Journal of Nuclear Medicine, published on July 14, 2022 as doi:10.2967/jnumed.122.264106

APPLICATION OF THE LUGANO CLASSIFICATION FOR INITIAL EVALUATION, STAGING, AND RESPONSE ASSESSMENT OF HODGKIN AND NON-HODGKIN LYMPHOMA: THE PROLOG CONSENSUS INITIATIVE (PART 1- CLINICAL).

Short running title

Lugano classification: PRoLoG clinical

Authors list

Fabien Ricard* (1), Bruce Cheson* (2), Sally Barrington (3), Judith Trotman (4), Annette Schmid (5), Guenther Brueggenwerth (6), Gilles Salles (7), Larry Schwartz (8), Greg Goldmacher (9), Rudresh Jarecha (10), Jayant Narang (5), Florence Broussais (11), Paul Galette (12), Min Liu (13), Surabhi Bajpai (14), Eric Perlman (15), Julie Gillis (16), Ira Smalberg (17), Pierre Terve (18), Gudrun Zahlmann (19), Ron Korn (20)
*contributed equally to this work

Institution/Affiliation

(1) Relay Therapeutics, Cambridge, Massachusetts
(2) Lymphoma Research Foundation, New York, NY, USA
(3) Sally Barrington: King's College London and Guy's and St Thomas' PET Centre, School of Biomedical Engineering and Imaging Sciences, King's College London, King's Health Partners, London, UK
(4) Concord Repatriation General Hospital, University of Sydney, Concord, Australia
(5) Takeda Pharmaceutical Company Ltd, Cambridge, MA, USA
(6) Bayer AG, Berlin, Germany
(7) Department of Medicine, Memorial Sloan Kettering Cancer Center and Weil Cornell Medicine, New York, NY, US
(8) Department of Radiology, Columbia University College of Physicians and Surgeons and Radiologist-in-Chief, The New York Presbyterian Hospital, Columbia, NY, USA
(9) Merck & Co, Inc. Kenilworth, NJ, USA
(10) Deciphera Pharmaceuticals LLC, Waltham, MA, USA

(11) Lymphoma Study Association Research Center LYSARC, Pierre Benite, France
(12) GSK, Collegeville, PA, USA
(13) Autolus Therapeutics, London, UK
(14) Calyx international, Billerica, MA, USA
(15) Perlman Advisory Group LLC, Boynton Beach, FL, USA
(16) Merigold LLC, Branford, CT, USA
(17) Associate Professor of Radiology, Saint John's Cancer Institute and Tower imaging medical group, Sherman Oaks, CA, USA
(18) KEOSYS Medical Imaging, Saint Herblain, France
(19) QIBA/RSNA, Radiological Society of North America, Oak Brook, IL, USA
(20) Adjunct Faculty TGEN/City of Hope and Imaging Endpoints Core Lab, Scottsdale, AZ, USA
*contributed equally to this work

First and Corresponding Author

Fabien Ricard MD MS
399 Binney St,
Cambridge, MA 02139
fabienricardmd@gmail.com
1 (917) 862-2616

Financial support

no

Word count manuscript (main text)

6479 words

ABSTRACT

Rationale: To provide consensus recommendations from a consortium of academic and industry experts in the field of lymphoma and imaging for the consistent application of the Lugano classification.

Methods: Consensus was obtained through a series of meetings from July 2019 until July 2021 sponsored by the PINTaD (Pharma Imaging Network for Therapeutics and Diagnostics) as part of the ProLoG (PINTaD RespOnse criteria in Lymphoma wOrking Group) consensus initiative.

Results: Consensus recommendations clarified technical considerations for PET-CT and diagnostic CT from the Lugano Classification including updating the FDG-avidity of different lymphoma entities, clarifying the response nomenclature and refining lesion classification and scoring, especially with regards to scores 4 and 5 and the X category of the 5 point scale. Combination of metabolic and anatomic responses is clarified and response assessment in case of discordant or missing evaluations. Usage of clinical data in the classification, especially the requirement for bone marrow assessment is further updated based on lymphoma entities. Clarification is provided with regards to spleen and liver measurements and evaluation, as well as nodal response.

Conclusion: Consensus recommendations are made to comprehensively address areas of inconsistency and ambiguity in the classification encountered during response evaluation by end users and such guidance should be used as a companion to the Lugano 2014.

Keywords: Lugano classification, clinical recommendations, consensus, standardization

INTRODUCTION

In 2014, the Lugano classification (*1*) and the companion report (*2*) (referred together as Lugano 2014) provided a standardized approach to classifying response based upon 18F-Fluorodeoxyglucose (18F-FDG) Positron Emission Tomography-Computed Tomography (PET-CT). In particular, Lugano 2014 emphasized the importance of a 5 Point Scale (5-PS) for FDG-avid lymphomas along with a well-defined characterization of splenomegaly, while maintaining many of the anatomic elements of the Revised Response Criteria for Malignant Lymphoma published in 2007 (*3*) (referred as Cheson 2007).

The Lugano classification has been widely adopted by academia, the pharmaceutical industry, and in clinical practice for evaluation of Hodgkin Lymphoma (HL) and non-Hodgkin Lymphoma leading to acceptance by regulatory agencies for drug approval and treating physicians alike. Currently, hundreds actively recruiting and ongoing investigational trials use the Lugano classification (*4*).

As with any criteria, the application of the Lugano classification has uncovered some challenges in implementation resulting in non-uniform usage, variable interpretation, and customized modifications with the potential to undermine effective comparisons between patient groups, treatment regimens and outcomes analyses.

To address these challenges, volunteer leaders from industry and academia, including original authors of the Lugano classification, referred to as the PRoLoG committee (PINTaD RespOnse criteria in Lymphoma wOrking Group), sponsored by the PINTaD (*5*) (Pharma Imaging Network for Therapeutics and Diagnostics), engaged in discussions from July 2019 until

September 2021 to provide expert guidance for the consistent application of the Lugano classification.

This manuscript is not intended to replace the classification but is proposed as a companion to the Lugano 2014. While other lymphoma response criteria have since been published (e.g., LYRIC 2016 (*6*) for immunomodulatory therapies and RECIL 2017 (*7*)), most of the current recommendations may also apply to the newer criteria as well.

The recommendations in this document focus on imaging aspects rather than implementation in clinical practice for treatment decisions. They will hopefully facilitate consistent imaging interpretation and response assessment during clinical trials and may be a valuable addition for healthcare providers.

METHODS

Task forces (TF) were created to evaluate technical and clinical considerations and descriptive ambiguities within the Lugano classification. A steering committee (FR, RK, AS, GB) was formed to oversee the activities of each TF and to summarize, reconcile, and consolidate the recommendations from the regularly scheduled TF meetings. The TF members included independent research leaders (BC, SFB, JT, GS, LS) and representatives from academic/scientific organizations (n=3), industry (n=9), Clinical Research Organizations (n=13) and other clinical trial specialists (n=4). All meetings were held virtually, from July 2019 to September 2021.

TF meetings were recorded, minutes transcribed and approved by TF members. Recommendations were based on a hierarchical approach, with evidence-based decisions providing the strongest level of support, followed by best practices, then expert consensus

opinions. Where there was lack of evidence-based data, or consensus, a call for future research on that topic was suggested. Additional recommendations from the TF, primarily for advanced imaging technical considerations will be available elsewhere ("Application of the Lugano Classification for Initial Evaluation, Staging, and Response Assessment of Hodgkin and Non-Hodgkin Lymphoma: The PRoLoG Consensus Initiative (Part 2- technical)").

Finally, the term "end-user" in this manuscript refers to any individual involved in the implementation of the Lugano classification – e.g., clinical trialists, physicians, scientists, data managers, statisticians, scientific and medical writers, healthcare providers, program coders, and regulatory personnel. The term "reviewer" in this document is defined as any physician responsible for assessing response in lymphoma, such as an imaging specialist (radiologist and/or nuclear medicine physician) or a clinical specialist (oncologist, hematologist, radiation oncologist).

IMAGING ACQUISITION CONSIDERATIONS

The use and frequency of acquisition of PET-CT and/or a diagnostic CT with contrast enhancement depends upon several factors. These include the clinical question, lymphoma histology and stage, FDG-avidity, and efficacy endpoints. In FDG-avid lymphomas, a diagnostic CT scan may not be required at each scheduled tumor assessment where 18F-FDG PET-CT is scheduled, e.g. when it is prespecified in a clinical trial protocol that 18F-FDG PET-CT is required for each imaging visit, then no additional diagnostic CT examination may be needed. Similarly, an 18F-FDG PET-CT scan may not be required at each time point, e.g. 18F-FDG PET-CT is usually discouraged for surveillance (*8*,*9*). While the role of surveillance imaging is not established

in clinical practice (*10-13*), a diagnostic CT may still be required in follow-up of clinical trials using time-dependent endpoints (e.g. progression free survival (*1*)).

It should be noted that the terminology PET-CT-based versus CT-based response criteria in Lugano refers to PET-CT as PET corrected for attenuation by CT, i.e. for metabolic assessment and localization of lesions, and to define CT as diagnostic quality CT for morphologic assessment.

