages I-II on CT</td>
<td>29</td>
<td>0</td>
<td>45</td>
</tr>
<tr>
<td>Le Dortz et al119</td>
<td>2010</td>
<td>PET-CT</td>
<td>45</td>
<td>FL</td>
<td>8</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Luminari et al118</td>
<td>2013</td>
<td>PET-CT</td>
<td>142</td>
<td>FL</td>
<td>11</td>
<td>1</td>
<td>NS</td>
</tr>
</tbody>
</table>

Abbreviations: CT, computed tomography; FL, follicular lymphoma; HL, Hodgkin lymphoma; NHL, non-Hodgkin lymphoma; NS, not stated; PET, positron emission tomography.

*False negative.
lymphomas with low FDG avidity. PET-CT may be used to select the best site to biopsy (type 1).

**Role of Contrast-Enhanced CT**

The CT part of a PET-CT scan may be performed with contrast enhancement (ceCT) at full dose to obtain a high-quality CT examination or without contrast using a lower dose. Lower-dose CT is used to correct for the attenuation of radioactivity within the patient and to localize abnormalities seen on PET, with less radiation than a full diagnostic examination. Whichever protocol is used, CT must be acquired during shallow breathing or end of expiration to avoid misregistration and artifacts.

Direct comparison of unenhanced lower-dose PET-CT and cePET-CT suggests management is rarely altered by ceCT, although ceCT may identify additional findings and improve detection of abdominal or pelvic disease. However, full-dose ceCT involves additional radiation, which should be considered when deciding which examination to perform. ceCT is desirable for RT planning performed in the treatment position and is required for accurate nodal measurements for trial purposes.

Small errors in the measurement of FDG uptake in tumor may occur with contrast media because of an effect on attenuation correction; these errors are unlikely to be clinically important. Contrast may cause errors in comparison of uptake between tumor and reference sites by causing FDG uptake to be overestimated in the mediastinum and liver by 10% to 15%. Several organizations (eg, European Association Nuclear Medicine, Society Nuclear Medicine, and Radiological Society North America) recommend that a low-dose CT scan with normal breathing be performed before a PET scan, followed by full diagnostic high-dose ceCT with repositioning of the arms and breath hold, if quantitative measures and ceCT are required.

In practice, many patients undergo separate ceCT before PET-CT. If baseline ceCT demonstrates no additional relevant findings, lower-dose CT during PET-CT examination will be sufficient for response assessment.

**Recommendation.** ceCT when used at staging or restaging should ideally occur during a single visit in combination with PET-CT, if not already performed. The baseline findings will determine whether cePET-CT or lower-dose unenhanced PET-CT will suffice for additional imaging examinations (type 2).

**Relevance of Initial Disease Bulk**

The presence of bulky disease is a negative prognostic factor in some lymphomas. Bulk is considered an adverse factor in early-stage HL but not in advanced HL. In DLBCL, bulk is predictive of inferior survival in favorable-prognosis disease but not in poor-prognosis disease, probably because its influence is supersedes by other factors reflecting disease burden. The longest diameter of the largest involved node is included in the FL International Prognostic Index. Unidimensional measurements are used for bulk, but these do not assess total tumor burden. Newer methods of contouring are being developed for CT and PET to measure the total tumor volume. The prognostic value of these methods remains to be evaluated.

**Recommendation.** Bulk remains an important prognostic factor in some lymphomas. Volumetric measurement of tumor bulk and total tumor burden, including methods combining metabolic activity and anatomic size or volume, should be explored as potential prognosticators (type 3).

**Role of iPET**

Interim imaging is frequently performed in clinical practice and trials and is recommended by some international guidelines. The purpose is to ensure the effectiveness of treatment and exclude the possibility of progression. PET-CT shows metabolic response earlier than anatomic response and has the potential to replace CT. Studies have shown that iPET is a strong prognostic indicator in HL and aggressive NHL, outperforming the International Prognostic Score and International Prognostic Index. These findings highlight the potential of using iPET to tailor treatment according to individual response. However, it is important to emphasize that there is no conclusive evidence that changing treatment according to iPET improves outcome, a question currently being addressed in clinical trials worldwide.

