

Response Criteria In Lymphoma The Core Lab's Perspective 26-MAY-2018 Ronald Korn MD, PhD Founder and Chief Medical Officer Imaging Endpoints Research and Core Lab



O Brief review of current lymphoma response assessment criteria

- Lugano, LyRIC and RECIL will be major focus
- Time does not allow for discussion of other criteria (IMWG, Olsen, etc)
- O Understand the evolution of criteria updates with more to come!
- Discuss challenges and strategies for Imaging Core Labs to adopt to new criteria

The Evolution of Lymphoma Assessment Criteria

Response Assessments In Lymphoma



IWG 1999 Criteria



Bruce Cheson M.D.

The 1999 criteria was developed under the direction of Bruce Cheson MD



Evolution of Response Assessments in Lymphoma



2014 Lugano Classification



Lugano, Switzerland

Consensus 2011-12 ICML reached agreement on PET evaluation in lymphoma trial including iPET and 5PS especially in HD/Aggressive NHL 2016 LYRIC Classification



Washington DC

One day conference to address the unique responses to IOT and recommend adaptations to the current criteria

2017 RECIL Classification



San Diego, California

Consensus based upon retrospective review of over 47,000 imaging measurements from 2983 patients in 10 clinical trials into a "RECIST" like measurement platform.

	Target Lesion	BL	TP1	TP2
AT AN	L1	20 mm	30 mm	40 mm
	L2	40 mm	30 mm	nromal
	L3	60 mm	40 mm	20 mm
	SLD	120	90 (SD)	60 (PR)



(Gordon) Moore's Law: 1965

7



More's Law



"Each year we will see *More* criteria"





- O Common Elements of Response Criteria
 - Radiologic: Structural vs Functional assessments
 - Tumor burden estimate at baseline follow up
 - New lesions
 - Organ (spleen) involvement
 - Bone Marrow status
 - Laboratory
 - Pathology
 - Clinical

3

- Improve patient evaluation and staging as lymphoma treatments evolve
- Eliminate ambiguity of previous criteria
- Facilitate the comparison of patients and results across studies
- Simplify the evaluation of new therapies by regulatory agencies
- Allow for atypical lesion behavior (tumor flare) due to novel therapies
- Converge with common focal point using RECIST core assessment criteria

2

Cheson Core Concepts

CT Assessments	What's New in Lugano?		
Target disease			
Non-target disease	Organomegaly Splean > 12cm is only	What's New in LYRIC?	
New lesions	 Spleen >13cm is enlar Liver size not assessed 	Lugano until PD	What's New in RECIL?
Spleen assessment		Indeterminate Response (IR)	Convergence towards RECIST
Liver assessment	• PET: 5 PS per timepoint		Up to 3 Target Lesions
PET Assessments	. Dono Marrow	Provisional and not mechanistic	 <u>></u>15mm in LD all lesion types avoiding long
Target disease/New Lesions M	 Bone Marrow Use PET for HL, DLBCL 	3 Categories of IR	narrow nodes in inguinal canal, portocaval and axillary
Bone Marrow Assessment	Biopsy for other histol	 IR1= Global increase within 12 IR2 = ≥1 lesion increase or nev 	Spleen measurements defined
Assessed by biopsy	Constitutional Symptoms:	without original SPD change > <u>!</u> response type pattern) at <u>any t</u>	
Constitutional Symptoms		 IR3 = increase in PET activity v 	diameter
Presence prevents CR	Overall Assessments	growth (size or number)	 Response incorporates CT and PET (avid histologies)
Overall Assessments	 PET trumps for avid his 		Minor Response: -10 to -30 % change in
CR, PR, SD, PD, Not Evaluable	CT for non-avid histolo	IP2 now locion added to SPD	SLD from baseline without new lesions and <i>Negative</i> PET
	single lesion growth ca PD if due to tumor	increase	A negative PET must be associated with at least a degraded in SLD by 20% for CB
		 IR3 increase in size or new lesic Encourages Biopsy 	 least a decrease in SLD by 30% for CR When LN <10mm use "0 mm"
	Referenced 1395		
		Referenced 528	Encourages waterfall plots for assessment
			Referenced 2

