

Criteria-Based Imaging and Response Evaluation of Lymphoma 20 Years After Cheson: What is New?

A Review of the Current Classifications

Kriterien-basierte Bildgebung und Responsebeurteilung bei Lymphomen 20 Jahre nach Cheson: Was gibt es Neues?

Eine Übersicht zu den aktuellen Klassifikationen

Authors

Christopher Skusa¹, Marc-André Weber¹, Sebastian Böttcher², Kolja M. Thierfelder¹

Affiliations

- 1 Institute of Diagnostic and Interventional Radiology, Pediatric Radiology and Neuroradiology, Rostock University Medical Center, Rostock, Germany
- 2 Department of Medicine, Clinic III – Hematology, Oncology, Palliative Medicine, Rostock University Medical Center, Rostock, Germany

Key words

lymphoma, CT, PET-CT, staging

received 07.09.2019

accepted 10.12.2019

Bibliography

DOI <https://doi.org/10.1055/a-1091-8897>

Published online: 26.3.2020

Fortschr Röntgenstr 2020; 192: 657–667

© Georg Thieme Verlag KG, Stuttgart · New York

ISSN 1438-9029

Correspondence

Dr. Christopher Skusa

Institute of Diagnostic and Interventional Radiology, Pediatric Radiology and Neuroradiology, Rostock University Medical Center, Ernst-Heydemann-Str. 6, 18055 Rostock, Germany
Tel.: ++49/381/4949201
skusis@outlook.de

ABSTRACT

Background The rapid progress in oncology research requires numerous new scientific publications. This article aims to provide an overview of criteria-based imaging and response evaluation of lymphoma according to the current status of knowledge. In fact, common criteria for evaluating data, especially imaging response evaluation, are essential for comparability of studies. While criteria-based classifications of solid tumors have been established for some time, there are now increasing classifications of lymphoma diseases. The purpose of this review is to describe the devel-

opment of criteria-based evaluation of lymphoma diseases with a special focus on imaging up to current guidelines.

Methods Literature review based on PubMed including the languages English and German was performed. This review article includes the most important criteria-based response evaluations of lymphoma published between January 1999 and July 2019.

Results and Conclusion The two latest classifications of response evaluation of lymphoma are: The Lugano classification, which has been steadily developed over the past 20 years and has been specially adapted to technical progress, as well as the evaluation method RECIL (Response Evaluation Criteria In Lymphoma), which is based on the RECIST (Response Evaluation Criteria in Solid Tumors) classification already established for solid tumors. Significant imaging components of both classifications are the anatomical measurement and measurement of the metabolic response of the manifestation of lymphoma using positron emission tomography (PET/CT).

Key Points:

- Standardized criteria-based response evaluations are essential for the objective and comparable analysis of new drugs for the treatment of lymphoma diseases.
- The latest classification RECIL has significantly simplified treatment evaluation and has established a better comparability to the therapeutic evaluation of solid tumors according to RECIST.
- Further studies will show the most appropriate classifications depending on study settings.

Citation Format

- Skusa C, Weber M, Böttcher S et al. Criteria-Based Imaging and Response Evaluation of Lymphoma 20 Years After Cheson: What is New?. Fortschr Röntgenstr 2020; 192: 657–667

ZUSAMMENFASSUNG

Hintergrund Der schnelle Fortschritt in der onkologischen Forschung bedingt eine Vielzahl neuer wissenschaftlicher Publikationen. Dieser Beitrag möchte daher einen Überblick

über den aktuellen Wissensstand bezüglich der Kriterien-basierten Bildgebung und Responsebeurteilung bei Lymphomen liefern. So sind gemeinsame Kriterien in der Auswertung von Daten, insbesondere der bildgebenden Responseevaluation, für die Vergleichbarkeit der Studien essenziell. Während Kriterien-basierte Klassifikationen für solide Tumoren bereits länger etabliert sind, gibt es nun zunehmend auch Klassifikationen für Lymphomerkrankungen. Als Ziel dieser Übersichtsarbeit soll die Entwicklung der Kriterien-basierten Auswertung von Lymphomerkrankungen mit besonderem Schwerpunkt auf die Bildgebung bis hin zu den derzeit gültigen Leitlinien beschrieben werden.

