SUPPLEMENTAL 1: EXAMPLE STRUCTURE OF AN IMAGING CASE REPORT FORM (CRF) TO BE USED WITH LUGANO 2014

A. MAIN DATA TO BE CAPTURED FOR DIAGNOSTIC CT ASSESSMENT

- timepoint number, anatomical coverage, quality*
- Up to 6 Target Lesions (TL)
    - Per lesion [record Long axis Diameter (LDi) and perpendicular greatest Short axis Diameter (SDi)]
        - Nodal (must be ≥15mm in Long Axis to be selected as baseline TL)
            - Product of perpendicular diameter (LDi x SDi), or too small to measure (TSTM)
            - Anatomic Location (including laterality, if applicable)
        - Non-nodal (must be ≥10mm in Long Axis to be selected as baseline TL)
            - Product of perpendicular diameters (LDi x SDi), or too small to measure (TSTM)
            - Anatomic location (including laterality and organ)
    - All TL Assessments
        - Sum of perpendicular diameters (LDi x SDi)

        - Overall TL response (note: single lesion increase may trigger progressive disease [PD])
- Non-Target Lesions (NTL)
    - Qualitative Assessments of all other lesions: additional measurable nodal or non-nodal disease, non-measurable but assessable disease (lesions that are poorly defined or difficult to measure).
    - Per organ or lesion group to include
        - Site of involvement/organ system
        - Laterality,
        - Lesion status (present, resolved to normal, progressed, not evaluable)
    - All NTL Assessments
        - NTL response
- Spleen
    - Cranio-Caudal length
    - Spleen Size Response
- New lesions
    - Present or Absent
    - If Present

        - Lesion location, organ, laterality
        - Lesion size for nodal disease to confirm that it meets criteria for new or recurrent nodal disease
- Other findings (e.g. pleural effusions, ascites, etc)
- Anatomic Response
    - Complete anatomic response (CAR)
    - Partial anatomic response (PAR)
    - Stable anatomic disease (SAD)
    - Progressive anatomic disease (PAD)
    - Not evaluable (NE)
- Comments, clarifications or justifications for overriding Lugano Classification

B. MAIN DATA TO BE CAPTURED FOR FDG PET-CT ASSESSMENT

- timepoint number, anatomical coverage, quality*
- Most metabolically active lesion at Baseline (can include nodal, extra nodal, splenic or bone marrow tissue)
    - Anatomic Location
    - FDG uptake compared to Mediastinum and Liver

        - Optional 5 point-scale (5-PS) = 1,2,3,4,5, NE

- Most metabolically active lesion at follow-up (can include nodal, extra nodal, splenic or bone marrow tissue)
    - Anatomic location
    - 5-PS = 1,2,3,4,5, NE
        - if score of 4 or 5 at follow-up, compare uptake intensity (hottest lesion) and extent (hypermetabolic tumor burden) with baseline**
            - Increase in intensity
            - Increase in extent
            - Decrease in intensity
            - Decrease in extent
            - No change in disease intensity
            - No change in disease extent
        - If score of 5 at follow-up, uptake markedly greater than liver or New Lesion (or both)
        - X (new areas of uptake unlikely to be related to lymphoma). Note: the X could be specified in the comments section so there is no question about the interpretation of it not being considered lymphomatous. Note 2: X should not be

considered a category for metabolic response and a visual score should always be assigned in addition to X (e.g. complete resolution of all uptake with new lesions due to another etiology would be 5-PS=1, X).

- Bone marrow uptake
    - Baseline
        - No evidence of lymphoma
        - Evidence of lymphoma
        - Indeterminate
    - Follow-up
        - No evidence of FDG uptake consistent with lymphoma
        - FDG uptake greater than normal but less than baseline and consistent with lymphoma
        - No change in FDG uptake
        - Higher than normal and significantly increased in comparison to nadir
        - new or recurrent FDG-avid foci
        - Increased uptake due to reactive changes including therapy but not lymphoma
        - Indeterminate uptake
- Metabolic response

- Complete metabolic response (CMR)
- Partial metabolic response (PMR)
- No metabolic response (NMR)/Stable metabolic disease (SMD)
- Progressive metabolic disease (PMD)
- NE

- Comments, clarifications or justifications for overriding Lugano Classification

C. IMAGING TIME POINT RESPONSE BASED UPON IMAGING FINDINGS

- Complete response (CR)
- Partial response (PR)
- Stable disease (SD)
- Progressive disease (PD)
- NE

- Comments, clarifications or justifications for overriding Lugano Classification

D. COMMENTS SECTIONS

It is recommended that the comments sections include both a structured listing and free comments and that reviewers be trained to standardize the usage of the comments sections with:

- equivocal lesions: a formal documentation of equivocal lesions is required (e.g. accurately back-dating progression may be recommended when lesion is subsequently confirmed to be unequivocal)
- findings/uptake considered non-malignant lesions: it is recommended that concise comments are provided on significant findings that may potentially be misinterpreted as malignant lesions, to facilitate monitors and/or regulatory reviews
- free comments: scan quality, spleen status, not available (NA) timepoints, NE lesions, impact of change in modality…

* Image quality can be assessed as follows:

- Adequate: the scan has acceptable image quality such that reliable quantitative and qualitative assessments can be made.
- Suboptimal but interpretable: the scan has reduced image quality which may allow for reliable qualitative assessments but unreliable quantitate assessments. The reviewer should provide comments indicating the reason(s) for selecting this category.
- Inadequate (i.e., uninterpretable and/or non-diagnostic): the scan has unacceptable image quality that would neither allow for reliable qualitative and quantitative assessments. The reviewer should provide comments indicating the reason(s) for selecting this category.

Reason(s) for reduced image quality may be related to e.g. patient preparation, patient body habitus, dose infiltration/contrast extravasation, missing anatomy, motion artifacts, attenuation correction artifacts (misregistration with PET emission scans, high density material affecting attenuation correction such as metal), other.

** more information on intensity and extent of the overall metabolic uptake can be found in “Ricard F., Cheson B., Barrington S., et al. 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). Under review.”

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

Issue	PRoLoG Recommendations	Comments
Required Image Series and Viewing Stations	• PET AC and NAC images • CTAC for attenuation correction and localization purposes • Reconstructed images (AC MIP and PET-CT fusion) • Diagnostic CT images with standard soft tissue and lung reconstruction algorithms • Appropriate viewing stations for image review and interpretation with multiplanar display • PET images should be scaled to a set SUV range and color table	• Allows for thorough evaluation of areas of FDG uptake and distribution • Helps to assess technical factors (e.g. poor attenuation correction, patient motion, lesion size, etc) that may affect image interpretation • Reconstructed images are required unless the viewing software enables to create such images from AC images. • CT of suitable image quality assists in lesion characterization and size measurement • Interpretations should be performed on viewing stations that allow multiplanar display of PET, CT and fused PET-CT images for better lesion identification,

		characterization and comparison of lesion FDG uptake to reference tissues
Body Coverage and Scan Visits	• The entire tumor burden needs to be evaluated • The direction of image acquisition should commence below the pelvis/ upper thigh region and conclude at the eye area but adjusted to ensure that all areas of disease are included • Use consistent patient positioning and breathing instructions across all imaging visits (instruction may differ for PET-CT and diagnostic CT)	• Good communication between ordering physician and image specialist is needed to ensure that all areas of known or suspected disease are visualized at each image visit • High FDG uptake in bladder can lead to reconstruction artifacts that may affect the visual assessment of lesion activity. This can be avoided by having the patients void prior to scanning and beginning image acquisition from pelvis • Consistent use of careful patient positioning for comfort and to avoid CTAC artifacts from arm positioning, body and respiratory motion is important to reduce technical

	• Use the same scanner for PET-CT throughout the study. Ideally, use the same CT scanner at each scheduled visit. • Avoid use of IV contrast and positive bowel contrast for CTAC • Use IV contrast for diagnostic CT. Oral contrast is helpful for hollow viscus evaluation • When both diagnostic CT and PET-CT are required, PET-CT with low dose CT for attenuation correction should be acquired first	changes in PET and CT signal between image visits and avoids PET-CT fusion errors • Consistent use of the same scanner for PET-CT and CT across each time point reduces scanner variability that could affect image assessments • IV and positive oral contrast during CTAC acquisitions can introduce attenuation correction errors (over correction) leading to technical alterations in FDG uptake that may affect visual interpretation • Use of IV contrast for diagnostic CT (oral contrast per site standard) is essential for accurate lesion measurement and characterization • Non-contrast low-dose CT for PET-CT being acquired before diagnostic CT avoids effect of contrast on semi-quantitative PET assessment

PET Acquisition and Image Reconstruction	• Ensure quality performance of PET and CT scanners per manufacture's and/or institutional recommendation or as per specific clinical trial recommendations • Ensure acquisition and reconstruction methods be kept rigorously consistent throughout the trial and between patient visits • PET 3D mode acquisition with ToF is preferred • Reconstruction methods that are not widely available and whose effect on the 5-PS is not yet known should be avoided in multi-institutional clinical trial settings • Rigorous compliance to ensure that the FDG uptake time is maintained for each imaging	• A well calibrated PET-CT scanner is needed for accurate measurement and representation of PET uptake • Phantom-based validation and harmonization of scanner reconstructions should be encouraged for studies relying on quantitative (SUV-based) assessments • Maintaining the use of the same scanner with identical image acquisition parameters and reconstruction methods is important to avoid introducing variability between imaging visits. Hence ideally all scans for a same patient should be done at the same institution throughout the trial. • The impact of newer techniques (e.g. PSF, regularized reconstructions, AI-based algorithms) for image acquisition and reconstruction and Lugano assessments is evolving but remains unknown and yet to be validated.