Summary of Core Criteria Changes

Cheson Core Concepts

CT Assessments		
Target disease	6 Lesions Measured bi-dimensionally	
Non-target disease	Followed qualitatively	What's New in Lugano?
New lesions	Identified and followed qualitatively	
Spleen assessment	Qualitative	Spleen: Quantitative cranial to caudal measurement
Liver assessment	Qualitative	Liver: size no impact
PET Assessments		
Target disease/New Lesions	Must be PET positive by qualitative visual assessment	PET: 5 Point Scale assessment per timepoint
Bone Marrow Assessment		
Assessed by biopsy		Bone Marrow: Assessed by PET in FDG avid histologies
Constitutional Symptoms		Constitutional Symptoms: Not Applicable
Presence prevents CR		
Overall Assessments		Overall Assessments: Radiologic Responses
CR, PR, SD, PD, Not Evaluable	Based on lesion size changes (CT) and PET activity (negative for CR), PD based on SPD	(CT) – Single Lesion can be used for progression and Metabolic Responses (PET/CT)



LyRIC- A Response to Pseudo Progression (PsP)



- O Lymphoma Response to Immunomodulatory Therapy Criteria
- See PsP with novel lymphoma treatments
 - Agents induce flare reactions in lymphoma
 - O Lenalidomide, Rituximab, Brentuximab vedotin, Ibrutinib, CPI
- Workshop to recommend aligning PsP using the immune response criteria principles for solid tumors progression recognizing the following:
 - Lugano vs RECIST rules are different
 - Lymphomas are not solid tumors
 - Tumor mass is always abnormal whereas lymphoma is mainly dealing with RES
 O CR for RECIST is tumor disappearance while it is normalization in lymphoma
 - Lymphoma leads to Organomegaly and marrow infiltration while RECIST does not take account these two features
 - Confirmation of PD in 4 weeks in solid tumors vs 12 weeks in lymphoma

Introduction of Indeterminate Response

2

Indeterminate Response (IR)

- IR1 = Global growth of existing lesions and/or new lesions
 - o SPD > 50% within 12 weeks of therapy
 - No clinical deterioration
 - Must repeat within 12 weeks and if there is >10% growth from IR1SPD then PD
 - o "Global Tumor Burden Swelling"
- IR2 = New Iesion(s) or ≥ 1 Iesion growth in setting of general stability
 - o Total SPD (including new lesions) has increased < 50% from Nadir
 - Prevents a single lesion from causing progression
 - Mixed response
 - Occurs at any time
- IR3 = Increase in FDG uptake
 - o No concomitant increase in lesion size meeting criteria for PD
- Updates to IR determination
 - If IR is followed by PD then update date of IR to PD
 - If IR is followed by improvement then consider IR as PsP
 - If IR is still within 10% of SPD on follow up then keep following until either PD or Response

Summary of Core Criteria Changes



What's New in Lugano?

- Organomegaly
 - Spleen >13cm is enlarged
 - Liver size not assessed
- PET: 5 PS per timepoint
- Bone Marrow
 - Use PET for HL, DLBCL
 - Biopsy for other histologies
- Constitutional Symptoms: removed
- Overall Assessments
 - PET trumps for avid histologies
 - CT for non-avid histologies but single lesion growth can cause PD if due to tumor

What's New in LYRIC?

- Lugano until PD
- Provisional and not mechanistic
- Indeterminate Response (IR)
- 3 Categories of IR:
 - IR1= Global increase within 12 wks of C1D1
 - IR2 = >1 lesion increase or new lesion without original SPD change >50% (mixed response type pattern) at <u>any time</u>
 - *IR3* = increase in *PET* activity without tumor growth (size or number)
- PD on follow up (no later than 12wks) if:
 - IR1 further increase in SPD
 - IR2 new lesion added to SPD for >50% increase
 - IR3 increase in size or new lesions
- Encourages Biopsy

Response Criteria in Lymphoma (RECIL)