FDG-AVIDITY OF LYMPHOMA ENTITIES

Although most lymphomas are FDG-avid, metabolic imaging may be less reliable for response assessment in some histologies because of greater inter- or intra-individual lesion variability in FDG uptake. Entities of lymphoma can be categorized as:

- Routinely FDG-avid lymphoma (*2,14*) (e.g. HL, diffuse large B-cell lymphomas [DLBCL], follicular lymphoma [FL] (*15-18*), mantle cell lymphoma [MCL] (*19-21*), nodal peripheral T-cell lymphoma (*22-24*), lymphoblastic (*25-27*) and Burkitt lymphoma (*28,29*)), to be assessed by 18F-FDG PET-CT and, when anatomic assessment is required, by diagnostic CT.
- Lymphomas that are generally not FDG-avid should be assessed with diagnostic CT and not with 18F-FDG PET-CT (e.g. small lymphocytic lymphoma [SLL], chronic lymphocytic leukemia [CLL]), unless for suspected or documented transformation.
- Other lymphomas, while commonly FDG avid, have variability in FDG uptake, either inter-patient and/or inter-lesional (e.g. some marginal-zone lymphoma (*30-32*), some T-cell (*33*) notably cutaneous T-cell lymphomas). There is no formal recommendation for which type of imaging to be performed; it should be based on the lymphoma entity and can be aligned with Health Authorities. In general, baseline may include 18F-FDG PET-CT and diagnostic CT.

Patients without FDG-avid lesions at baseline should be followed with diagnostic CT (unless transformation is suspected). In patients with FDG-avid lesions at baseline, PET-CT may be used for response assessment and rules for combination of metabolic and anatomic response should be prespecified in the protocol.

LESION CLASSIFICATION, SCORING AND RESPONSE NOMENCLATURE

Common Lesion Classification and Response Nomenclature

CT: tumor lesions should be referred to as either Target Lesion (TL, assessed quantitatively) or Non-Target Lesions (NTL, assessed qualitatively). Nodal and extra-nodal lesions should ideally be documented as separate classifications since they have different assessment rules.

18F-FDG PET-CT: assessment nomenclature is designated as the 5-point scale (5-PS). The 5-PS is based on the single most metabolically active lesion (with visual or semi-quantitative assessment), which can vary at each time point. SUV values that are captured (e.g. most hypermetabolic lesion, reference regions) usually represent the maximum values (SUVmax), in alignment with the Lugano classification. Other types of measurements (e.g. SUVpeak, SUVmean) are being explored for use in clinical trials (*34*,*35*) and further work is required in this field to identify the optimal measure. Besides, metabolic assessments (e.g. metabolic tumor volumes) or radiomics may become more important in the future.

Both CT and PET-CT responses should be reported when available, and designated “M” for metabolic or “A” for anatomic, as well as the overall response (i.e. the response to be used for determining endpoints, integrating imaging response [metabolic, anatomic, or combination of both, when available] and clinical data when available). See supplemental Table 1.

Scoring of lesions on CT and PET-CT and metabolic response category

TL selected on CT at baseline should be FDG-avid, with higher uptake than normal liver, for FDG-avid lymphoma. While the 5-PS was not originally intended to be applied at baseline, this means 5-PS > 3.

Protocol inclusion criteria for FDG-avid lymphoma should state that eligible subjects must have at least one FDG-avid lesion; it is recommended that it means at least one lesion with higher intensity than normal liver (for FDG-avid lymphomas) and at least one CT measurable lesion (when anatomic measurements are required).

A score of 4 should be applied to lesion uptake greater than the uptake in a large region of normal liver, i.e. not only to uptake moderately greater than liver as this was originally stated in the Lugano classification. When a semi-quantitative approach is used, this means an uptake greater than the SUVmax in a large region of normal liver.

A score of 5 should be applied to lesion uptake markedly greater than liver or to hypermetabolic new lesions (NL), or both, and the reason for assessing a score of 5 should be collected (uptake or NL or both). When a semi-quantitative approach is used, this means an uptake at least 2 times SUVmax in the liver since both thresholds of 2 or 3 times have been used in published reports (*2,34,36-41*) and there is no further evidence to recommend one or the other. If such a semi-quantitative approach is used, the threshold value that will be used should be defined a priori in clinical study documents and reported (*42*).

The optimal threshold for response likely depends on lymphoma entity, treatment and timing; further research is recommended to define score 4 and score 5 (*36,43-45*).

When evaluating the metabolic response category, the overall metabolic uptake (i.e. intensity and extent) has to be considered along with the 5-PS. A visual score of 4 or 5 with reduced

intensity and no increase in the extent, is Partial Metabolic Response (PMR), while increased intensity and/or increased extent is Progressive Metabolic Disease (PMD). Stable intensity uptake with no increase in extent is No Metabolic Response (NMR).

The “X” defined in the Lugano classification as areas of uptake unlikely to be related to lymphoma should not be considered as a category by itself and the reviewer should always assign a 5-PS in addition to “X”.

RESPONSE ASSESSMENT

Imaging Response Assessment

When using the Lugano classification for FDG-avid lymphomas, the metabolic response assessed on 18F-FDG PET-CT should take precedence over the anatomic response assessed on diagnostic CT.

While the definition of a PMR lacks an objective quantitative cut-off in the Lugano classification, there is insufficient evidence at this time to further define a PMR for most lymphoma, and efforts at further standardization are warranted (e.g. delta SUV, change in metabolic tumor volume).

The below defines rules for combining metabolic and anatomic responses in FDG-avid lymphomas when both modalities are available:

- Acceptable assessments of imaging Complete Response (CR)
    - PET Complete Metabolic Response (CMR) + CT response (either Complete Anatomic Response [CAR] or partial anatomic response [PAR]) or stable anatomic disease (SAD)

    - PET CMR + CT progressive anatomic disease (PAD) within the same visit window, i.e. the progressive disease seen on CT does not correlate to any metabolic progression

- Acceptable assessments of imaging Partial Response (PR)
    - PET PMR + CT response (either CAR or PAR) or SAD
    - PET PMR + CT PAD within the same imaging window, i.e. the progressive disease seen on CT does not correlate to any metabolic progression
- Acceptable Assessments of imaging Stable Disease
    - PET NMR + CT response (either CAR or PAR) or SAD
    - PET NMR + CT PAD by new anatomic lesion within the same imaging window, i.e. the new lesion seen on CT is not showing hypermetabolism suspicious for lymphoma

The TF recognizes (but does not recommend, since metabolic response should take precedence over anatomic response) the practice by which imaging response based on PET-CT (e.g. CMR/PMR) is downgraded when CT shows PAD. Such cases may be reassigned as overall CR/PR based on clinical review (i.e. the hematology-oncology review that is performed in some clinical trials after the imaging review), after biopsy or follow-up imaging.

In the Lugano classification (*1*), CMR, PMR and NMR require the absence of new lesions for FDG-avid lymphomas. For clarification, any new lesion not considered to be lymphoma, whether metabolically active or not, does not represent disease progression.

Further Considerations for Discordant Cases Between diagnostic CT and 18F-FDG PET-CT

In routinely FDG-avid lymphomas when the 18F-FDG PET results are discordant with diagnostic CT, the PET results should supersede CT interpretation with the caveat that the overall time point response can be overridden during a clinical review by integration of clinical data into the imaging assessment, if applicable, or if additional data, such as biopsy/imaging follow-up are subsequently provided. For example, when the 18F-FDG PET is CMR (or PMR or NMR), but CT demonstrates new or growing metabolically inactive lesions, it is unlikely that this represents lymphoma in a routinely FDG-avid histology, and CR (or PR or Stable Disease respectively) can usually be assigned.

Biopsy of a growing or new lesion and/or follow-up should be strongly encouraged, as clinically appropriate, as well as search for alternative causes. A positive biopsy (including via endoscopy if gastro-intestinal lesion), or cytology (if effusion), or follow-up confirmation of disease would preclude an overall time point response of CR (or PR or Stable Disease respectively) and be considered Progressive Disease (PD). PD would then need to be backdated to the first appearance of the growing/new lesion. This should be prespecified in study documents.

In non-FDG-avid lymphomas, CT results should supersede PET for the imaging time point response assessments and the CT-based response as per the Lugano classification (*1*) should be used. If CT scans visits are missing, the imaging time point response would be not evaluable (NE) unless a PET-CT has been performed and the CT portion of the PET-CT is of diagnostic quality, based on reviewer judgement, to permit accurate tumor burden assessments.

Please refer to section “FDG-avidity of lymphoma entities” for recommendations in lymphomas with variability in FDG uptake.

Assessment of Response When PET-CT or diagnostic CT Imaging Visits Are Missing or not done as per protocol

Best practice recommendations for PET scheduling in pivotal clinical trials, when acceptable and reasonable, are to time the frequency of PET-CT acquisitions, with the anticipated response to the intervention and provide details for superseding rules (i.e. how to carry over responses when one or the other modality is not done at every visit).

When PET-CT is not available, but a diagnostic CT is, the PET-CT response can be carried forward from the prior visit to provide an imaging response assessment as long as the diagnostic CT scan does not suggest disease deterioration (nor clinical status, with regards to overall response, in the case where clinical review is performed).

When diagnostic CT is not available, but there has been no substantial change on 18F-FDG PET-CT, the results of the prior CT can be carried forward. On occasion, the CT portion of PET (CTAC) can be used to assess the CT disease burden if considered of suitable diagnostic quality.

Note: It is very common for clinical trials to be using a modified Lugano classification (i.e. with variations from the original publication). In such case, it should be required to define what "modified" means.

INCORPORATION OF CLINICAL DATA

Imaging and Clinical Response Assessments

Best practice opinions suggest that a paradigm of independent review by imaging specialists followed by clinical oncology review to update results according to clinical and laboratory data introduces the least amount of bias into the process while providing the most reliable and consistent results.

For studies not using an independent clinical oncology review, it is suggested that imaging reviewers be provided with some limited clinical information, to be prospectively defined in the protocol.

Clinical Data Requirements

There is no requirement for integrating clinical information per Lugano guidance, except for bone marrow (BM) biopsy (BMB) and aspiration for lymphoma histologies where PET-CT may not be a substitute for this information.

