New response criteria for lymphomas in clinical trials

B. D. Cheson

Hematology, Georgetown University Hospital, Lombardi Comprehensive Cancer Center, Washington, DC, USA

Introduction

Standardized staging and response assessment are critical to the successful conduct of clinical trials. In turn, clinical trials are essential to the development of new and more effective therapy for patients with lymphomas. In the absence of effective agents, response criteria are almost irrelevant. However, with the increasing number of effective therapies, standardized criteria are necessary to reliably assess and compare results of studies.

Variability in how patients were evaluated led to an International Working Group (IWG) that developed guidelines to standardize normal lymph node size, when and how responses are assessed, and definitions for response categories and endpoints [1]. These recommendations were widely adopted by clinical trials groups and regulatory agencies. However, with their application over time, it became clear that revisions were indicated. For example, the IWG criteria relied on physical examination, with its marked inter- and intraobserver variability, CT scans and SPECT gallium scans, the latter no longer being widely used.

A major problem with the original IWG criteria was the misinterpretation of the term complete remission/unconfirmed (CRu). CRu was originally proposed to designate patients with curable histologies, such as Hodgkin’s lymphoma or diffuse large B-cell lymphoma, with a large mass prior to therapy for whom treatment eradicated all detectable tumor except for persistence of the single mass, which had decreased by at least 75% on CT scan, recognizing that these lesions are scar tissue or fibrosis in >90% of cases [2, 3]. Instead, CRu was often applied to situations in which the sum of the product of the diameters (SPD) of multiple nodes decreased by at least 75%, which would more appropriately designated partial response (PR). A second type of CRu indicated patients who fulfilled all of the conditions for a CR following therapy except that the bone marrow was considered morphologically indeterminate. Instead, the term was also assigned to patients who did not undergo a repeat biopsy to confirm response.

FDG-PET has resulted in a major shift in lymphoma patient management. PET is not useful for diagnosis because it lacks specificity. However, it has been considered for staging, prognosis, directing therapy, restaging and post-treatment surveillance. The strongest evidence for the usefulness of PET is in post-treatment restaging [4–27]. The Ann Arbor system most commonly used was based on physical examination and bone marrow assessment, with CT scans subsequently incorporated. PET is highly sensitive in detecting nodal and extranodal involvement by most histologic subtypes of lymphoma and may provide complementary information to the Ann Arbor staging system [4, 8, 10, 11, 16, 21, 28–38].

Most common lymphoma histologies are routinely FDG avid with a sensitivity and specificity that is superior to CT [28, 29, 32, 38]. Whereas PET and CT are 80–90% concordant in staging of diffuse large B-cell lymphoma, follicular lymphoma and mantle cell lymphoma [10, 32], PET results in upstaging by identifying additional sites of disease. Concordance of PET and CT is lower in Hodgkin’s lymphoma [4, 8, 16, 21, 34–36].

PET can detect bone or bone marrow involvement in lymphoma patients with a negative iliac crest bone marrow biopsy [39–41], although being more sensitive with diffuse disease and less reliable with limited involvement [41]. Thus, PET cannot substitute for bone marrow biopsy in lymphoma staging. PET is currently not part of standard lymphoma staging primarily because of its expense and the generally small percentage of patients (~15–20%) where PET modifies clinical stage, with a change in management in only ~10–15% [11, 32, 42]. Thus, at present PET alone should not replace CT for staging [4, 8, 16, 35].

PET/CT offers important advantages compared with contrast-enhanced, full-dose diagnostic CT or PET alone [37, 43, 44]. The CT portion of the PET/CT exam for initial staging using i.v. contrast may permit a more accurate assessment of liver and spleen compared with unenhanced CT [25]. PET/CT may be valuable in patients with clinical stage I or II disease who are being considered for local radiation therapy.

Numerous studies have demonstrated that PET scans performed after one or more cycles of chemotherapy predict progression-free and overall survival [5–7, 20, 21, 23, 24, 45–47]. Unfortunately, no available data demonstrate that altering treatment on the basis of PET results improves patient outcome. This critically important issue is currently being addressed in a number of clinical trials.

The use of PET in clinical trials

Juweid et al. [19] were the first to demonstrate that integrating PET into the IWG criteria in non-Hodgkin’s lymphoma increased the number of patients with diffuse large B-cell NHL classified as a CR, eliminated CRus, with a clearer separation of the progression-free survival curves between CR and PR patients.
The International Harmonization Project was convened to standardize performance and interpretation of PET in lymphoma clinical trials [25], recommend when PET scans were appropriate in clinical trials considering variability in FDG avidity among the various lymphoma histologic subtypes and the relevant endpoints of clinical trials (Table 1), and develop new response criteria incorporating PET and bone marrow immunohistochemistry (Table 2) [26].