- IWG aim was to harmonize lymphoma response with RECIST
 - Younes et al Annals of Oncology 28: 1436–1447, 2017
 - Simplify application of response assessment
 - Evidence based upon large scale data analysis rather than expert opinion
- Hypothesis: Can unidimensional measurements replace SPD?
 - Tested on retrospective database (Adult and Peds)
 - Compare Sum Diameter (SLD vs SSD) to "SPD" using landmark outcomes of response and PFS
 - Prior pilot studies demonstrated that unidimensional analysis equivalent to standard criteria
 - > 47,000 measurements, 2900 patients, 10 clinical trial, Phase I-III, different histologies and treatments
 - Did not incorporate FDG PET
- O Results
 - $\sim 95\%$ agreement in outcomes with SLD c/w SPD
 - 3 TLs is the sweet spot
 - Use RECIST thresholds for response (-30% for PR, >20% for PD)
 - Minor Response categories may be useful since patients with MR and PR tend to have similar K-M curves in some lymphoma subtypes
 - LDi is as good as SDi for response and PFS



Younes et al Annals of Oncology 28: 1436-1447, 2017



Updates of Lymphoma Response Criteria- RECIL

Younes et al Annals of Oncology 28: 1436–1447, 2017

PRIMA N= 365 Follicular lymphoma Untreated Phase 3 R-chemo +/- Rituximab maintenance⁵

 Table-S1.
 Studies used as a source for collecting imaging measurements. Five trials were from first line treatments, and 5 were from treatment of relapsed/refractory lymphoma.

Study	Number of patients included in this cohort	Lymphoma Subtype	Prior treatment status	Study Phase	Study Treatment/References
COG AHOD0031	855	Hodgkin Lymphoma	Untreated	3	Randomized response adapted therapy with ABVE-PC ¹
EORTC-20021	24	DLBCL	Untreated	Randomized phase 2	Gem-RCHOP vs RCHOP ²
EORTC-20921	352	Indolent lymphoma (FL, SLL)	Untreated	3	Fludarabine vs CVP ³
JNJ-LYM2034	141	DLBCL (non-GCB)	Untreated	3	RCHOP vs VR-CAP ⁴
PRIMA	365	Follicular lymphoma	Untreated	3	R-chemo +/- Rituximab maintenance ⁵
EORTC-20981	443	Follicular lymphoma	Relapsed	3	CHOP vs RCHOP +/- R maintenance ⁶
CL1-78454-001	34	HL and NHL	Relapsed	1	Phase-I Abexinostat7
CL2-78454-001	61	HL and NHL	Relapsed	2	Phase-2 Abexinostat ⁷
Gilead 101-09	122	Indolent (FL and SLL, marginal zone, Waldenstrom)	Relapsed	2	Phase-I of Idelalisib ⁸
JNJ-LYM3001	586	Follicular lymphoma	Relapsed	3	Rituximab vs Rituximab + Bortezomib ⁹
TOTAL	2983				

COG, Children Oncology Group; EORTC, European Organisation for Research and Treatment of Cancer; DLBCL, Diffuse large B cell lymphoma; FL, Follicular lymphoma; SLL, Small lymphocytic lymphoma; HL, Hodgkin lymphoma; NHL, non-Hodgkin lymphoma; MCL, Mantle cell lymphoma



Source	Rank correlation		
	Short diameter vs. area	Long diameter vs. area	
COG	0.936	0.913	
EORTC-20021	0.949	0.971	
EORTC-20921	0.998	0.997	
EORTC-20981	0.931	0.945	
CL1-78454-001	0.969	0.957	
CL2-78454-001	0.919	0.928	
Gilead	0.949	0.954	
JNJ-LYM2034	0.931	0.957	
JNJ-LYM3001	0.960	0.951	
PRIMA	0.893	0.947	
All Trials	0.947	0.945	

Updates of Lymphoma Response Criteria-RECIL

- Selecting Lesions
 - Target Lesions
 - O Up to 3 lesions maximum
 - ≥15mm in LDi for nodes and extra nodal disease
 - Non Target Lesions
 - O Extra nodal disease similar to RECIST
 - o 10-14 mm for nodes (axis inferred to be SDi)
 - PET Activity
 - o Keep 5PS
 - Bone Marrow Assessment
 - o Keeps Lugano recommendation
 - Spleen
 - o Allows for measurements on coronal images