Methode Die Literaturrecherche erfolgte in PubMed in den Sprachen Englisch und Deutsch. Dieser Übersichtsartikel enthält die wichtigsten Kriterien-basierten Responseevalua-

tionen von Lymphomerkrankungen, die von Januar 1999 bis Juli 2019 publiziert wurden.

Ergebnisse und Schlussfolgerung Es gibt aktuell 2 Klassifikationen zur Evaluation des Therapieansprechens von Lymphomerkrankungen: Die seit 20 Jahren stetig weiterentwickelte, insbesondere dem technischen Fortschritt angepasste Lugano-Klassifikation sowie die Auswertmethode RECIL (Response Evaluation Criteria In Lymphoma), die sich an der bereits für solide Tumoren etablierten RECIST (Response Evaluation Criteria in Solid Tumors)-Klassifikation orientiert. Die wesentlichen Bestandteile beider Klassifikationen sind die anatomische Messung von Lymphomanifestationen in der Bildgebung und die Messung der metabolischen Response des Tumorgewebes durch die Positronen-Emissions-Tomografie (PET)/CT.

Introduction

Lymphoma treatment is constantly progressing. New treatment approaches utilize antibodies, immunotherapeutics, and antibody-drug conjugates. To be able to provide optimal treatment for every patient, an individualized treatment approach is needed [1]. In addition to clinical parameters, oncological imaging is often decisive for evaluating the effectiveness of new therapeutic agents. The goal of criteria-based evaluation of imaging is to ensure objective, standardized, comparable, and uniform evaluation during and after treatment. This review article describes the development as well as the current state of follow-up of Hodgkin and Non-Hodgkin lymphoma. The integration of standardized morphological imaging with metabolic response (PET/CT) and highly sensitive methods for detecting residual cells in blood and bone marrow (minimal residual disease, MRD) is particularly important. On the whole, there is a trend from purely expert-based consensus criteria to criteria based on study data.

Cheson criteria (IWG criteria) [2]

While there has been a committee-based classification for Hodgkin lymphoma (HL) since 1989 [3], the first generally recognized standardized classification of non-Hodgkin lymphoma (NHL) was the International Working Group (IWG) classification published by Cheson et al. The classification focused on anatomical measurements of the manifestations of lymphoma. [2]

For the total response evaluation, the Cheson classification takes clinical and laboratory criteria as well as bone marrow biopsy results into consideration in addition to the radiological criteria. Computed tomography (CT) is recommended as the imaging modality of choice. Chest, abdominal, and pelvic CT examination is recommended. CT scan of the neck is typically also performed in clinical application although this examination was not required in the original publication. The following parameters are relevant for the radiological evaluation and corresponding classification (► Fig. 1):

- Great transverse diameter (GTD): Greatest lesion dimension in the transverse section.
- Short axis diameter (SAD): Shortest lesion dimension in the transverse section, orthogonal to the GTD. A lymph node > 1 cm in the SAD is considered suspicious. [4–8]
- Sum of the product of the diameters (SPD): The product of the GTD and SAD from a maximum of six “dominant” nodal manifestations is calculated. These products are then totaled. The selected lesions should be distributed in the body to the greatest extent possible and at least one lesion from the retroperitoneum and one from the mediastinum should be selected if present. Lesions in other organs outside the liver and spleen can be detected but not measured as a dominant lesion.

The results of every follow-up imaging examination are classified into one of the six response categories. In addition to the SPD and the occurrence of any new lesions, clinical parameters are essential here. For the first Cheson classification, the size of the spleen and liver was characteristically determined only by clinical examination and not imaging. Complete response unconfirmed (CRu) is a special category. The individual response categories of the Cheson classification are provided in ► Table 1.