PET scans should be performed at least 6–8 weeks following therapy to reduce false-positive results [25]. PET is essential for restaging the potentially curable lymphoma histologies following completion of therapy since therapeutic intervention is generally indicated if residual disease is present.

In clinical trials where PET is unavailable to the vast majority of participants, or where PET is not deemed necessary or appropriate, response should be assessed as above, but only using CT scans. In this setting, residual masses should not be considered as CRu, but should be designated as PRs.

### Table 1. Recommended timing of PET (PET/CT) scans in lymphoma clinical trials

<table>
<thead>
<tr>
<th>Histology</th>
<th>Pre-treatment</th>
<th>Mid-treatment</th>
<th>Response assessment</th>
<th>Post-tx surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Routinely FDG avid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DLBCL</td>
<td>Yes*</td>
<td>Clinical trial Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>HL</td>
<td>Yes*</td>
<td>Clinical trial Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Follicular NHL</td>
<td>No b</td>
<td>Clinical trial No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>MCL</td>
<td>No b</td>
<td>Clinical trial No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Variably FDG avid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other aggressive NHls No b</td>
<td>Clinical trial No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Other indolent NHls No b</td>
<td>Clinical trial No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

*Recommended but not required pre-treatment.

bRecommended only if ORR/CR is a primary study endpoint.

From Cheson et al. [26].

### Table 2. Response definitions for clinical trials

<table>
<thead>
<tr>
<th>Response remission (CR)</th>
<th>Definition</th>
<th>Nodal masses</th>
<th>Spleen, liver</th>
<th>Bone marrow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete remission (CR)</td>
<td>Disappearance of all evidence of disease</td>
<td>(a) FDG avid or PET+ before therapy: mass of any size permitted if PET–; (b) variably FDG avid or PET–: regression to normal size on CT</td>
<td>Not palpable, nodules disappeared</td>
<td>Infiltrate cleared on repeat biopsy, if indeterminate by morphology immunohistochemistry should be negative</td>
</tr>
<tr>
<td>Partial remission (PR)</td>
<td>Regression of measurable disease and no new sites</td>
<td>≥50% decrease in SPD of up to six largest dominant masses. No increase in size of other nodes</td>
<td>≥50% decrease in SPD of nodules (for single nodule in greatest transverse diameter), no increase in size of liver or spleen</td>
<td>Irrelevant if positive before therapy, cell type should be specified</td>
</tr>
<tr>
<td>Stable disease (SD)</td>
<td>Failure to attain CR/PR or PD</td>
<td>(a) FDG avid or PET+ prior to therapy: PET+ at prior sites of disease and no new sites on CT or PET; (b) variably FDG avid or PET–: no change in size of previous lesions on CT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relapsed or progressive disease</td>
<td>Any new lesion or increase from nadir by ≥50% of previously involved sites</td>
<td>Appearance of a new lesion &gt;1.5 cm in any axis ≥50% increase in the longest diameter of a previously identified node &gt;1 cm in short axis or in the SPD of more than one node; lesions PET+ if FDG-avid lymphoma or PET+ before therapy</td>
<td>≥50% increase from nadir in the SPD of any previous lesions</td>
<td>New or recurrent involvement</td>
</tr>
</tbody>
</table>

SPD, sum of the product of the diameters.

From Cheson et al. [26].
follow-up evaluation

Although widely used in clinical practice, there is no evidence to support regular surveillance CT or PET scans [48, 49]. A number of studies in the pre-PET era demonstrated that it is the patient or physician who identifies the relapse >80% of the time [50–53].

issues with PET(CT)

Assessment of clinical trials incorporating FDG-PET must take into consideration differences in equipment, technique and variability in interpretation. PET/CT makes comparisons with older data difficult. Moreover, there are many causes of false-positive and false-negative PET scans [19, 25, 54, 55]. The International Harmonization Project provided guidance for the interpretation of FDG-PET and generated response definitions to improve interpretation of response, comparability between studies, leading to availability of more effective therapies, and enhancing outcome for patients with lymphoma.

references

33. Blum RH, Seymour JF, Wirth A, MacManus M et al. Frequent impact of [18F] Fluorodeoxyglucose positron emission tomography on the staging and