There is a preponderance of data reporting the predictive value of iPET, most often after two cycles in HL (Appendix Table A1, online only). In DLBCL, early indication of poor response is especially important because salvage treatment of progressive or relapsed disease is less effective in the rituximab era. However, although early data favored iPET, more recent data have suggested iPET is less predictive for response with immunochemotherapy (Appendix Table A2, online only), and end-of-treatment PET is a better predictor.

Visual assessment with iPET in HL results in consistently high NPV, with ≥2-year PFS of approximately 95%, and acceptable PPV, with PFS between 13% and 27%, for advanced disease treated with ABVD. Initial reports using visual analysis for iPET in DLBCL were favorable, but more recent studies have demonstrated good NPV, with ≥2-year PFS rates of 73% to 86% for patients with so-called negative scans, but more variable PPV. PPV for PET-positive patients in recent studies has ranged from 18% to 74%. The drop in PPV may be related to improved outcomes with rituximab or better supportive care or may possibly occur because so-called false-positive metabolic activity is more frequent with immunotherapy. A different cutoff or combination of factors may be required for modern management of DLBCL.

**Recommendation.** If midtherapy imaging is performed, PET-CT is superior to CT alone to assess early response. Trials are evaluating the role of PET response–adapted therapy. Currently, changing treatment solely on the basis of iPET-CT is not recommended, unless there is clear evidence of progression (type 1).

The use of quantitation to improve on visual assessment has been explored in DLBCL. Change in the maximum SUV (SUVmax) in tumor before and after treatment has been evaluated as a measure of response. Receiver operator curve analysis in 92 patients with DLBCL scanned after two cycles and 80 patients scanned after four identified optimum thresholds for percentage change in SUVmax for predicting event-free survival (EFS). A retrospective analysis applied to a trial where treatment was adapted according to visual assessment with iPET reported that SUVmax at two and four cycles was predictive of PFS, whereas visual analysis was not. Other groups have also reported that SUVmax predicts response, but with thresholds ranging from 66% to 91%, suggesting that consistency in scanning protocols, matching conditions for serial scans, and proper calibration and scanner maintenance are mandatory for
general application. The optimum cutoff is also likely influenced by timing, with a tendency for a higher cutoff later during treatment. Although the goal of quantitation is more objective assessment, it remains necessary to integrate with clinical information to exclude confounding variables.

The $\Delta$SUVmax analysis is being prospectively applied in the PETAL (Positron Emission Tomography Guided Therapy of Aggressive Non-Hodgkin’s Lymphomas) and GAINED (GA in Newly Diagnosed Diffuse Large B Cell Lymphoma) studies exploring response-adapted treatment with immunochemotherapy. Combining $\Delta$SUVmax with CT metrics in early nonbulky HL and with age-adjusted International Prognostic Index in DLBCL has been reported to improve response prediction. Another measure proposed is SUVpeak, a 1-cm$^3$ volume containing the hottest area of tumor, which may be less sensitive to noise and resolution and possibly more reproducible. Changes in the metabolic tumor volume (MTV) and total lesion glycolysis (TLG) calculated as MTV × SUVmean are additional exploratory measures. However, preliminary reports have suggested changes in MTV and TLG are not predictive in DLBCL. The results of the UK National Cancer Research Institute PET R-CHOP (rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone) sub-study measuring PFS in 200 patients with DLBCL where clinicians were blinded to iPET are awaited.

This and other studies may provide insight into whether quantitation will improve the performance of iPET in DLBCL.