Revised Cheson criteria [9]

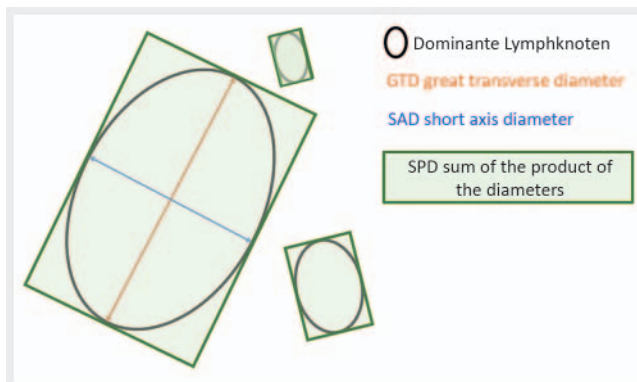
The Revised Response Criteria for Malignant Lymphoma published in 2007 was a major revision of the Cheson criteria [9]. In particular, technical advancements in the field of hybrid imaging were taken into consideration. The main changes were:

- The treatment of Hodgkin lymphoma is now also evaluated according to the Cheson criteria. The classification according to Lister et al. was previously used [3].
- The “complete response unconfirmed” (CRu) category was eliminated.
- 18F-FDG-PET or PET/CT was integrated in the criteria as a functional-metabolic part of imaging

► **Table 1** Response evaluation according to the criteria of the International Working Group (IWG) (Cheson classification).

	CR	CRu	PR	SD	RD after CR/CRu	PD
general	<ul style="list-style-type: none"> disappearance of all radiological signs of the disease 		<ul style="list-style-type: none"> no new lesions 	<ul style="list-style-type: none"> reduction <PR, PD criteria not fulfilled 	<ul style="list-style-type: none"> new lesions 	<ul style="list-style-type: none"> new lesions
lymph node	<ul style="list-style-type: none"> previously > 1.5 cm to ≤ 1.5 cm GTD previously > 1.0 cm and ≤ 1.5 cm to ≤ 1.0 cm GTD or reduction of SPD > 75 % 	<ul style="list-style-type: none"> lymph node > 1.5 cm GTD despite reduction of SPD > 75 % separation of nodal masses with reduction of SPD > 75 % 	<ul style="list-style-type: none"> reduction of SPD ≥ 50 % of the 6 most dominant lymph nodes or nodal masses no increase in size of other lymph nodes 		<ul style="list-style-type: none"> increase in the GTD ≥ 50 % of an existing lymph node > 1.0 cm SAD or increase in SPD ≥ 50 % in multiple lymph nodes/nodal masses 	<ul style="list-style-type: none"> increase in SPD ≥ 50 % of the target lesions
extranodal manifestation	<ul style="list-style-type: none"> spleen normalized and lesion-free other organs enlarged prior to therapy have returned to normal size 		<ul style="list-style-type: none"> no increase in the size of the liver, spleen reduction of SPD ≥ 50 % of the lesions in the spleen and liver 			

CR: complete response, CRu: complete response unconfirmed, PR: partial response, SD: stable disease, RD: relapsed disease, PD: progressive disease, GTD: great transverse diameter, SPD: sum of the product of the diameters.



► **Fig. 1** Schematic drawing of the anatomical parameters (in this example, a total of three “dominant” nodal manifestations were chosen to calculate the SPD. This matches with the sum of all green rectangles).

However, PET examination should only be performed in the case of lymphoma with sufficient avidity for ¹⁸F-FDG. Known cases of PET-avid lymphoma include diffuse large B-cell lymphoma (DLBCL), Hodgkin lymphoma (HL), follicular lymphoma (FL) and mantle cell lymphoma (MCL), while chronic lymphatic leukemia (CLL), for example, is a lymphoma with only minimal FDG uptake. To date, FDG uptake in other types of Non-Hodgkin lymphoma has been examined in a variable manner or only minimally. The PET scan is categorized qualitatively only a purely visual basis as positive or negative. The PET scan is categorized as positive when

the focal or diffuse uptake of ¹⁸F-FDG is higher than the background activity in a location that cannot be explained purely anatomically/physiologically.

In the case of FDG-avid lymphoma, CR is achieved when post-treatment residual lesions are PET negative – regardless of their size. In patients without a pretreatment PET/CT examination, the size of the lymph nodes is decisive as in the Cheson criteria. Original lesions with a GTD > 1.5 cm must have a GTD ≤ 1.5 cm after treatment. Lymph nodes with a GTD of 1.1–1.5 cm and an SAD > 1.0 cm must have an SAD ≤ 1.0 cm after treatment. The size of the spleen and liver must be normal (again) and the spleen must not have any lesions. The morphological imaging criteria as well as the bone marrow finding are also taken into consideration in the revised Cheson criteria. As a result, despite radiological classification as CR, the treatment response can still be classified as PR if corresponding morphological changes in the bone marrow finding are seen. The additional categories are summarized in ► **Table 2**.