Younes et al Annals of Oncology 28: 1436–1447, 2017

Updates of Lymphoma Response Criteria- RECIL



- O Unique Response Category Highlights
 - Complete Response
 - PET alone can not drive CR because many novel agents alter metabolism without impacting tumor viability
 - \geq 30% decrease SLD (partial response) plus 5PS \leq 3 on FDG-PET
 - SLD \leq 30% with negative PET <u>should not</u> be CR unless in negative tissue biopsy
 - o If lymph node normalizes then record as 0 mm
 - Partial Response
 - Mixed responses are allowed where Lugano requires a PD, PET is positive (4-5)
 - Minor Responses
 - o New category
 - $o \geq$ -10 but <-30% in SLD
 - O Doesn't matter what's on PET- "irrespective of PET Scan Results"
 - SD
 - o <10% to increase <20% <u>when MR</u> is used
 - Standard RECIST range if MR is not used
 - PD
 - o RECIST rules except that new lesion needs to be >10mm LDi <u>regardless</u> of PET
 - Small new PET avid lesions should be followed to determine significance and/or biopsied



	% Change in sum of diameters of target lesions from nadir				
	CR	PR	MR ^a	SD	PD
% change from baseline	 Complete disappear- ance of all target le- sions and all nodes with long axis <10mm. ≥30% decrease in the sum of longest diam- eters of target lesions (PR) with normaliza- tion of FDG-PET 	≥30% decrease in the sum of longest diam- eters of target lesions but not a CR	≥10% decrease in the sum of longest diam- eters of target lesions but not a PR (<30%)	<10% decrease or ≤ 20% increase in the sum of longest diameters of target lesions	 >20% increase in the sum of longest diameters of target lesions For small lymph nodes measuring <15 mm post therapy a minimum absolute increase of 5 mm and the long diameter should exceed 15 mm Appearance of a new lesion
FDG-PET	Normalization of FDG- PET (Deauvile score 1- 3)	Positive (Deauville score 4-5)	Any	Any	Any
Bone marrow involvement	Not involved	Any	Any	Any	Any
New lesions	No	No	No	No	Yes or No

CR, complete response; CT, computerized tomography; FDG-PET, [¹⁸F]2-fluoro-2-deoxy-D-glucose; MR, minor response; PD, progression of disease; PR, partial response; SD, stable disease.

^aA provisional category.

Younes et al Annals of Oncology 28: 1436–1447, 2017

2

O Additional Highlights

- Progressive Disease After Initial Response
 - Measured from nadir not baseline and growth of new LN to >15mm and 5mm absolute
 - Ex: CR with a 9mm node (normal) that grows to 16 mm would lead to PD as a single lesion
- Response Assessment in with Immune Modulating Treatments
 - "To account for potential 'pseudoprogression,' immune-related response criteria should be used, requiring confirmation of progressive disease on two consecutive scans at least 4 weeks apart and inclusion of new lesion measurements in the total tumor burden"

• Appearance of New Extranodal Lesion

- A minimum of 1 cm in largest diameter of new extranodal lesions is required to confirm progressive disease.
- New smaller but suspicious lesions should be designated as equivocal; if later confirmed (by CT or biopsy) as due to lymphoma, the documented date of disease progression should be the date of identification as equivocal.

2

O Additional Comments

- Disseminated Disease
 - The status of nontarget lesions should be taken into account before formulating the final response status
- **O** Frequency of Response Assessment
 - In phase I/II clinical trials in previously treated patients
 - every 2 to 3 months x 1 year, then every 3 to 4 months in year 2, then every 6 months in year 3 until end of trial
 - In randomized phase III imaging may be less frequent without recommendation to timing

Determining Response and Progression



Example: Zeroing out normalized Nodes





Subject Evaluated using Lugano and RECIL. A total of 6 lesions (SPD) were included in Lugano assessment whereas 3 lesions with longest diameter (SLD) were included in the RECIL assessment. Although the largest target lesions were different in the two calculations the category of response was the same as was the time to progression when measured from Nadir. The RECIL estimate shown above includes absolute nodal size in SLD. In determining SLD TL1, 2 and 4 from the 6 TL used in Lugano had the largest LDi and were selected as RECIL TL.