**Recommendation.** Standardization of PET methods is mandatory for the use of quantitative approaches and desirable for routine clinical practice (type 1). Data suggest that quantitative measures (eg, $\Delta$SUVmax) could be used to improve on visual analysis for response assessment in DLBCL, but this requires further validation in clinical trials (type 2).

### Role of PET at the End of Treatment

End-of-treatment remission assessment is more accurate with PET-CT than CT alone in patients with HL, DLBCL, and high–tumor burden FL (Appendix Table A3, online only). High accuracy for PET-CT has been reported in patients after treatment with ABVD and BEACOPP (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone) for advanced HL. In a study where PET was used to guide RT, patients treated with BEACOPP (but not RT) with a PET-negative positive response had PFS equivalent to that of patients with complete response (CR) or unconfirmed CR. In aggressive NHL, studies involving > 300 patients have reported consistently high NPV of 80% to 100% but more variable PPV of 50% to 100%. In the presence of residual metabolically active tissue, if salvage treatment is being considered, a biopsy may be required. If residual disease is considered unlikely, the scan could be repeated later.

**Recommendation.** PET-CT is the standard of care for remission assessment in FDG-avid lymphoma. For HL and DLBCL, in the presence of residual metabolically active tissue, where salvage treatment is being considered, a biopsy is recommended (type 1).

The significance of a residual mass if CMR is achieved is unclear, with some reports suggesting improved outcomes when CMR is associated with a radiologic CR in HL and DLBCL, whereas others suggest outcomes are unaffected by the presence of a residual mass. It is proposed that the size of the residual mass be recorded where possible, and if relapse occurs, it should be documented whether this occurred within the residual mass.

**Recommendation.** Investigation of the significance of PET-negative residual masses should be collected prospectively in clinical trials. Residual mass size and location should be recorded on end-of-treatment PET-CT reports where possible (type 3).

In FL, PET predicts inferior outcomes in patients with high tumor burden who remain PET positive after first-line immunochemotherapy. Post-treatment PET seems to be a better predictor than iPET. Currently, data are insufficient regarding assessment after maintenance therapy. This suggests a potential role for PET in evaluating new approaches in response-adapted studies in FL after first-line treatment with rituximab-containing chemotherapy.

**Recommendation.** Emerging data support the use of PET-CT after rituximab-containing chemotherapy in high–tumor burden FL. Studies are warranted to confirm this finding in patients receiving maintenance therapy (type 2).

### Assessment Before High-Dose Chemotherapy and Autologous Stem-Cell Transplantation

Various studies have reported that PET-CT using FDG is prognostic in patients with relapsed or refractory HL or DLBCL after salvage chemotherapy before high-dose chemotherapy and autologous stem-cell transplantation (ASCT) and is superior to CT alone. Three-year PFS and EFS rates of 31% to 41% have been reported for patients with PET-positive scans, compared with 75% to 82% for patients with PET-negative scans.

PET may have a role in selecting patients for high-dose chemotherapy and ASCT after salvage treatment and in identifying patients with poor prognosis who could benefit from alternative regimens or consolidation. PET could also be used as a surrogate endpoint to test the addition of novel therapies to current reinduction regimens.

**Recommendation.** Assessment with PET-CT could be used to guide decisions before high-dose chemotherapy and ASCT, but additional studies are warranted (type 3).

### PET-CT in Subtypes Other Than HL, DLBCL, and FL

Small retrospective studies have suggested that post-treatment scans can predict survival in treatment of mantle-cell lymphoma. In primary mediastinal B-cell lymphoma, a recent prospective study reported that 54 (47%) of 115 patients achieved CMR after first-line chemotherapy, and a PET response–adapted approach is currently being tested. However, another study involving 51 patients treated with dose-adjusted EPOCH (etoposide, prednisone, vincristine, cyclophosphamide, doxorubicin) plus rituximab reported that 10 of 15 patients had FDG uptake 6 weeks after treatment, which later diminished or stabilized, suggesting treatment-related inflammation with this regimen. There are limited data regarding T-cell lymphomas, with higher uptake reported in more aggressive subtypes and lower uptake in cutaneous lymphomas. In mycosis fungoides, higher uptake has been reported in the presence of large-cell transformation and extracutaneous disease, which adversely affects prognosis.