Lugano classification [10, 11]

A further revision of the Cheson criteria was published in 2014 [10]. This is the currently widely used “Lugano” classification of treatment response in lymphoma. The most important changes regarding imaging were:

- Replacement of the dichotomous PET evaluation with the Deauville 5-point scale (5-PS)
- Introduction of the interim PET scan during treatment

► **Table 2** Response classification according to the Revised Response Criteria for Malignant Lymphoma (Revised Cheson).

	CR	PR	SD	PD
CT	<ul style="list-style-type: none"> lymph node > 1.5 cm to ≤ 1.5 cm GTD lymph node > 1.0 cm and ≤ 1.5 cm GTD and > 1.0 cm SAD to ≤ 1.0 cm SAD spleen/liver: <ul style="list-style-type: none"> spleen and liver normalized and lesion-free 	<ul style="list-style-type: none"> reduction of SPD ≥ 50 % of the max. 6 most dominant lymph nodes no increase in other lesions and no new lesions spleen/liver: <ul style="list-style-type: none"> reduction of SPD ≥ 50 % in multiple lesions reduction of GTD ≥ 50 % in solitary lesions no increase in size 	<ul style="list-style-type: none"> no increase in size of existing lesions 	at least 1 criterion met: <ul style="list-style-type: none"> new lesions > 1.5 cm regardless of axis increase in SPD ≥ 50 % of existing lesions in case of SAD > 1.0 cm: Increase in GTD ≥ 50 % in case of SAD ≤ 1.0 cm: increase in SAD ≥ 50 % to a minimum size of > 1.5 cm × 1.5 cm or GTD > 1.5 cm increase in SPD ≥ 50 % of spleen/liver lesions
PET in FDG-avid or PET(+) lymphomas	<ul style="list-style-type: none"> negative without size limitation 	<ul style="list-style-type: none"> positive existing lesion(s) 	<ul style="list-style-type: none"> positive existing lesion(s) and no new lesions 	<ul style="list-style-type: none"> positive

CR: complete response, PR: partial response, SD: stable disease, PD: progressive disease, GTD: great transverse diameter, SAD: short axis diameter, SPD: sum of the product of the diameters, FDG: fluorodeoxyglucose, PET: positron-emission tomography:

The 5-PS is a purely visual qualitative assessment that allows a more differentiated classification than was possible with the previous method of evaluating FDG uptake. The criteria for the 5-PS are summarized in ► **Table 3** [12–14].

The interim PET examination makes it possible to adjust treatment earlier in some cases. Therefore, it is possible to respond more quickly in the case of a lack of treatment response so that corresponding treatment escalation can be initiated or, in the case of early treatment response, treatment can be deescalated to minimize toxicity and secondary diseases [15–18]. The German Hodgkin lymphoma study group was recently able to convincingly prove this concept in the HD18 study: treatment de-escalation on the basis of the interim PET resulted in better overall survival due to lower toxicity [19]. In the recently published PETAL study, the prognostic significance of the interim PET scan for diffuse large B-cell lymphoma was also convincingly shown even if an attempt to escalate treatment with a more aggressive protocol for PET-positive patients was not successful [20]. Studies were able to confirm the advantages of 5-PS compared to the previously recommended dichotomous PET evaluation [21–23].

The “product of the perpendicular diameters” (PPD) is introduced in the anatomical evaluation as a new term. It is the product of the orthogonal diameter, i. e., the product of the GTD and the SAD. While the PPD relates to an individual lesion, the SPD is the sum of multiple PPDs. In prior classifications, an increase in the PPD of an individual lesion could be offset in the sum, the SPD, by the absent or smaller increase in size of the other lesions. The PPD is particularly important for the PD classification. An increase in the size of an individual lesion by ≥ 50 % in the PPD can now result in classification as progress. For categorization as SD or PR, a change in the SPD is still the decisive parameter.