The clinical data that should be provided to the reviewer must be defined in study documents and be consistently recorded and provided as a structured report or dossier with pertinent clinical information (e.g. BMB results, lesion biopsy/fluid evaluation if performed, concomitant therapy that could affect scan results, such as the use of colony stimulating factors (CSF), infection/inflammation or other information that can confound PET-CT and diagnostic CT findings, clinical and laboratory information).

In general, physical examination data should not be provided to the central reviewer since imaging should take precedence over clinical examination for lesion measurement; except for lesions that would not be captured on imaging (e.g. scalp, lower extremities).

As well, when feasible, appropriate but limited clinical history and information should be provided to imaging reviewers to better select lesions at baseline (e.g. prior radiation therapy).

Recommendations for assessment of Bone Marrow involvement

Although BM samples should usually be obtained prior to the start of therapy, many patients with relapsed/refractory disease have BM results in the pre-baseline period that could eliminate the need for a repeat biopsy prior to receiving therapy, especially when it was positive. In general, it should be discussed if BM results from the pre-baseline period may be used for the baseline within a timeframe to be prespecified per protocol (typically BM results should be dated no longer than 3 months prior to start of therapy, and unless clinical changes suggest otherwise).

Requirement for repeat BMB in a clinical trial are based on the setting (e.g. lymphoma entity, FDG-avidity, study phase and endpoints) and should be prespecified in study documents.

BM involvement in DLBCL and HL tends to be focal in appearance, whereas diffuse avidity suggests an inflammatory process. Rarely, predominant BM based disease in DLBCL can present with intense diffuse uptake. Involvement by follicular and other low-grade lymphomas may not be apparent because of the indolent nature of the diseases (*46*).

- In FDG-avid lymphomas

- In cases where BM sampling is negative at baseline, it is reasonable to assign a CR as overall response if patient achieves a metabolic CMR without repeating the BM.
- In HL and DLBCL, a baseline BMB may not be required in all patients as PET-CT may substitute for BM evaluation as per Lugano classification (*1,42,47-50*). Wherein the patient achieves a CMR, it is reasonable to assign a CR as overall response, whatever the status of BM sampling at baseline. Requirement for BMB should be prespecified in the clinical study protocol.
- In FL, while there is new evidence that BM sampling may not be mandatory in all trials (*51-53*), PET-CT does not uniformly substitute for BM biopsy for staging and response assessment and may still need to be obtained, especially in patients without BM uptake on 18F-FDG PET-CT at baseline. It should be pre-specified in study documents if a patient who had positive BM uptake on PET at baseline and achieves a PET-CT CMR can be assigned a CR as overall response if BM sampling is not done.
- In FDG-avid lymphomas when a BMB during or at the end of treatment shows lymphoma involvement the best response can be PR, even with an otherwise CMR.

- In non-FDG-avid lymphomas or lymphomas with variability in FDG uptake
  - In cases where BM sampling is negative at baseline, it is reasonable to assign a CR as overall response if the patient achieves a CAR (and CMR if PET-CT is available)
  - In cases where BM sampling is positive or unknown at baseline, and BM sampling is not obtained or is positive during/at end of treatment, but the patient achieves a CAR (and CMR if PET-CT is available), it should be downgraded to a PR as overall response.

- In situations where BMB findings are indeterminate, it is reasonable to downgrade a PET CMR to PR for lymphomas where PET cannot substitute to BMB.

EVALUATION OF SPLEEN, LIVER AND NODAL INVOLVEMENT

Spleen and Liver size and nodules

Spleen: Expert judgement of the reviewer should be used in instances where the size measurement is inconsistent with the rest of the tumor burden. Spleen size can vary with factors unrelated to lymphoma involvement including patient age, body dimensions and sex (*54*), non-malignant conditions (e.g. enlargement from portal hypertension, splenic vein thrombosis), technical factors such as respiratory motion on CT and prior injury/trauma. Thus, expert reviewer should determine the status of the spleen with respect to splenomegaly when measurements are close to the 13cm threshold before, during or following treatment.

Liver: in alignment with the Lugano classification, liver size should no longer be considered as part of the assessment.

Nodules/masses in the spleen and liver should be recorded as part of the anatomic tumor lesion assessment (TL/NTL). When standard diagnostic CT is acquired, intravenous injection of contrast during anatomic imaging, unless contra-indicated, is paramount for the evaluation of lesions in solid organs, which may not be visible without contrast.

Modality for Spleen Measurement

The TF recommends that when splenic size assessments are required, a diagnostic CT should be used and vertical length be reported (Supplemental Figure 1).

If a diagnostic CT is not available, the splenic measurement from the CTAC may be used if considered to be of acceptable quality by the reviewer; if the CTAC is considered of unacceptable quality for measurement (e.g. major breathing motion artifacts), splenic measurements on PET should be discouraged and, unless splenic size would not have an impact on the outcome, it should be reported as NE.

Clinical palpation is not considered adequate for determination of splenic length.

New and Recurrent Splenomegaly

As defined in the Lugano classification, an increase of at least 2.0 cm should be applied to both new and recurrent splenomegaly. Progression should be assessed compared to the nadir (which can be the baseline).

Liver Used as a Reference for the 5-PS

When the liver is used as a reference site, the reference region in the liver should avoid the liver margins and any focal hepatic involvement. Where diffuse hepatic involvement occurs, reviewers should use their expert judgment to decide whether the liver can be used as a reference organ, though the TF was not able to provide an alternative organ reference tissue in this scenario due to lack of available publications on the matter and the rarity of the circumstance.

Of note, uptake higher than liver in areas with high physiological uptake may not always preclude the assessment of a CMR, such as in Waldeyer's ring, or extranodal sites with high physiologic uptake (e.g. gastro-intestinal tract, oesophago-gastric junction) or with activation within spleen or marrow (e.g. with chemotherapy or granulocyte CSF).

New Nodal Lesions and Regrowth of Nodal Lesions on CT

In addition to the size threshold (i.e. >15 mm in the longest transverse diameter), it is recommended to apply a 5 mm absolute increase from nadir to declare new or recurrent nodal lesions and be careful when assessing progression in small nodes for which limited variation in size may represent physiological or post-therapeutic changes (e.g. nodes replenished with B cells months after discontinuation of antiCD19/20 therapies) in order to avoid overcalling progression due to small size variation.

Discordance Between Splenic and Nodal Disease Outcomes

In cases of nodal response but unequivocal new/recurrent splenomegaly presumed due to lymphoma (e.g. with FDG uptake on PET-CT, suspicious for lymphoma involvement), it is recommended to report disease progression.

Conversely, in situations where FDG-avid lymphomas have sustained splenomegaly on CT without FDG uptake higher than normal liver but complete resolution of FDG activity in nodal tissue, a CMR (and thus an imaging CR) may be declared per Lugano classification.

Additionally, consideration of other conditions that may cause diffuse increase in organ FDG uptake is suggested since several pharmaceutical products (e.g. granulocyte CSF) or other treatments given to support blood counts may increase splenic activity.

Further recommendations for the evaluation of spleen and nodes can be found in Supplemental 1 and Supplemental Figures 2-5. Summary tables of recommendations can be found in Supplemental Table 2.

CONCLUSION

The PRoLoG initiative has created a platform to gather recommendations from an international group of recognized imaging and clinical experts from industry and academia in the field of lymphoma response assessment to standardize application of the Lugano classification in clinical trials and beyond.

These recommendations are intended for clinical users, at local sites and central facilities, in academic and pharmaceutical clinical trials and should be used as a companion to the Lugano 2014 to enhance assessment of response, and facilitate clinical trial conduct and regulatory review, ultimately leading to improved lymphoma patient outcome.

ACKNOWLEDGEMENTS

The authors thank all the members from the PINTaD who participated to the PRoLoG initiative, especially Klaus Noever, Melissa Burkett, Anand Devasthanam, Nick Enus, Andres Forero Torres, Nick Galante, Sayali Karve, Katarina Ludajic, Michael ONeal, Ravikanth Mankala, Daniel Mollura, and Jason Vilardi. In addition, the authors thank Tina Nielsen and John Sunderland for their review, as well as the original authors of the classification and companion paper, especially Richard Fisher, Rodney Hicks, Otto Hoekstra, Lale Kostakoglu, Michel Meignan, George Mikhaeel, Michael O'Doherty and Emanuele Zucca. Finally, the authors thank Deming Litner for her administrative support and the CDISC group for discussion on the terminology to be used for anatomic and metabolic responses.

Professor Barrington acknowledges support from the National Institute for Health and Care Research (NIHR) [RP-2-16-07-001] and by core funding from the Wellcome/EPSRC Centre for Medical Engineering at King's College London [WT203148/Z/16/Z] and the NIHR Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London and the NIHR Clinical Research Facility. The views expressed by Professor Barrington are not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

DISCLOSURE

This project was sponsored by PINTaD and the views expressed are those of the authors, not necessarily from their institution. No financial support was received for this project. Fabien Ricard, employed by Bayer at the time of first submitting this manuscript, is now an employee and has shares of Relay Tx. Paul Galette is employed by and shareholder of GSK. Greg Goldmacher is employed by Merck and has stocks in Merck, ImmunoGen and Aveo. Julie Gillis, formerly employed by Imaging Endpoints, is at Merigold LLC. Pierre Terve is with Keosys medical imaging. Min Liu has shares at Autolus. Larry Schwartz reports BMS, Regeneron, Merck as member of independent review panel and data safety monitoring board for clinical trials. Jayant Narang and Rudresh Jarecha, formerly employed by Calyx, are with Takeda Pharmaceuticals and Deciphera Pharmaceuticals, respectively. Ron Korn is owner and CMO at Imaging Endpoints Core Lab, consultant for Virginia G Piper Cancer Center, ImaginAB Technologies, and shareholder of Verve Medical, Telelite Health, Globavir, Renibus. No other potential conflicts of interest relevant to this article exist.