There are few data on response assessment; one report in noncutaneous mature natural killer/T-cell lymphoma suggested iPET was predictive of response, whereas another found that neither...
interim nor end-of-treatment PET were predictive. Prospective studies are warranted.

**DISCUSSION**

In response to developments involving PET-CT, recommendations from the ICML imaging group have been made to update practice. These include guidance on reporting of PET-CT for staging and response assessment of HL, DLBCL, and aggressive FL using the 5-PS. PET-CT is recommended for midtreatment assessment in place of CT alone, if imaging is clinically indicated, and for remission assessment. Quantitative imaging parameters for assessing disease burden and response should be explored as potential prognosticators. The standardization of PET-CT methods is mandatory for quantitative analysis and desirable for best clinical practice.

**AUTHORS’ DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST**

Although all authors completed the disclosure declaration, the following author(s) and/or an author’s immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a “*” are those for which no compensation was received; those relationships marked with a “C” were compensated. For a detailed description of the disclosure categories, or for more information about ASCO’s conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

**Employment or Leadership Position:** None Consultant or Advisory Role: None Stock Ownership: None Honoraria: Lawrence H. Schwartz, Pfizer Research Funding: None Expert Testimony: None Patents, Royalties, and Licenses: None Other Remuneration: Lawrence H. Schwartz, Roche: Richard J. Fisher, Johnson & Johnson, Gilead Sciences, Coherus Bioscience

**REFERENCES**


**AUTHOR CONTRIBUTIONS**

Conception and design: Sally F. Barrington, N. George Mikhail, Lale Kostakoglou, Michel Meignan, Martin Hutchings, Stefan P. Mueller, Lawrence H. Schwartz, Emanuele Zucca, Richard I. Fisher, Rodney J. Hicks, Michael J. O’Doherty, Alberto Biggi, Bruce D. Cheson

Collection and assembly of data: Sally F. Barrington, N. George Mikhail, Lale Kostakoglou, Michel Meignan, Martin Hutchings, Stefan P. Mueller, Lawrence H. Schwartz

Data analysis and interpretation: Sally F. Barrington, N. George Mikhail, Lale Kostakoglou, Michel Meignan, Martin Hutchings, Stefan P. Mueller, Lawrence H. Schwartz, Judith Trotman, Otto S. Hoekstra, Roland Hustinx

Manuscript writing: All authors

Final approval of manuscript: All authors
ICML Recommendations for Using PET-CT in Lymphoma


22. Barrington et al


ASCO’s Quality Assessment Tool for Community Research Programs

As part of ASCO’s mission to ensure the highest quality of cancer care and research, the Community Research Forum has developed the ASCO Research Program Quality Assessment Tool. Access the tool today to learn about key components of an internal quality assessment program, conduct a quality assessment of your site, and use the checklist tool, templates, references, and resources.

Visit asco.org/communityresearchforum for more information.
Acknowledgment


We thank R. Boellaard, D. Caballero, L. Ceriani, M. Coronado, F. Cicone, A. Gallamini, M. Gregianin, E. Itti, T.A. Lister, C. Moskowitz, H. Schöder, and J. Zijlstra for their contributions to improve the manuscript; Paul Smith, tumor group lead (hematology and brain trials), Cancer Research United Kingdom and University College London Cancer Trials Centre, and the UK National Cancer Research Institute Lymphoma Clinical Study Group for assistance with funding the imaging task group meeting in London, United Kingdom, in 2012; and the European School of Oncology and European Society for Medical Oncology for their support of the workshops at the International Conference on Malignant Lymphoma in 2011 and 2013.