A maximum of six dominant lesions are still measured. If possible, the measurement should include a retroperitoneal and a mediastinal lesion. The definition of (extra-) nodal lesions was

► **Table 3** Deauville five-point scale (5-PS).

score	criterion
1	no ¹⁸ F-FDG uptake
2	uptake ≤ mediastinum
3	uptake > mediastinum and ≤ liver
4	uptake moderately above liver level
5	uptake significantly above liver level and/or new lesion(s)
X	new uptake region(s), most likely not part of lymphoma

FDG: fluorodeoxyglucose.

simplified. Nodal lesions with a diameter > 1.5 cm and extranodal lesions > 1.0 cm are now considered measurable regardless of the axis. An SAD > 1.0 cm is no longer a requirement for defining a measurable lesion.

Enlargement of the spleen was defined more precisely. Per definition, spleen involvement is present starting at a vertical diameter > 13.0 cm. The percentage change in the enlargement of the spleen is monitored during and after treatment. A decrease in a vertical spleen diameter over the limit of 13 cm by > 50 % is classified as PR. A spleen with a vertical diameter of 18 cm is 5 cm over the normal size (5 cm = absolute spleen enlargement). A decrease in the spleen diameter after treatment to 15 cm thus means a reduction in the spleen enlargement of 60 % (3 cm/5 cm * 100), corresponding to PR. It is important that only a spleen that is enlarged above the limit value of 13 cm needs to be evaluated. A reduction in the size of the spleen only from 18 cm to 16 cm, i. e., a 40 % decrease in the spleen enlargement (2 cm/5 cm * 100), can at best be classified as SD.

► **Table 4** Response evaluation according to Lugano.

	CR	PR	SD	PD
CT: radiological response	<p>all criteria met:</p> <ul style="list-style-type: none"> no new lesions reduction ≤ 1.5 cm GTD of target lesions no extranodal manifestation <p>spleen/liver:</p> <ul style="list-style-type: none"> spleen and liver normalized <p>new lesions: None</p>	<ul style="list-style-type: none"> reduction of SPD of the max. 6 measurable target lesions (nodal/extranodal) $\geq 50\%$ lesions that are too small/not measurable are assigned a size of 0.5×0.5 cm lesions that are not visible are assigned a size of 0×0 cm <p>spleen/liver:</p> <ul style="list-style-type: none"> length reduction $> 50\%$ of the enlarged spleen <p>new lesions: none</p>	<ul style="list-style-type: none"> reduction of SPD of the max. 6 measurable target lesions (nodal/extranodal) $< 50\%$ PD criteria not met <p>new lesions: none</p>	<p>at least 1 criterion met: abnormal lymph nodes must meet the following criteria:</p> <ul style="list-style-type: none"> GTD > 1.5 cm and PPD increase $\geq 50\%$ and increase of GTD or SAD by <ul style="list-style-type: none"> 0.5 cm for ≤ 2.0 cm 1.0 cm for > 2.0 cm <p>spleen/liver:</p> <ul style="list-style-type: none"> in the case of existing splenomegaly: Increase in the size of the spleen $> 50\%$ in the case of newly occurring splenomegaly: Increase in the length of the spleen ≥ 2.0 cm with respect to baseline <p>new lesions:</p> <ul style="list-style-type: none"> recurrence of lesions that had already returned to the normal range or new nodal lesion > 1.5 cm regardless of the axis or new extranodal lesions > 1.0 cm regardless of the axis or ≤ 1.0 cm in the case of clear assignment to lymphoma
PET/CT: metabolic response	<ul style="list-style-type: none"> score 1/2/3 according to 5-PS without residual disease <p>new lesions: none</p>	<ul style="list-style-type: none"> score 4/5 according to 5-PS with reduced ^{18}F-FDG uptake with respect to baseline and residual lymph nodes of any size <p>new lesions: none</p>	<ul style="list-style-type: none"> score 4/5 according to 5-PS <p>new lesions: None</p>	<ul style="list-style-type: none"> score 4/5 according to 5-PS with increase in ^{18}F-FDG uptake with respect to baseline <p>new lesions: New ^{18}F-FDG-avid typical lymphoma lesions</p>