KEY POINTS

Question: How can the Lugano classification be consistently applied among clinical end users?

Pertinent Findings: These consensus recommendations should be used as a companion to the Lugano Classification with regards to FDG-avidity of different lymphoma entities, response nomenclature and lesion classification and scoring. Response assessment, usage of clinical data and spleen, liver and nodal evaluation are clarified.

Implications for patient care: This guidance will enhance usage of the Lugano Classification, facilitating clinical trial conduct and regulatory review, ultimately leading to improved lymphoma patient outcome.

REFERENCES

1. Cheson BD, Fisher RI, Barrington SF et al. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32:3059–3067.
2. Barrington SF, Mikhaeel NG, Kostakoglu L, 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*. 2014;32:3048–3058.
3. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007;25:579–586.
4. https://clinicaltrials.gov. Accessed on 23Apr2022.
5. https://www.pintad.net. Accessed on 23Apr2022.
6. Cheson BD, Ansell S, Schwartz L, et al. Refinement of the Lugano classification lymphoma response criteria in the era of immunomodulatory therapy. *Blood.* 2016;128:2489-2496.
7. Younes A, Hilden P, Coiffier B, et al. International working group consensus response evaluation criteria in lymphoma (RECIL 2017). *Ann Oncol. 2017;*28:1436–1447.
8. Cheah CY, Hofman MS, Dickinson M, et al. Limited role for surveillance PET–CT scanning in patients with diffuse large B-cell lymphoma in complete metabolic remission following primary therapy. *Br J Cancer. 2013;*23:312-317.
9. Thanarajasingam G, Bennani-Baiti N, Thompson CA, et al. PET-CT in staging, response evaluation, and surveillance of lymphoma. *Curr Treat Options Oncol. 2016;*17:24.
10. Thompson CA, Ghesquieres H, Maurer MJ, et al. Utility of routine post-therapy surveillance imaging in diffuse large B-cell lymphoma. *J Clin Oncol. 2014;*32:3506-3512.

11. Rutherford SC. Surveillance scanning in lymphoma. *Clin Adv Hematol Oncol*. 2019;17:352-359.
12. Kallam A, Adusumalli J, Armitage JO. surveillance in patients with diffuse large B cell lymphoma. *Mayo Clin Proc. 2020;*95:157-163.
13. Zelenetz AD, Gordon LI, Abramson JS, et al. NCCN guidelines insights: B-cell lymphomas, Version 3.2019. *J Natl Compr Canc Netw*. 2019;17:650-661.
14. Weiler-Sagie M, Bushelev O, Epelbaum R, et al. (18)F-FDG avidity in lymphoma readdressed: a study of 766 patients. *J Nucl Med*. 2010;51:25-30.
15. Trotman J, Fournier M, Lamy T, et al. Positron emission tomography-computed tomography (PET-CT) after induction therapy is highly predictive of patient outcome in follicular lymphoma: analysis of PET-CT in a subset of PRIMA trial participants. *J Clin Oncol. 2011;*29:3194-3200.
16. Luminari S, Biasoli I, Arcaini L, et al. The use of FDG PET in the initial staging of 142 patients with follicular lymphoma: a retrospective study from the FOLL05 randomized trial of the Fondazione Italiana Linfomi. *Ann Oncol. 2013;*24:2108-2112.
17. Dupuis J, Berriolo-Riedinger A, Julian A, et al. Impact of [(18)F]fluorodeoxyglucose positron emission tomography response evaluation in patients with high-tumor burden follicular lymphoma treated with immunochemotherapy: a prospective study from the Groupe d'Etudes des Lymphomes de l'Adulte and GOELAMS. *J Clin Oncol*. 2012;30:4317-4322.
18. Trotman J, Barrington SF, Belada D, et al. Prognostic value of end-of-induction PET response after first-line immunochemotherapy for follicular lymphoma (GALLIUM): secondary analysis of a randomised, phase 3 trial. *Lancet Oncol. 2018;*19:1530-1542.

19. Bodet-Milin C, Touzeau C, Leux C, et al. Prognostic impact of 18F-fluoro-deoxyglucose positron emission tomography in untreated mantle cell lymphoma: a retrospective study from the GOELAMS group. *Eur J Nucl Med Mol Imaging. 2010;*37:1633-1642.
20. Bailly C, Carlier T, Berriolo-Riedinger A, et al. Prognostic value of FDG-PET in patients with mantle cell lymphoma: results from the LyMa-PET Project. *Haematologica. 2020;*105:e33–36.
21. Albano D, Bosio G, Bianchetti N, et al. Prognostic role of baseline18F-FDG PET/CT metabolic parameters in mantle cell lymphoma. *Ann Nucl Med. 2019;*33:449–458.
22. Cottereau AS, Becker S, Broussais F, et al. Prognostic value of baseline total metabolic tumor volume (TMTV0) measured on FDG-PET/CT in patients with peripheral T-cell lymphoma (PTCL). *Ann Oncol. 2016;*27:719-724.
23. Cottereau AS, El-Galaly TC, Becker S, et al. Predictive value of PET response combined with baseline metabolic tumor volume in peripheral T-cell lymphoma patients. *J Nucl Med. 2018;*59:589-595.
24. Mehta-Shah N, Ito K, Bantilan K, et al. Baseline and interim functional imaging with PET effectively risk stratifies patients with peripheral T-cell lymphoma. *Blood Adv*. 2019;22:187-197.
25. Fox TA, Carpenter B , Khwaja A, et al. Utility of FDG-PET/CT in lymphoblastic lymphoma. *Blood*. 2019;134(Suppl 1):2890
26. Fox TA, Carpenter B, Taj M, et al. Utility of 18F-FDG-PET/CT in lymphoblastic lymphoma. *Leuk Lymphoma. 2021;*62:1010-1012.

27. Becker S, Vermeulin T, Cottereau AS, et al. Predictive value of 18F-FDG PET/CT in adults with T-cell lymphoblastic lymphoma: post hoc analysis of results from the GRAALL-LYSA LLO3 trial. *Eur J Nucl Med Mol Imaging. 2017;*44:2034–2041.
28. Zeng W, Lechowicz MJ, Winton E, et al. Spectrum of FDG PET/CT findings in Burkitt lymphoma. *Clin Nucl Med. 2009;*34:355-358.
29. Carillo-Cruz E, Marin-Oyaga VA, Rodriguez MS, et al. Role of 18F-FDG-PET/CT in the management of Burkitt lymphoma. *Eur J Haematol. 2015;*94:23-30.
30. Vaxman J, Bernstine H, Kleinstern G, et al. FDG PET/CT as a diagnostic and prognostic tool for the evaluation of marginal zone lymphoma. *Hematol Onco*. 2019;37:168-175.
31. Carrillo-Cruz E, Marín-Oyaga VA, de la Cruz Vicente F, et al. Role of 18F-FDG-PET/CT in the management of marginal zone B cell lymphoma. *Hematol Oncol*. 2015;33:151-158.
32. Ceriani L, Meignan M. Present role and future perspective of PET-CT in marginal zone lymphoma. *Ann Lymphoma*. 2020;4:13.
33. Feeney J, Horwitz S, Gönen M, et al. Characterization of T-cell lymphomas by FDG PET/CT. *Am J Roentgenol*. 2010;195:333–340.
34. Barrington SF, Kluge R. FDG PET for therapy monitoring in Hodgkin and non-Hodgkin lymphomas. *Eur J Nucl Med Mol Imaging. 2017;*44(Suppl 1):97-110.
35. Lodge MA. Repeatability of SUV in Oncologic 18F-FDG PET. *J Nucl Med*. 2017;58:523-532.
36. Barrington SF, Phillips EH, Counsell N, et al. Positron emission tomography score has greater prognostic significance than pretreatment risk stratification in early-stage Hodgkin lymphoma in the UK RAPID study. *J Clin Oncol. 2019;*37:1732-1741.

37. Johnson P, Federico M, Kirkwood A, et al. Adapted treatment guided by interim PET-CT scan in advanced Hodgkin's lymphoma. *N Engl J Med*. 2016;374:2419-2429.

38. Gallamini A, Tarella C, Viviani S, et al. Early chemotherapy intensification with escalated BEACOPP in patients with advanced-stage Hodgkin lymphoma with a positive interim positron emission tomography/computed tomography scan after two ABVD cycles: long-term results of the GITIL/FIL HD 0607 trial. *J Clin Oncol.* 2018;36:454-462.

39. Mikhaeel NG, Cunningham D, Counsell N, et al. FDG-PET/CT after 2 cycles of RCHOP in DLBCL predicts complete remission but has limited value in identifying patients with poor outcome - Final result of UK-NCRI prospective study. *Br J Haematol.* 2021;192:503-513.

40. Eertink JJ, Burggraaff CN, Heymans MW, et al. Optimal timing and criteria of interim PET in DLBCL: a comparative study of 1692 patients. *Blood Adv*. 2021;5:2375-2384.

41. Ceriani L, Martelli M, Gospodarowicz MK, et al. Positron emission tomography/computed tomography assessment after immunochemotherapy and irradiation using the Lugano classification criteria in the IELSG-26 study of primary mediastinal B-cell lymphoma. *Int J Radiat Oncol Biol Phys. 2017;* 97:42-49.

42. Barrington SF, Trotman J. The role of PET in the first-line treatment of the most common subtypes of non-Hodgkin lymphoma. *Lancet Haematol.* 2021;8:e80-93.