Appendix

Table A1. Studies Including \( \geq 50 \) Patients With HL Reporting Outcomes According to Visual Assessment With Interim PET

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>No. of Patients</th>
<th>Disease Stage</th>
<th>Chemotherapy</th>
<th>No. of Cycles Before PET</th>
<th>No. PET Negative</th>
<th>PFS/EFS At (years)</th>
<th>PET Negative (%)</th>
<th>PET Positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hutchings et al(^1)</td>
<td>2005</td>
<td>85</td>
<td>I-IV</td>
<td>Mostly ABVD (n = 79)</td>
<td>2-3</td>
<td>72</td>
<td>5</td>
<td>92</td>
<td>39</td>
</tr>
<tr>
<td>Hutchings et al(^2)</td>
<td>2006</td>
<td>77</td>
<td>I-IV</td>
<td>Mostly ABVD (n = 70)</td>
<td>2</td>
<td>61</td>
<td>2</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>Gallamini et al(^3)</td>
<td>2007</td>
<td>260</td>
<td>IIb-IV</td>
<td>Mostly ABVD (n = 249)</td>
<td>2</td>
<td>210</td>
<td>2</td>
<td>96</td>
<td>6</td>
</tr>
<tr>
<td>Markova et al(^4)</td>
<td>2009</td>
<td>50</td>
<td>IIb-IV</td>
<td>BEACOPP</td>
<td>4</td>
<td>36</td>
<td>2</td>
<td>97</td>
<td>86</td>
</tr>
<tr>
<td>Cerci et al(^5)</td>
<td>2010</td>
<td>104</td>
<td>II-IV</td>
<td>ABVD</td>
<td>2</td>
<td>74</td>
<td>3</td>
<td>90</td>
<td>53</td>
</tr>
<tr>
<td>Barnes et al(^6)</td>
<td>2011</td>
<td>96</td>
<td>I-II (nonbulky)</td>
<td>ABVD</td>
<td>2-4</td>
<td>79</td>
<td>4</td>
<td>91</td>
<td>87</td>
</tr>
<tr>
<td>Zinzani et al(^7)</td>
<td>2012</td>
<td>304</td>
<td>I/IIA (n = 147)</td>
<td>ABVD</td>
<td>2</td>
<td>128</td>
<td>9</td>
<td>89</td>
<td>31</td>
</tr>
<tr>
<td>Zinzani et al(^8)</td>
<td>2012</td>
<td>304</td>
<td>I/IIb (n = 157)</td>
<td>ABVD</td>
<td>2</td>
<td>123</td>
<td>9</td>
<td>89</td>
<td>29</td>
</tr>
<tr>
<td>Biggi et al(^9)</td>
<td>2013</td>
<td>260</td>
<td>IIb-IV</td>
<td>ABVD</td>
<td>2</td>
<td>215</td>
<td>3</td>
<td>95</td>
<td>28</td>
</tr>
</tbody>
</table>

Abbreviation: ABVD, doxorubicin, bleomycin, vinblastine, and dacarbazine; BEACOPP, bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; EFS, event-free survival; HL, Hodgkin lymphoma; PET, positron emission tomography; PFS, progression-free survival.

Table A2. Studies Including \( \geq 50 \) Patients With Aggressive NHL Reporting Outcomes According to Visual Assessment With Interim PET