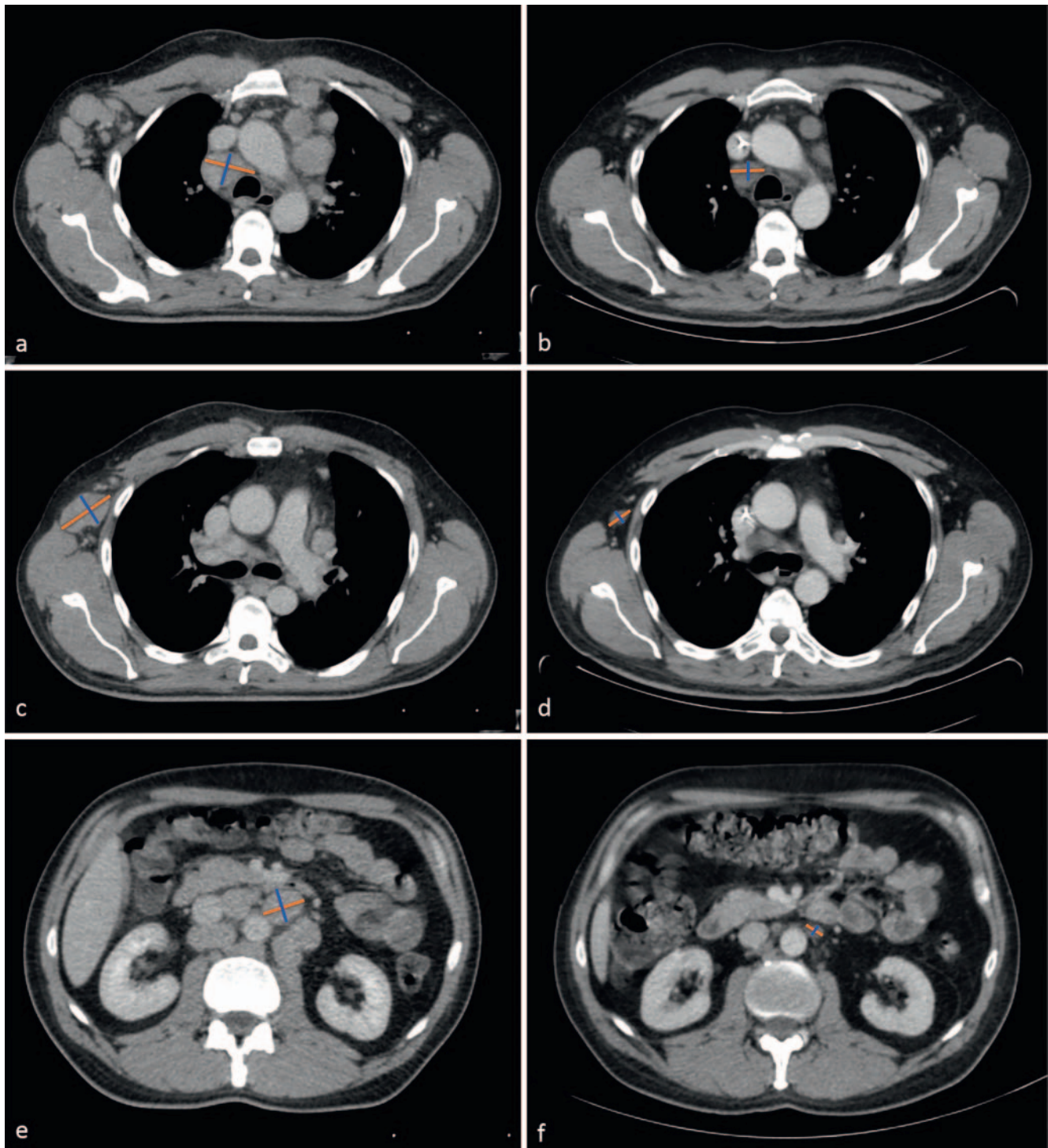
CR: complete response, PR: partial response, SD: stable disease, PD: progressive disease, GTD: great transverse diameter, SAD: short axis diameter, SPD: sum of the product of the diameters, PPD: product of the perpendicular diameters, FDG: fluorodeoxyglucose, PET: positron-emission tomography, 5-PS: Deauville 5-point scale.