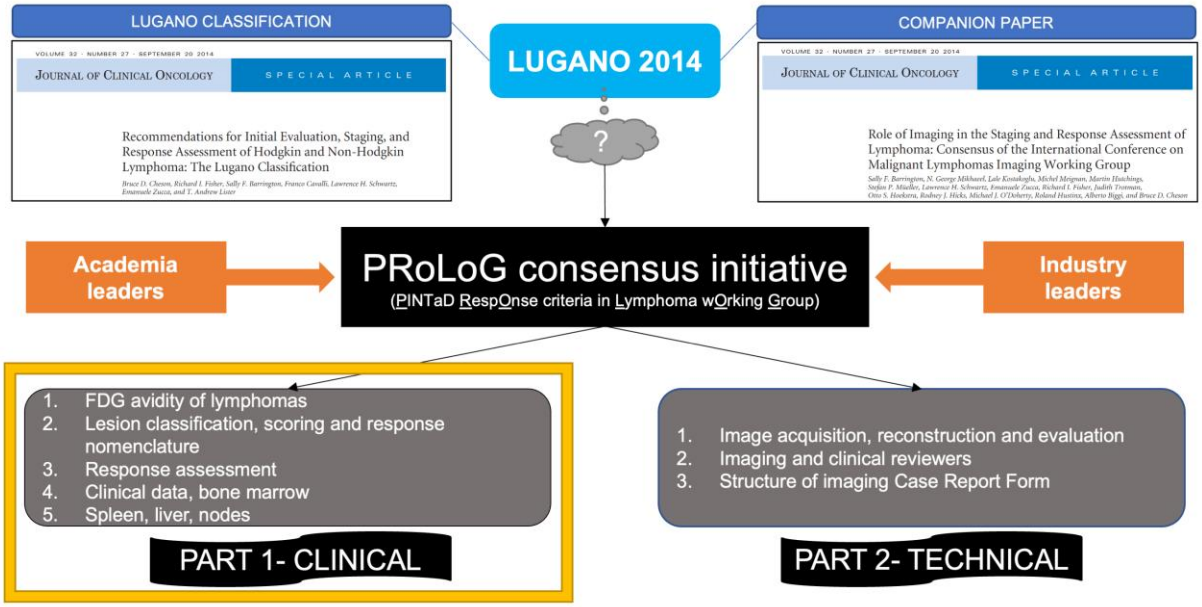
43. Itti E, Juweid ME, Haioun C, et al. Improvement of early 18F-FDG PET interpretation in diffuse large B-cell lymphoma: importance of the reference background. *J Nucl Med.* 2010;51:1857–1862

44. Eertink JJ, Burggraaff CN, Heymans MW, et al. Optimal timing and criteria of interim PET in DLBCL: a comparative study of 1692 patients. *Blood Adv.* 2021;5:2375-2384

45. Ceriani L, Martelli M, Gospodarowicz, et al. Positron emission tomography/computed tomography assessment after immunochemotherapy and irradiation using the Lugano classification criteria in the IELSG-26 study of primary mediastinal B-cell lymphoma. *Int J Radiat Oncol Biol Phys.* 2017;97:42-49
46. Barrington SF and Meignan M. Time to prepare for risk adaptation in lymphoma by standardizing measurement of metabolic tumor burden. *J Nucl Med*. 2019;60:1096-1102.
47. Trotman J, Barrington SF. The role of PET in first-line treatment of Hodgkin lymphoma. *Lancet Haematol*. 2021;8:e67-79.
48. Voltin CA, Goergen H, Baues C, et al. Value of bone marrow biopsy in Hodgkin lymphoma patients staged by FDG PET: results from the German Hodgkin Study Group trials HD16, HD17, and HD18. *Ann Oncol*. 2018;29:1926-1931.
49. Alzahrani M, El-Galaly TC, Hutchings M, et al. The value of routine bone marrow biopsy in patients with diffuse large B-cell lymphoma staged with PET/CT: a Danish-Canadian study. *Ann Oncol*. 2016;27:1095-1099.
50. Barrington SF, Kirkwood AA, Franceschetto A, et al. PET-CT for staging and early response: results from the Response-Adapted Therapy in Advanced Hodgkin Lymphoma study. *Blood*. 2016;127:1531-1538
51. Rutherford SC, Yin J, Pederson L, et al. Relevance of bone marrow biopsies for response assessment in NCTN follicular lymphoma clinical trials. *J Clin Oncol. 2020;*38(Suppl):8038-8038
52. Rutherford SC, Li V, Ghione P, et al. Bone marrow biopsies do not impact response assessment for follicular lymphoma patients treated on clinical trials. *Br J Haematol. 2017;*179:242-245.

53. Rutherford SC, Herold M, Hiddemann W, et al. Impact of bone marrow biopsy on response assessment in immunochemotherapy-treated lymphoma patients in GALLIUM and GOYA. *Blood Adv*. 2020;4:1589-1593.
54. Chow KU, Luxembourg B, Seifried E, et al. Spleen size is significantly influenced by body height and sex: establishment of normal values for spleen size at us with a cohort of 1200 healthy individuals. *Radiology. 2016;*279:306-313.

GRAPHICAL ABSTRACT



SUPPLEMENTAL 1: ADDITIONAL RECOMMENDATIONS FOR THE EVALUATION OF SPLEEN AND NODES

Method for Splenic Measurements

A unified approach to the methods used to measure spleen size does not exist and the TF has reviewed several methods. This includes direct measurement of the spleen on coronal (or sagittal) images from tip to tip (i.e. along the long axis of the spleen) or by placing a vertical measurement from the top of the spleen to inferior splenic tip in a vertical direction. However, these techniques suffer from user variability and free hand placement of the electronic cursor, as the superior and inferior splenic tip may be located on different coronal (or sagittal) slices. Measuring splenic size using the so-called "axial" method seems to provide the most consistent and objective method. In this approach the reviewer identifies the number of axial slices on which the spleen is visible and multiplies the number of slices by the reconstruction interval or slice thickness (including gap as applicable). In some cases, it may be appropriate to add an additional 1 slice thickness to the calculation to correct for partial volume averaging. The axial slices for splenic measurement should be determined from axial series in which a standard reconstruction algorithm has been applied. During this process the reviewer must ensure that the most cranial and the most caudal slices with significant visible spleen tissue is included in the size determination. If coronal or sagittal methods are used, they should only be applied to images that have been acquired as isotropic voxels. In this instance, efforts should be made to capture the true vertical length of the spleen and avoid oblique measurement (Supplemental Figure 2). Newer and more exploratory software algorithms that automatically segment and calculate splenic size using AI approaches are promising but will require validation prior to endorsement by the TF. The same

method should be consistently applied for a same patient throughout the trial and ideally be the same for all patients.

New/recurrent splenomegaly

The spleen must increase by at least 2 cm from the nadir and be over 13 cm to declare both new/recurrent splenomegaly. For example, a spleen that measures 12 cm at nadir will require at least a 2 cm growth to 14 cm while a spleen that measures 9 cm at nadir must grow more than 4 cm to cross the >13 cm threshold to be considered new/recurrent splenomegaly.

Progressive Splenomegaly

Progressive splenomegaly (Supplemental Figures 3 and 4) requires several conditions to be met:

- The splenic length must increase by > 50% in the enlarged portion of the spleen
- There needs to be at least a 1 cm of absolute change
- Progression is always measured from nadir (note: baseline can be the nadir when there has been no response in spleen)

Responsive Splenomegaly

Splenic based response (Supplemental Figure 5) requires the following conditions to be met:

- Regression of >50% in length beyond normal in comparison to baseline

- No minimum absolute size change required
- Response is measured in comparison to baseline

Not Evaluable Spleen and Splenectomy Patients

If the spleen cannot be evaluated at a specific timepoint (e.g. missing images), it is recommended that the anatomic response be NE at that timepoint, except if it was documented that the spleen was uninvolved with lymphoma since baseline, or unless otherwise specified in a clinical trial protocol.

In the case of splenectomy before baseline, it is recommended that disease be assessed without the spleen category. In the case of splenectomy while on trial, the patient assessment should be censored at the time of splenectomy, except if it was documented that the spleen was uninvolved with lymphoma at this time.

There is no specific recommendation how accessory spleen in patients with splenectomy should be assessed but the reviewer should be cautious that an increase in size in accessory spleens may reflect physiologic growth of the splenic tissue.