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>No. of Patients</th>
<th>Chemotherapy</th>
<th>No. of Cycles of Therapy</th>
<th>No. PET Negative</th>
<th>PFS/EFS At (years)</th>
<th>PET Negative (%)</th>
<th>PET Positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spaepen et al(^10)</td>
<td>2002</td>
<td>70</td>
<td>Mostly CHOP (n = 56)</td>
<td>3-4</td>
<td>37</td>
<td>2</td>
<td>85</td>
<td>0</td>
</tr>
<tr>
<td>Haioun et al(^11)</td>
<td>2005</td>
<td>90</td>
<td>CHOP or ACVBP/ACE (n = 53) plus rituximab (n = 37)</td>
<td>2</td>
<td>54</td>
<td>2</td>
<td>82</td>
<td>43</td>
</tr>
<tr>
<td>Michaeel et al(^12)</td>
<td>2007</td>
<td>121</td>
<td>Mostly CHOP (n = 97)</td>
<td>2-3</td>
<td>65</td>
<td>5</td>
<td>89</td>
<td>16</td>
</tr>
<tr>
<td>Cashen et al(^13)</td>
<td>2011</td>
<td>50</td>
<td>R-CHOP</td>
<td>2-3</td>
<td>26</td>
<td>2</td>
<td>85</td>
<td>63</td>
</tr>
<tr>
<td>Micallef et al(^14)</td>
<td>2011</td>
<td>76</td>
<td>ER-CHOP</td>
<td>2</td>
<td>60</td>
<td>2</td>
<td>73</td>
<td>60</td>
</tr>
<tr>
<td>Yang et al(^15)</td>
<td>2011</td>
<td>159</td>
<td>R-CHOP</td>
<td>3-4</td>
<td>116</td>
<td>3</td>
<td>86</td>
<td>29</td>
</tr>
<tr>
<td>Yoo et al(^16)</td>
<td>2011</td>
<td>155</td>
<td>R-CHOP</td>
<td>2-4</td>
<td>100</td>
<td>3</td>
<td>84</td>
<td>66</td>
</tr>
<tr>
<td>Zinzani et al(^17)</td>
<td>2011</td>
<td>91</td>
<td>Mostly R-CHOP (n = 66), rituximab (n = 91)</td>
<td>Midtreatment</td>
<td>56</td>
<td>5</td>
<td>75</td>
<td>18</td>
</tr>
<tr>
<td>Safar et al(^18)</td>
<td>2012</td>
<td>112</td>
<td>R-CHOP (n = 81), R-ACVBP (n = 31)</td>
<td>2</td>
<td>70</td>
<td>3</td>
<td>84</td>
<td>47</td>
</tr>
<tr>
<td>Pregno et al(^19)</td>
<td>2012</td>
<td>88</td>
<td>R-CHOP</td>
<td>2-4</td>
<td>66</td>
<td>2</td>
<td>85</td>
<td>72</td>
</tr>
<tr>
<td>Nols et al(^20)</td>
<td>2013</td>
<td>73</td>
<td>R-CHOP (n = 48), R-miniCHOP (n = 8), ACVBP (n = 17), CHOP (n = 1)</td>
<td>3-4</td>
<td>53</td>
<td>2</td>
<td>84</td>
<td>47</td>
</tr>
</tbody>
</table>

Abbreviations: ACE, doxorubicin, cyclophosphamide, and etoposide; ACVBP, doxorubicin, cyclophosphamide, vincristine, bleomycin, and prednisone; CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone; E, etoposide; EFS, event-free survival; NHL, non-Hodgkin lymphoma; PET, positron emission tomography; PFS, progression-free survival.

*Prospective study.
### Table A3. Studies, Including ≥ 50 With Homogenous Patient Populations With HL or Aggressive NHL or FL, Reporting Outcomes According to Visual Assessment With End-of-Treatment PET