Previously measured lymph node conglomerates (bulky disease) that separate over the course of treatment are handled slightly differently. The PPDs of the individual lesions are now calculated. These separated individual lesions are also part of the SPD in the further course even if that means that the maximum number of 6 is exceeded [11]. The response criteria according to Lugano are summarized in ► **Table 4**.

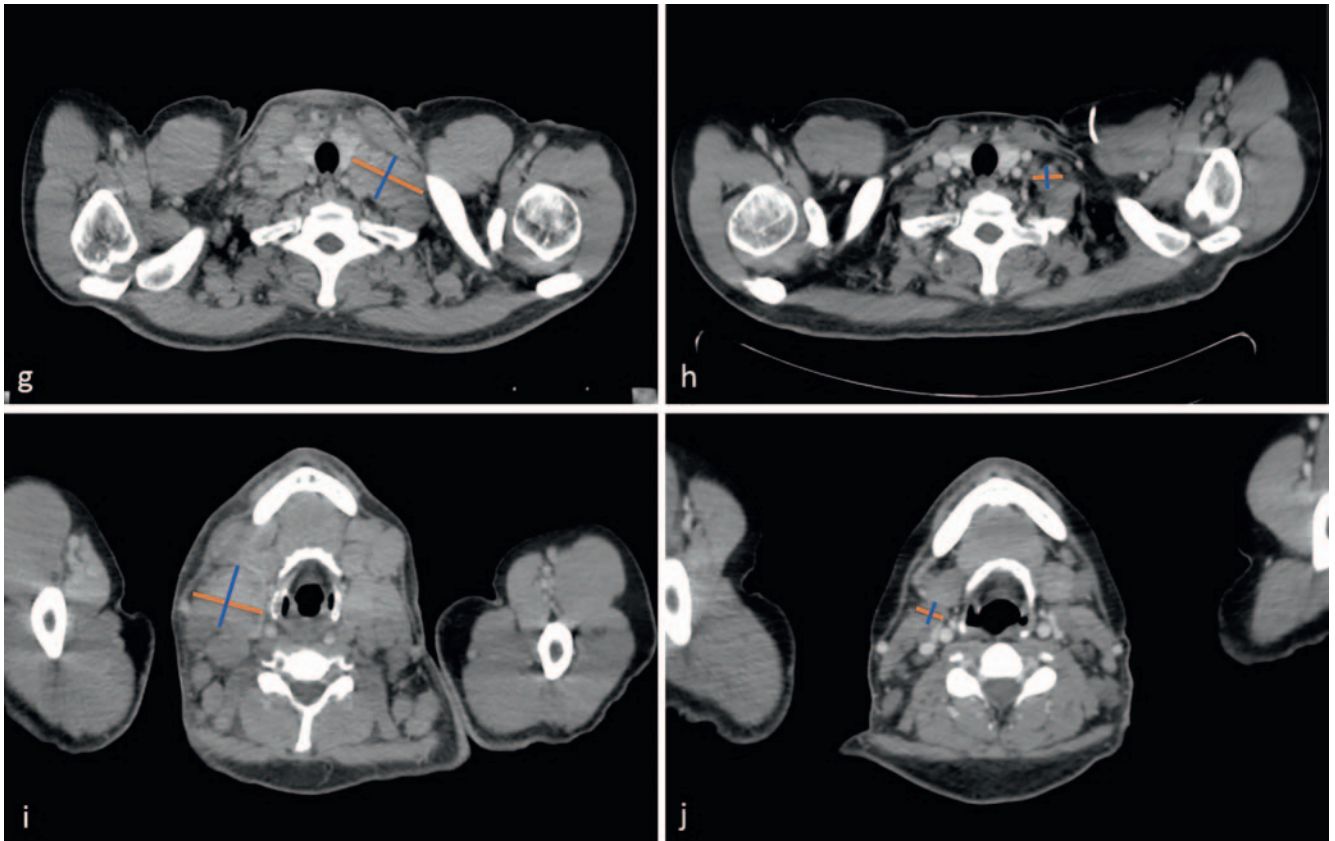
The Lugano classification is primarily used in cases of malignant lymphoma with nodal involvement. However, as a rule, it can be used for all malignant lymphomas, even primary extranodal diffuse large B-cell lymphoma. The Response Criteria for Primary CNS Lymphoma [24] published by Abrey et al. are used for CNS lymphomas, the European Society for Oncology (ESMO)

Clinical Practice Guidelines for diagnosis, treatment and follow-up are used for MALT (Mucosa-Associated Lymphoid Tissue) lymphomas [25], and the Clinical End Points and Response Criteria in Mycosis Fungoides and Sézary Syndrome published by Olsen et al. are used for cutaneous lymphomas [26]. The recently published response criteria for chronic lymphatic leukemia are also based on the Lugano classification [27].