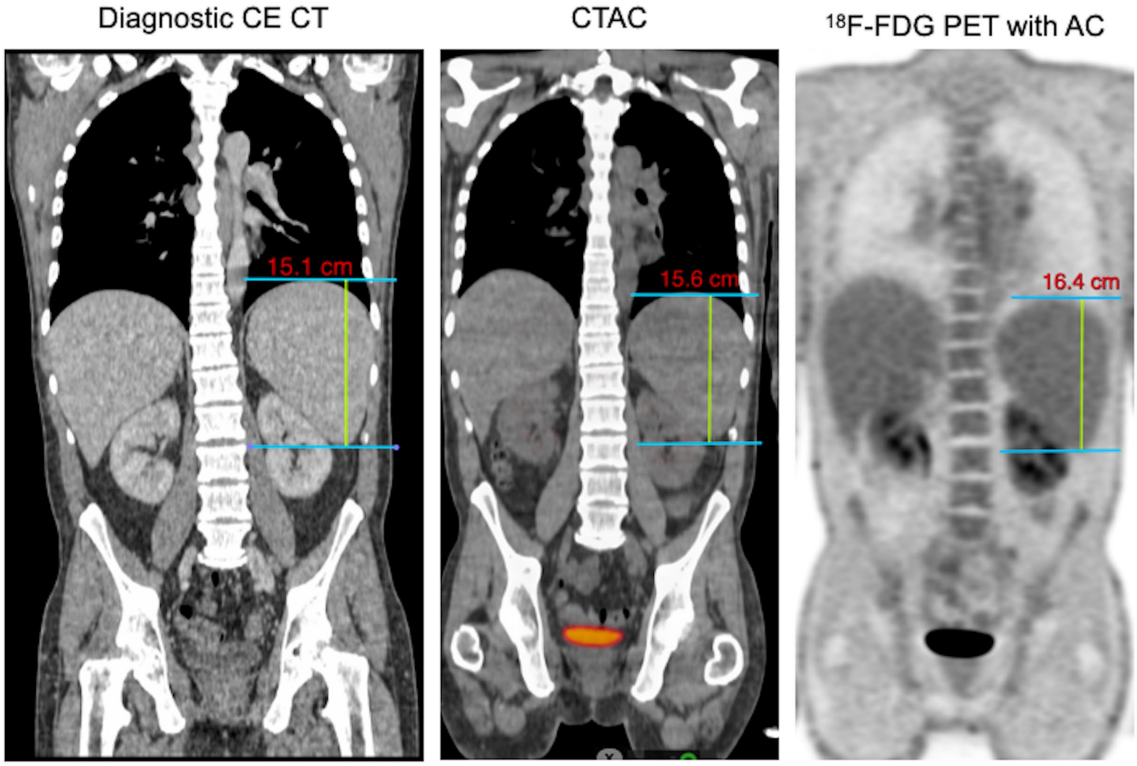
Lymphoma Involving the Spleen Only

In cases of spleen-only lymphoma (e.g. splenic marginal zone lymphoma, primary splenic lymphoma), the TF recommends that spleen assessments evaluate for the presence of nodules/masses using CT (for all histologies) and PET-CT (when FDG-avid) with the standard thresholds for spleen size applied.

New/Regrowth of a nodal lesion

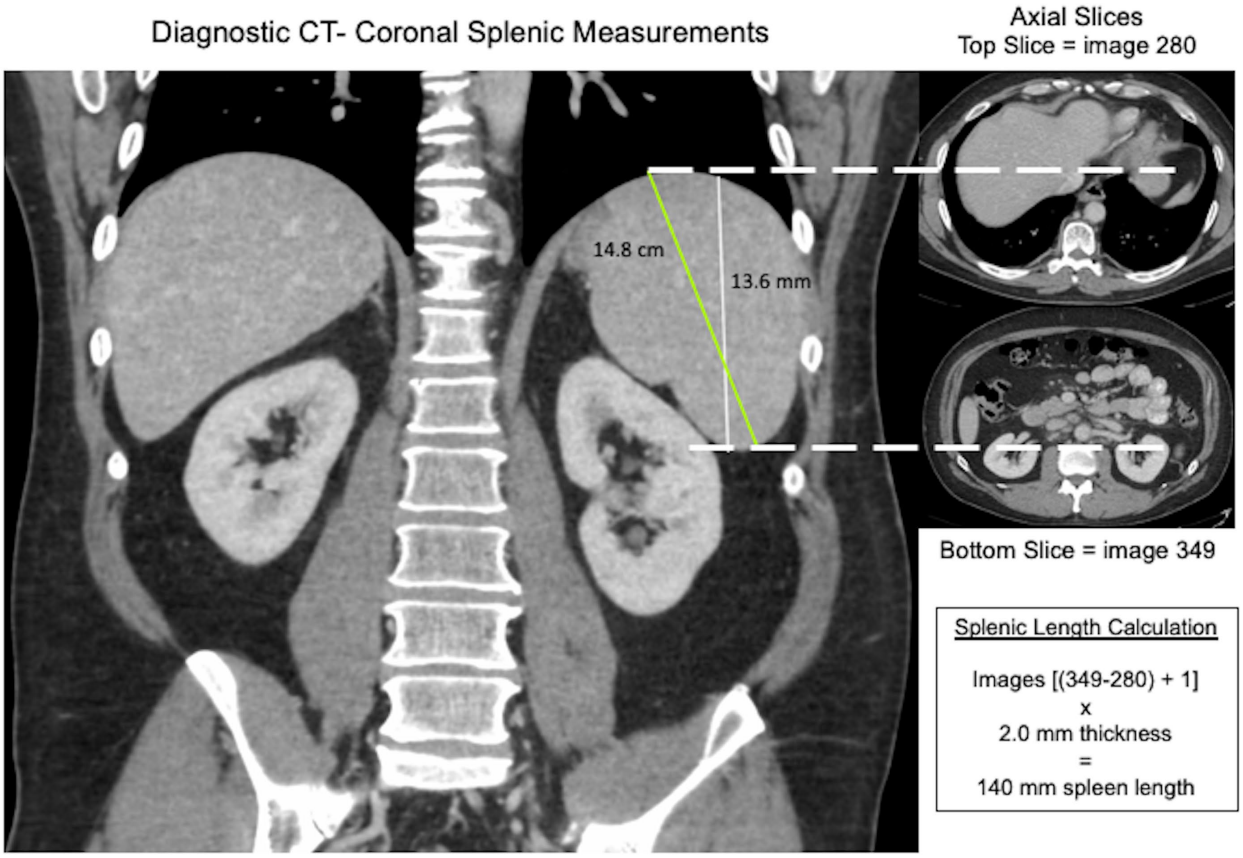
For example, a node that measures 12 mm in long axis diameter (non-pathological) would need to increase to 17 mm to be considered a new nodal lesion (or regrowth, if the node was pathological prior to shrinkage to 12 mm).

SUPPLEMENTAL FIGURE 1: Modality for measuring Splenic Length



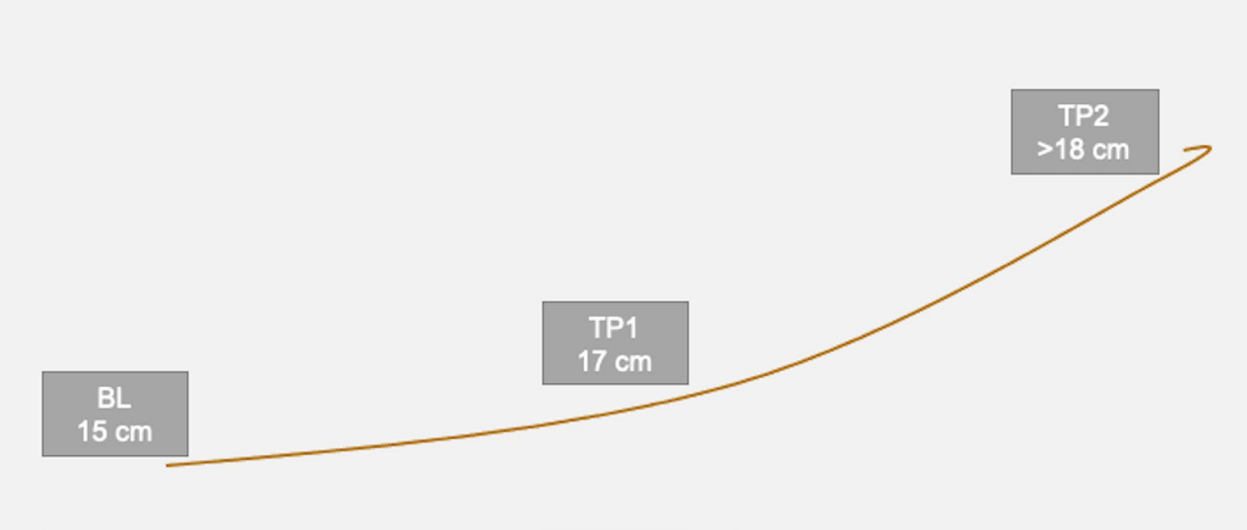
Differences in splenic length measurements are dependent upon imaging acquisition techniques. This patient with DLCBL had same day diagnostic contrast enhanced CT (CE CT; left panel) and FDG PET-CT scans (middle and right panels). Using the diagnostic CE CT (left panel) as the gold standard for splenic length measurements, the ^{18}F-FDG PET with attenuation correction (AC; right panel) clearly overestimates splenic size. Courtesy of Imaging Endpoints Core Lab, USA.

SUPPLEMENTAL FIGURE 2: Methods Commonly Used to Measure Splenic Length



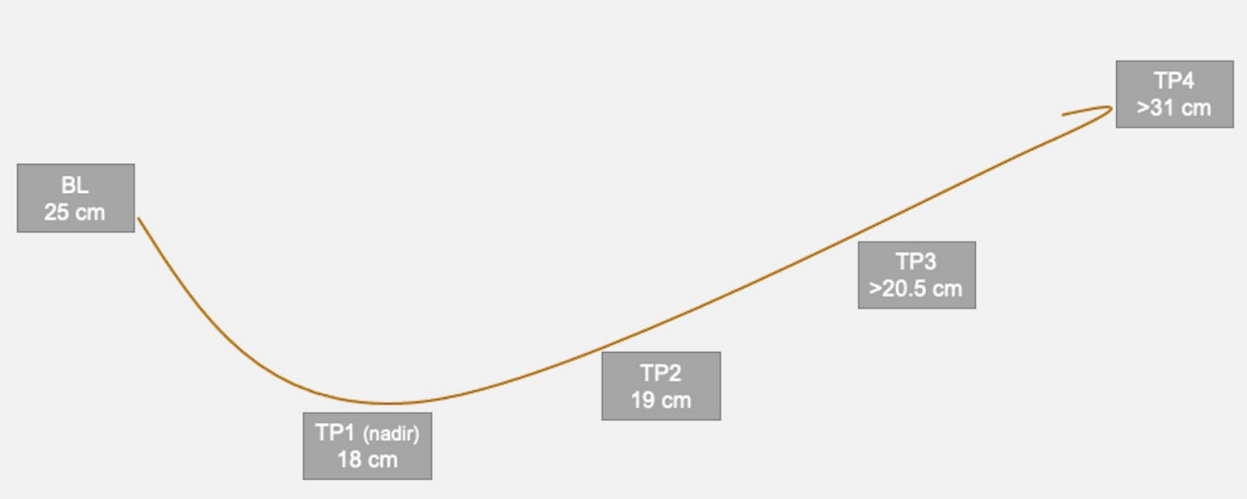
Coronal and Axial methods to measure spleen size on CT. If the coronal length of the spleen is used to measure the spleen size, it should be measured from the most cranio-caudal portions of the spleen (13.6cm). Oblique coronal length measurement should be avoided (14.8cm). Courtesy of Imaging Endpoints Core Lab, USA.