	visit, and that factors affecting SUV calculation that are manually entered are carefully checked and documented • Implement the comprehensive QIBA FDG profile as a guideline for standardization of the FDG PET workflow	• The uptake of FDG into tumor lesions and reference tissues is influenced by scan start time following FDG tracer injections. Longer FDG uptake periods can result in higher lesion and lower liver uptake which can affect visual scoring. An imaging timing window of +/- 5 minutes, up to +/- 10 minutes, compared to time used at baseline_is required and imaging should not start earlier than 55 minutes post FDG administration (55-75 min is acceptable) • Semi-quantitative SUV read-outs can be of interest in trials using the Lugano classification, therefore guidelines for standardization of FDG PET workflow are highly recommended
Technical Influence of Lesion Size and	• Further research to include the impact of lesion size and lesion background on the visual assessment of FDG PET is encouraged	• FDG uptake in smaller lesions can appear (and measure) lower than larger lesions given the same amount of radiotracer activity concentration. This is due to the

Background Activity		limitations of PET scanners to recover full radiotracer activity from smaller lesions (partial volume averaging) • Lesion conspicuity is influenced by background activity of reference tissues due to contrast resolution considerations which could affect the "apparent" visual assessments of FDG uptake
SUV measurements	• Semi-quantitative measurements can be captured (e.g. most hypermetabolic lesion, reference regions) and may be used in some instances to confirm visual assessment. • SUV that are captured usually represent maximum values (SUVmax) • Change in SUVmax and Metabolic Tumor Volume are promising tools for response evaluation and prognosis in lymphoma	• Other types of measurement such as SUVpeak or reference region SUVmean are frequently recorded in clinical trials • Change in SUV measurement used for response and prognosis evaluation at interim PET, as well as for assessment in PET-guided therapy is ongoing further validation

Imaging Reviewers Qualifications and Experience	• Reviewers (radiologist and/or nuclear physician) should meet minimum requirements to serve as a BICR • BICR with clinical trial read experience should still undergo training on Lugano criteria and modifications specific to the protocol prior to interpretation of on-study reads	• BICR should be board eligible or board certified Nuclear Medicine Physician (or the regional/national equivalent) with experience and/or certification in CT/MRI, or board eligible or board certified Radiology physicians with experience and/or training in PET-CT imaging. • BICR who have no prior experience in clinical trial reads should undergo a training program with test imaging cases to fulfill the minimum requirements • Close monitoring of on-trial performance is recommended for all reviewers.
Role of the Imaging Reviewer	• Same reviewer throughout the reads of all timepoints for a patient • Provides independent review of cases without bias or unblinding • Serve as a primary reader	• When possible, the same reviewers should review all timepoint for a patient and ideally a same reviewer should provide assessment of both FDG PET-CT and diagnostic CT.

	• Serve as an Adjudicator when applicable to resolve differences in time point responses	• Modality specific knowledge (i.e. CT and/or PET-CT) and experience with lymphoma imaging characteristics and response criteria evaluation is essential • The adjudicator should usually select the primary reviewer assessments that they most closely align based on the entirety of the reads (including access and review of all time points obtained prior to the adjudication event) rather than provide a third independent assessment. Alternative adjudication workflows exist, which are beyond the scope of this manuscript
Reviewers Training and Monitoring	• Training on Lugano classification (and any protocol specified modifications or clarifications), imaging CRF completion, familiarization of workstation usage and group review of clinical cases for	• Standards for equipment, workstation and monitors • Additional training cases and determination of statistically acceptable agreement rates amongst readers can be instituted on a per protocol basis

	formulating consensus on scan interpretation and time point responses are recommended activities prior to on-study reads • Minimum of 3 complete lymphoma cases should be considered for reader training	

AC: attenuation correction, AI: artificial intelligence, BICR: blinded independent central reviewer, CRF: case report form, CT: computed tomography, CTAC: CT for attenuation correction, FDG: fluorodeoxyglucose, IV: intravenous, MIP: maximum intensity projection, MRI: magnetic resonance imaging, NAC: non attenuation correction, PET: positron emission tomography, PSF point spread function, QIBA: Quantitative Imaging Biomarkers Alliance, SUV: standardized uptake value, ToF: time of flight, VS: visual score, 5-PS: 5 point-scale