Subject Evaluated using Lugano and RECIL. A total of 6 lesions (SPD) were included in Lugano assessment whereas 3 lesions with longest diameter (SLD) were included in the RECIL assessment. Although the largest target lesions were different in the two calculations the category of response was the same as was the time to progression. The RECIL estimate shown above removes normal size node measurements (0 mm) from SLD

	RECIST 1.1	Lugano	RECIL 2017
Number of target lesions	Up to 5	Up to 6	Up to 3
Measurement method	Uni-dimensional: long diameter of non-nodal lesions, short diameter of lymph nodes	Bi-dimensional: perpendicular diameters	Uni-dimensional: long diameter of any target lesion
Incorporates PET results to describe CR	May be considered to confirm CR and/or to declare PD based on detecting new lesions	Yes	Yes
Minor response	No	No	Yes, reduction in sum of long diam- eters between ≥10% and <30%
Stable disease	-29% to + 20%	-50% to $+50%$	decrease <10% to increase \leq 20%
PD	Increase in sum of diameters by 20%	Increase in the sum of products of perpendicular diameters by > 50%, or any single lesion by > 50%	Increase in sum of the longest diameters by 20%. For relapse from CR, at least one lesion should measure 2 cm in the long axis with or without PET activity

CR, complete response; PD, progression of disease; PET, positron emission tomography.

Younes et al Annals of Oncology 28: 1436–1447, 2017

3

- Want to use the best criteria to determine treatment efficacy
 - Histology, mechanism of action, patient population, toxicity profile, atypical behaviors, early vs late stage, prior response evidence
- O Simplicity is preferred
 - Allows for more straight forward process mapping, coding and analysis and interpretation
 - Allows for better training
 - Provides more consistency in reads with lower adjudication rates
 - Less site vs central discordance
 - But the more complex a criteria is the more important it is to use Core Lab for BICR
- O Quantitative vs Qualitative
 - If qualitative then consistency and standardization in acquisition is paramount especially in PET
 - If quantitative assessments are primary goals then standardization, harmonization and calibration are important
 - Support QIBA standards!
- Want to avoid modification(s) to criteria
 - Reduces criteria effectiveness making comparisons with historical or contemporaneous data difficult
 - May not have been field tested yet
 - More challenging for implementation

Criteria Challenges



- General Challenges
 - Sponsors tend to "cut and paste" criteria into protocols without understanding downstream implications
 - Words matter and ambiguity can causes uncertainty in implementation
 - "When things go missing" Avoid NE
 - CT is non diagnostic but PET is adequate, alternating diagnostic CT with PET, etc
 - Can be handled with logistical (what ifs) scenarios upfront and train, train, train...
 - When there is "too much information"- Avoid unblinding
 - Additional or unscheduled exams can lead to unblinding of treatment arm or lead to premature PD
 - Ex: Engineered T-cell therapy patients get MRI or early imaging which can show telltale signs of CRS and Neurotoxicity events
 - Must include instructions for incorporating studies into response assessments

Criteria Challenges

- General Challenges
 - When things go in "different direction" Avoiding ambiguity
 - o CT and PET responses do not align
 - PET should trump CT in HD and DLBCL but other types CT rules
 - Making the final call?
 - Radiologist, Clinician, End Points Committee
 - Integration of molecular profiling information
 - "When organs are involved" Understand variants
 - What if you are 7'3"? Cirrhosis?
 - Unclear as to whether central readers can use judgement in regards to the 13 cm cut off (e.g. due to patient size or irregular spleen shape etc.)
 - RECIL provides good guidance for spleen but measure on MIPs, really?











O "When to measure"

- SUVs: To Capture or Not to Capture?
- Lugano doesn't address SUVs but Barrington (2014) provides guidance on SUVs to help with this assessment

OMTV vs TLG vs Radiomics or other types of analysis

- Size estimates on PET?
- Good idea?

Criteria Challenges

• When "publications errors arise" what to do?