SUPPLEMENTAL FIGURE 3: Representation of Splenomegaly-Based Disease Progression from Baseline Nadir Values



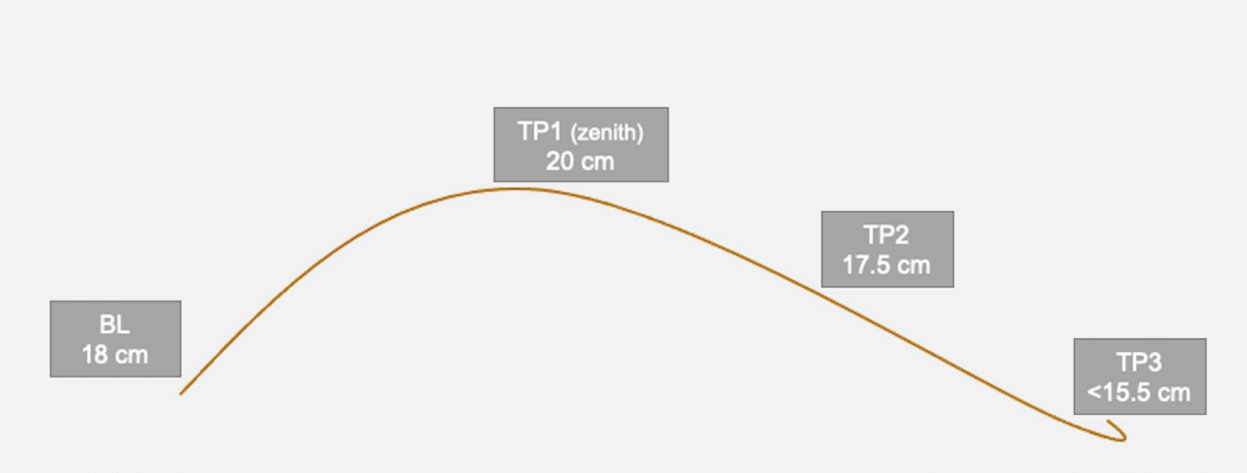
Anatomic progression in spleen can be called at TP1 (>50% increase of enlarged portion, i.e. > (15-13)/2 which means > 1cm, and at least 1 cm absolute growth compared to spleen size at baseline). BL = baseline; TP1= first timepoint after baseline; TP2 = second timepoint after baseline.

SUPPLEMENTAL FIGURE 4: Representation of Splenomegaly-Based Disease Progression from Non-Baseline Nadir Values



Disease progression based upon splenomegaly when there is a >50% increase in splenic length from Nadir. Note that disease progression will be declared when there is >50% increase in splenic length compared to nadir, or as shown in this example, a 2.5cm or more growth from TP1 since the spleen is 5.0 cm beyond the ULN). Based on this calculation, progression would be declared at TP3. Note that TP1 shows partial response in spleen in this example. Sizes in the illustration represent spleen size. BL= baseline; TP1 =first timepoint after baseline; TP2 = second timepoint after baseline, etc.)

SUPPLEMENTAL FIGURE 5: Representation of Spleen-Based Response when Splenomegaly is Present at Baseline



In this example, the enlarged portion of splenic length at BL is 5.0 cm (18.0 cm – 13.0 cm ULN). AT TP1, the spleen has only enlarged an additional 2.0 cm from baseline (or 40%) and does not meet the definition of disease progression. Partial spleen response at TP3 (reduction > 5/2, i.e. regression >2.5cm from baseline). BL: baseline, TP1: first timepoint after baseline, TP2: second timepoint after baseline, etc.)

SUPPLEMENTAL TABLE 1: Terminology for Metabolic, Anatomic (radiographic) and Overall responses

Type of Response*	Metabolic	Anatomic	Overall†
Complete response	CMR	CAR	CR
Partial response	PMR	PAR	PR
Stable disease/No response	NMR (preferred term, otherwise SMD)	SAD	SD
Progressive disease	PMD	PAD	PD

CMR: complete metabolic response; PMR: partial metabolic response; NMR: no metabolic response; SMD: stable metabolic disease; PMD: progressive metabolic disease; CAR: complete anatomic response; PAR: partial anatomic response; SAD: stable anatomic disease; PAD: progressive anatomic disease; CR: complete response; PR: partial response; SD: stable disease; PD: progressive disease

**this table is for terminology purpose only and does not describe how to combine metabolic and anatomic responses (see section "Response Assessment")*

†*overall response (used for determining endpoints) integrates imaging response (metabolic, anatomic, or combination of both, when available) and clinical data, when available*

SUPPLEMENTAL TABLE 2: SUMMARY TABLE OF PRoLoG RECOMMENDATIONS (PART 1- CLINICAL)

Issue	PRoLoG Recommendations	Comments
FDG-avidity of lymphoma entities and imaging guidance	FDG-avidity of lymphomas • FDG-avid (e.g. HL, DLBCL, FL, MCL, nodal peripheral T-cell lymphoma, lymphoblastic and Burkitt lymphoma) • Mostly not FDG-avid (e.g. SLL, CLL) • Lymphomas with inter-patient and/or inter-lesional variability in FDG uptake (e.g. some marginal-zone lymphoma, some T-cell notably cutaneous T-cell lymphomas)	• While routinely FDG-avid lymphomas should be assessed by FDG PET-CT and, when anatomic assessment is required, diagnostic CT, there is no formal recommendation for those with variability in FDG uptake. • Non FDG-avid lymphomas are not expected to be asssessed by FDG PET-CT unless transformed/suspected for transformation
Lesion Classification for Reporting	Anatomic Imaging (CT/MRI) • Target lesion (TL), non-target lesion (NTL), new lesion (NL) • May be useful to designate nodal and extranodal lesions separately	• Aligns with RECIST approach for lesion labeling • Keeps compliant with CDISC compliance standards • Different treatment regimens may affect nodal vs extranodal disease differently
	FDG PET Imaging • Individual hottest lesion does not require to be labelled as TL or NTL • NL designation may still be useful	• Only the lesion with the highest uptake on the 5-PS needs to be selected at each timepoint for determination of metabolic response, thus no TL/NTL labeling is required • Score 5 refers to uptake markedly higher than liver or NLs or both (to be specifically captured in imaging CRF)

Scoring of lesions	• A score of 4 should be applied to lesion uptake greater than the uptake in a large region of the liver (i.e. greater than the SUVmax in a large region of normal liver) • A score of 5 should be applied to lesion uptake markedly greater than liver (i.e. at least 2 times higher) or to hypermetabolic new lesions	• Visual or semi-quantitative assessment can be used • A score of 4 applies to uptake greater than liver, not only moderatly greater as originally stated in the classification • The reason for a score of 5 (uptake, or new lesion, or both) should be collected in the CRF • Both 2x and 3x thresholds have been used and the threshold to be used in score 5 for a specific study should be prespecified

Metabolic response category	• Overall metabolic uptake (i.e. intensity and extent) should be considered.
	• In case of 5-PS of 4 or 5
	  o With reduced intensity and no increae in extent, it is PMR
	  o With increased intensity and/or extent, it is PMD
	  o With stable intensity and no increase in extent, it is NMR/SMD
	• X is not a category by itself and a 5-PS must be assigned in addition to X
		• Besides the 5-PS assessment, based on the hottest lesion, both overall intensity and extent of FDG uptake must be considered to evaluate the metabolic response
		• X relates to areas of uptake unlikely to be due to lymphoma and should not be considered as a response category

Response Nomenclature	Response Assessment Designations • Anatomic responses should use a “A” designation to indicate a radiologic change (e.g. CAR, PAR, SAD and PAD) • Metabolic responses should use a “M” designation to indicate a metabolic	• Allows for separate reporting of anatomic vs metabolic vs overall changes • Highlights differences in imaging results for further investigation of discordance • Overall Time Point Response integrating imaging response with clinical data, when available, should

	(FDG) change (e.g CMR, PMR,NMR/SMD, PMD) • Overall Time Point response should continue to apply conventional designations (e.g. CR, PR, SD, PD)	continue to use conventional reporting to avoid confusion amongst investigators, patients and regulatory agencies
Imaging Response Assessment	• Maintain recommendations from Lugano classifications for CR and PR unless stated in trial protocol • Use caution when designating PD in patients with new lesions on anatomic imaging that are metabolically inert on PET-CT in FDG-avid lymphomas • Areas of uptake on FDG PET (existing and/or new) are not always malignant	• A reduction in tumor size by anatomic imaging does not need to accompany a CMR or PMR • Any modification of response categories from Lugano Classification should be clearly stated in the Protocol and other essential documents and preferably discussed with regulatory agencies prior to study start up • In routinely FDG-avid lymphomas (e.g., HL, DLBCL, FL, MCL…), any new finding on anatomic imaging that does not have increased uptake is considered most likely unrelated to the disease, until proven otherwise by a biopsy or follow-up • FDG uptake can be due to infections, inflammatory or adverse treatment events and providing information that is specified prospectively (e.g. relevant AE information) to the central reviewer should be considered.
Assessment of Imaging Response When Current Imaging Visits Are Missing	• In selected circumstances (e.g. FDG avidity of lymphoma entity), the results from prior time point assessments of PET or CT can be carried forward when anatomic or metabolic assessment at current time point is missing	• When either PET or CT is missing but the other scan is available for response assessment, the missing scan results can be carried forward from the prior visit to provide an overall response assessment as long as the current available scan does not suggest disease deterioration compared to the last available scan for the missing modality (nor clinical status, with regards to overall response, where a clinical review is performed)