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>No. of Patients</th>
<th>Disease and Stage</th>
<th>No. PET Negative</th>
<th>PET NPV</th>
<th>PPV</th>
<th>FTF/PFS at (years)</th>
<th>PFS/EFS</th>
<th>PET Negative (%)</th>
<th>PET Positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spaepen K et al: Br J Haematol 115:272-278, 2001</td>
<td>2001</td>
<td>60</td>
<td>IIA-IIB HL</td>
<td>55</td>
<td>100</td>
<td>91</td>
<td>2</td>
<td>91</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cerci et al26</td>
<td>2010</td>
<td>50</td>
<td>I-IV HL (patients in CRu/PR on CT)</td>
<td>23</td>
<td>100</td>
<td>92</td>
<td>—</td>
<td>—</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Engert et al27†</td>
<td>2012</td>
<td>739</td>
<td>IIB-IV HL</td>
<td>548</td>
<td>95</td>
<td>NA</td>
<td>5</td>
<td>92</td>
<td>86†</td>
<td></td>
</tr>
<tr>
<td>Barnes et al60</td>
<td>2011</td>
<td>96</td>
<td>I-II nonbulky HL</td>
<td>83</td>
<td>94</td>
<td>46</td>
<td>4</td>
<td>94</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>Spaepen et al61</td>
<td>2001</td>
<td>93</td>
<td>Aggressive NHL</td>
<td>90</td>
<td>100</td>
<td>70</td>
<td>—</td>
<td>—</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td>Micallef et al62*</td>
<td>2011</td>
<td>69</td>
<td>DLBCL</td>
<td>61</td>
<td>90</td>
<td>50</td>
<td>2</td>
<td>78</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Pregno et al63</td>
<td>2012</td>
<td>88</td>
<td>DLBCL</td>
<td>77</td>
<td>100</td>
<td>82</td>
<td>2</td>
<td>83</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Trotman et al63</td>
<td>2011</td>
<td>122</td>
<td>High-tumor burden FL</td>
<td>90</td>
<td>NS</td>
<td>NS</td>
<td>3.5</td>
<td>71</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Dupuis et al64*</td>
<td>2012</td>
<td>106</td>
<td>High-tumor burden FL</td>
<td>83</td>
<td>NS</td>
<td>NS</td>
<td>2</td>
<td>87</td>
<td>51</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** CRu, unconfirmed complete response; CT, computed tomography; DLBCL, diffuse large B-cell lymphoma; EFS, event-free survival; FL, follicular lymphoma; FTF, freedom from treatment failure; HL, Hodgkin lymphoma; NA, not applicable; NHL, non-Hodgkin lymphoma; NPV, negative predictive value; NS, not stated; PET, positron emission tomography; PFS, progression-free survival; PPV, positive predictive value; PR, partial response.

*Prospective study.
†Treatment guided by end-of-treatment PET.
Fig A1. (A) Pretreatment scan: computed tomography, positron emission tomography, and fused images showing disease in left neck (arrow). (B) Example of score 1: complete metabolic response with no uptake in normal-size lymph nodes at site of initial disease in left neck (arrow).
Fig A2. (A) Pretreatment scan: computed tomography, positron emission tomography, and fused images showing disease in left axilla. (B) Example of score 2: residual uptake of intensity < mediastinal blood pool in lymph nodes in left axilla (arrow). Maximum standardized uptake value (SUVmax) in lymph nodes was 1.2; SUVmax in mediastinal blood pool was 1.7.
Fig A3. (A) Pretreatment scan: computed tomography, positron emission tomography, and fused images showing disease in right neck and mediastinum (arrow). (B) Example of score 3: residual uptake of intensity > mediastinal blood pool but < liver in residual mediastinal mass (arrow). Maximum standardized uptake value (SUVmax) in mass was 1.7; SUVmax in liver was 2.2.
Fig A4. (A) Pretreatment scan: computed tomography, positron emission tomography, and fused images showing disease in mediastinum. (B) Example of score 4: residual uptake of intensity > liver in residual mediastinal mass (arrow). Maximum standardized uptake value (SUVmax) in mass was 4.5; SUVmax in liver was 3.2.
Fig A5. (A) Pretreatment scan: computed tomography, positron emission tomography, and fused images showing disease in right neck, mediastinum, and right axilla. (B) Example of score 5: residual uptake in mediastinum with intensity markedly higher than normal liver. Maximum standardized uptake value (SUVmax) in mass was 13.0; SUVmax in liver was 2.3.