• Example:

Progressive disease Individual target nodes/nodal masses Extranodal lesions

Progressive metabolic disease

Score 4 or 5 with an increase in intensity of uptake from baseline and/or

New FDG-avid foci consistent with lymphoma at interim or end-of-treatment assessment Cheson et al

Table 3. Revised Criteria for Response Assessment				
Response and Site	PET-CT-Based Response	CT-Based Response		
Complete	Complete metabolic response	Complete radiologic response (all of the following)		
Lymph nodes and	Score 1, 2, or 3" with or without a residual mass on SPS1	Target nodes/hodal masses must regress to ≤ 1.5 cm in LI		
'extralymphatic sites	It is recognized that in Waldwyr's ring or extranodal stles with high physiologic uptake or with activation within splean or marrow log, with charaotherapy or myeloid colony stimulating factors, uptake may be grader than normal mediastimum and/or liver. In this circumstance, complete metabolic response may be informed if uptake at stits of initial involvement is no groupot than surrounding normal tissue even if the tissue has high physiologic uptake	No extralymphatic sites of disease		
Nonmeasured lesion	Not applicable	Absent		

Progressive disease requires at least 1 of the following PPD progression:

An individual node/lesion must be abnormal with: LDi > 1.5 cm and Increase by ≥ 50% from PPD nadir and An increase in LDi or SDi from nadir 0.5 cm for lesions ≤ 2 cm 1.0 cm for lesions > 2 cm In the setting of splenomegaly, the splenic length must increase by > 50% of the extent of its prior increase beyond baseline (eg, a 15-cm spleen must increase to > 16 cm). If no prior splenomegaly, must increase by at least 2 cm from baseline



Criteria Challenges-LYRIC



- IR(1): "Increase in overall tumor burden (as assessed by sum of the product of the diameters [SPD]) of >50% of up to 6 measurable lesions in the first 12 weeks of therapy, without clinical deterioration"
 - Requires knowledge of timepoint dates (or at least weeks between time points) which central radiologists are typically blinded to.
 - Could handle through global type form/review, or education of sponsor on how to derive this response
- IR(3): "Increase in FDG uptake of 1 or more lesion(s) without a concomitant increase in lesion size or number"
 - Lack of clear guidance for central readers on what constitutes increase in FDG uptake (e.g. does this require a change in Deauville score or just a change in visual estimation of FDG uptake?)
 - Lack of guidance on what increase in lesion size means (i.e. what if there is a small increase but lesions still qualify as stable?)
 - Potential to lead to high reader variability



O Newest criteria is currently being field tested

- O High quality diagnostic CT becomes a must
- Thresholds for response and progression and size criteria for TL keeps on changing
 - PET responses have to carefully considered in assessment of outcomes
 - Recommendation to use MIP images to measure spleen has not been rigorously evaluated to our knowledge
 - Reading past PD is allowed but 4 weeks repeat recommended which timing is different from Lugano

Challenges of Updates to Lymphoma Criteria



• FAQs from Sponsor

- Experience Questions
 - o "What is the agency position on x criteria?"
 - "My phase 2 trial was a success using previous criteria but should I switch to new criteria"
 - Strategy:
 - "Talk to them" especially if it will be used for clinical treatment decisions
 - Include as secondary or exploratory endpoints
 - Road test on prior trials
 - Switching to new criteria while still in active investigation is perilous
- Modality Questions
 - o "Can I use the CT from PET or do I need a separate diagnostic scan?"
 - CTs from PET scanner is becoming commonplace and acceptable if truly diagnostic
 - 10-40% variability in size measurements
 - "Can I use PET/MRI?"
 - Rethink lesion assessments especially and Bone Marrow involvement and it is unclear how that will impact lymphoma studies
 - O Digital PET scans with Recovery Coefficient corrections is on the horizon
 - Need to plan for integration into clinical trial read outs
 - Impact on quantitation thresholds or Tumor : Reference ratios unknown



- We are in exciting times and cure may be on the horizon
- We don't work in a static environment and progress is rapid
- We have to be adaptive but we must be rigorous in our regulatory control and compliance
- Robust QMS are a critical element of a core lab
- Imaging will become more elegant, important and complex and we must prepare for that future
- Pintad and others play an crucial role in helping to lead the conversation which then helps to set the standards
- O Acknowledge: Annette Schmidt, Paul Galette and Julie Gillis

Thank You

9700 N. 91st Street, STE B-200 Scottsdale, Arizona 85258

460 Totten Pond Road, STE 420 Waltham, Massachusetts 02451

Phone 480.314.3070 www.ImagingEndpoints.com