		• On occasion, the CT portion of PET can be used as a substitute for the missing diagnostic CT if considered of suitable diagnostic quality
Assessment of Imaging Response in Discordant Cases Between CT and FDG PET	• PET responses should take precedence over CT when evaluating FDG-avid malignancies (e.g. HL, DLBCL, FL, MCL…) • CT responses should take precedence over PET when evaluating non-FDG-avid malignancies • No formal recommendation for lymphomas with variability in FDG uptake	• These are general guidelines as the task force recognizes that exceptions can exist based upon tumor biology • When in doubt, follow-up imaging and/or tissue biopsy may be warranted to resolve discordances by integrating the totality of clinical and radiographic data available for review
Imaging and Clinical Response Assessments	• The radiology reviewer may receive prespecified clinical data when providing responses to integrate clinical data into the time point response • Best practices suggests a read paradigm of independent review by imaging specialists followed by clinical oncology review to update results based upon clinical and laboratory data information	• If not carefully controlled, clinical information and data results may unblind or bias the radiology review during scan assessments. • Exceptions are e.g. information on treated lesions when selecting TLs for radiologic responses • In cases where an independent review is performed, a set of limited and pre-specified data may be appropriately provided to the reviewer (e.g. bone marrow biospy information, other pathology results, intercurrent infection, etc)
Clinical Data Requirements	• Integrating clinical information per Lugano guidance is not required except for bone marrow data, where required. When feasible, appropriate but limited clinical history and information should	• Recommendations that information supplied to the reviewer be carefully screened to prevent either treatment arm unblinding or undue reviewer bias.

	be provided to reviewers and be pre-specified in study documents.	
Clinical assessment of Bone Marrow	• The requirements for BM sampling should be clearly stated in the clinical trial protocol • Pre-baseline BM results (i.e. results obtained within e.g. 3 months before the screening period) should be discussed and prespecified if it can be used as baseline value, unless clinical changes suggest otherwise • The meaning of indeterminate BM results should be clearly defined for consideration of time point response assessments • BM involvement in DLBCL and HL tends to be focal in appearance, whereas diffuse avidity suggests an inflammatory process. • For HL and DLBCL, FDG results can substitute for bone marrow biopsy to define (BM) status as stated in Lugano Classification. Bone marrow sampling may not be needed. For FL, bone marrow sampling may not be mandatory for all efficacy trials (depending on the study endpoints) but is still recommended in some instances as PET may not substitute for BM biopsy, e.g. when BM is negative on PET at	• The type, timing and analysis of BM results needs to be defined in the trial protocol • BM biopsies may be avoided in certain patients based upon lymphoma entities (e.g. HL and DLBCL) and patient's prior treatment history. For example, it may be appropriate to avoid a screening bone marrow biopsy in a patient with DLBCL or HL where FDG PET can substitute for BMB. • Rarely, predominant BM based disease in DLBCL can present with intense diffuse uptake. • Involvement by follicular and other low-grade lymphomas may not be apparent because of the indolent nature of the diseases. • New treatment regimens (e.g., cellular therapy) may infiltrate the BM but should be recorded as indeterminate for lymphoma by IHC evaluation.

	baseline. All other lymphoma entities require negative BM results for CR unless BM sampling was negative prior to treatment	
Spleen Size	• Clear guidance on what should be considered splenomegaly is helpful (i.e. spleen > 13cm in vertical plane) • However, in some cases body habitus and other factors may be taken into account by expert reviewer judgment • Reviewer expert judgement may be used in select situations where a diagnostic CT is not available or if splenic size is close to the 13.0 cm threshold. • The splenic length does not always need to regress to under 13.0 cm for assessing an overall CR	• Splenic length can be influenced by patient characteristics (age, sex, body habitus) benign disease and method of evaluation. Therefore, rigid reliance on absolute thresholds for splenomegaly may not accommodate the expert understanding of the disease state. • Any modification of splenic measurement assessment should be clearly stated in clinical protocol and other essential documents • Persistent splenic enlargement without evidence of lymphoma involvement elsewhere (e.g. negative BM results, FDG response, no CT evidence of disease and no additional clinical or laboratory evidence of lymphoma) should not preclude an overall assessment of CR.
Modalities for Spleen Measurements	• Diagnostic CT with or without contrast enhancement is the modality of choice for splenic size assessments • CT component from PET-CT may be used with caution • Length measurements from FDG PET alone should not be done • Ultrasound should only be used in special situations (e.g. pediatric lymphomas)	• Considerable variability in splenic length measurements can occur between different imaging modalities • When precise spleen length measurements are required then diagnostic CT (and/or MRI) should be used as they provide the most accurate representation of spleen length • The same modality should be used for longitudinal assessments of spleen length within the same patient whenever possible

	• Newer techniques (e.g. AI, automated volume assessments) require further validation	
Methodology for Spleen Measurements	• A consistent method should be used for measuring the vertical (cranio-caudal) spleen length • Axial methodologies (extrapolating the vertical size of the spleen from the number of axial slices on which spleen is visible) may be more reproducible than coronal techniques	• Switching methodologies between imaging visits can create significant variabilities which could impact response assessments • The position of the spleen on cross-sectional imaging can be different between patients. The axial methods may provide the most consistent and reproducible approach to approximate the vertical length of the spleen since it is less dependent upon orientation of the spleen in the longest axis
Using Splenic Measurements For Response	New and Recurrent Splenomegaly • Progression is determined from nadir, which may be the baseline scan • Increase ≥ 2 cm and be over 13 cm	• Clarify recommendations from Lugano classification
	Progressive Splenomegaly • Progression is determined from nadir, which may be the baseline scan • Increase by > 50% in the enlarged portion of the spleen and ≥1 cm absolute change	• Clarify recommendations from Lugano classification
	Responsive Splenomegaly • Response is determined from baseline	• Maintain recommendations from Lugano classification

	• Clinical Palpation of the spleen is not recommended	• Maintain recommendations from Lugano classification
Splenectomy	Prior Splenectomy • Disease assessments should proceed without spleen category if splenectomy is performed before start of treatment On-study Splenectomy • Response assessment should be censored at the time of splenectomy unless the spleen is free from lymphoma	• Clarify rules for disease assessment in case of prior or on-study splenectomy
Discordance between spleen and extrasplenic disease response	• In cases of nodal disease response but unequivocal new/ recurrent splenomegaly due to lymphoma (e.g. suspicious splenic FDG PET uptake) recommend disease progression	• The judgment of the reviewer may be considered for borderline cases of new/recurrent splenomegaly for spleen measurements in cases of discordant response
New Nodal Lesions & Regrowth of nodal lesions	• it is recommended to apply a 5 mm absolute increase from nadir, in addition to the size threshold (i.e. >15 mm in the longest transverse diameter), to declare new or recurrent nodal lesions.	• Be careful when assessing progression in small nodes for which limited variation in size may represent physiological or post-therapeutic in order to avoid overcalling progression due to small size variation
Liver	Hepatic enlargement • Hepatomegaly is not considered part of Lugano classification assessments • Focal hepatic disease should be assessed as part of the TL/NTL and for metabolic lesion scoring	• Maintain recommendations from Lugano classifications • Focal hepatic lesions can be selected as TL if they meet the appropriate size criteria • Focal hepatic lesions can be selected for PET evaluation by 5-PS if they are the hottest focus

	Liver as the reference organ for 5-PS=3 • Uninvolved hepatic parenchyma serves as the reference tissue for scores 3, 4 and 5 • Avoid using uninvolved hepatic parenchyma near focal lymphoma lesions • No recommendations can be offered when there is diffuse hepatic involvement as an alternative reference source	• Technical considerations and FDG uptake times can affect hepatic uptake • Lesions with abnormal FDG uptake can affect the visual appearance of adjacent normal areas due to count spill-over. Therefore, the use of reference tissue uptake near lesions should be avoided • An alternative to hepatic reference tissue (5-PS = 3) is not readily apparent but diffuse hepatic lymphoma without areas of uninvolved tissue is a rare event

AI: artifical intelligence, BM: bone marrow, BMB: BM biopsy, CDISC: clinical data interchange standards consortium, CLL: chronic lymphocytic leukemia, CMR: complete metabolic response, CR: complete response, CRF: case report form, CRR: complete response radiographic, CT: computed tomography, DLBCL: diffuse large B cell lymphoma, FDG: FluoroDeoxyGlucose, FL: follicular lymphoma, HL: Hodgkin lymphoma, IHC: immunohistochemistry, MCL: mantle cell lymphoma, MRI: magnetic resonance imaging, MZL: Mantle zone lymphoma, NMR: no metabolic response, NL: new lesion, NTL: non target lesion, PD: progressive disease, PDR: progressive disease radiographic, PET: positron emission tomography, PMD: progressive metabolic disease, PMR: partial metabolic response, PR: partial response, PRR: partial response radiographic, RECIST: response evaluation criteria in solid tumors, SD: stable disease, SDR: stable disease radiographic, SLL: small lymphocytic lymphoma, SMD: stable metaboblic disease, SUVmax: maximum standardized uptake value, TL: target lesion, WM: Waldenstrom macroglobulinemia, 5-PS: 5 point scale,