An example of an evaluation according to the Lugano criteria is shown in ► **Fig. 2**. The measurement parameters and the calculated values are shown in ► **Table 5**. If no spleen and liver lesions are present, the anatomical measurements of the target lesions are decisive for the response classification. A decrease in the SPD of 82.3% results in a treatment response categorization as PR.



► **Fig. 2** Computed tomography of a 40-year-old man with Hodgkin lymphoma. The left side shows the pretherapeutic images (baseline), and the right side shows the images after chemotherapy (follow-up). **a, b** Mediastinal lymph node. **c, d** Right axillary lymph node. **e, f** Retroperitoneal lymph node.



► **Fig. 2** Computed tomography of a 40-year-old man with Hodgkin lymphoma. The left side shows the pretherapeutic images (baseline), and the right side shows the images after chemotherapy (follow-up). **g, h** Left supraclavicular lymph node. **i, j** Right cervical lymph node.

► **Table 5** Calculation of the anatomical measurement parameters regarding the patient case in ► **Fig. 2**.

target lesion	GTD in mm		SAD in mm		PPD in mm ²		SPD in mm ²	
	baseline	follow-up	baseline	follow-up	baseline	follow-up		
baseline/follow-up							7047	1246
lymph node, axillary right	43	18	28	10	1204	180		
lymph node, mediastinal	43	26	27	20	1161	520		
lymph node, retroperitoneal	30	13	24	8	720	104		
lymph node, supraclavicular left	57	20	34	13	1938	260		
lymph node, cervical right	46	14	44	13	2024	182		
								-82.3%

GTD: great transverse diameter, SAD: short axis diameter, SPD: sum of the product of the diameters.

LYRIC (Lymphoma Response to Immunomodulatory Therapy Criteria) modification [28]

Due to the increasing importance of biopharmaceuticals in the treatment of lymphomas, it became necessary to expand the Lugano classification in 2016 [29]. Immune-modulating biopharmaceuticals can have more diverse effects than the classic treatment consisting of chemo-(immuno-)therapy. In particular, an increase both in the size and number of lesions prior to actual

treatment response can initially occur (known as pseudoprogression). Moreover, a metabolic flare phenomenon, i. e., increased tracer uptake as part of (good) treatment response, can occur in ¹⁸F-FDG-PET. Indeterminate response (IR) is only used in the PD category. If the IR criteria are met in patients in category PD according to Lugano, treatment can initially be continued. Biopsy of the corresponding lesions can be helpful. Over the course of 12 weeks, imaging is performed again to determine whether it is a case of treatment-based transient PD or actual PD [28].

► **Table 6** shows the definition of the indeterminate response (IR) categories and ► **Table 7** shows the corresponding reevaluation of the LYRIC modification in the interval.

RECIL [30]

RECIL (Response Evaluation Criteria In Lymphoma) is based on the existing response evaluation for solid tumors RECIST 1.1 (Response Evaluation Criteria In Solid Tumors) [31]. A closely related classification of treatment response both of lymphomas and solid tumors provides the advantage of better comparability of new drugs in phase I and II studies [32]. The primary goal of the RECIL classification is to simplify the Lugano classification for lymphomas. For this purpose, a statistical comparison of the new uni-dimensional measurement and the old two-dimensional measurement according to the Lugano classification was performed. As a result, only the parameter GTD is included as an anatomical measurement parameter in the response categories according to RECIL. It was also shown that the selection of a maximum of 3 target lesions instead of the up to 6 dominant lymph nodes in the Lugano classification results in comparable assignment of patients to identical response categories. The PET/CT result was not taken into consideration in the RECIL study. Therefore, it continues to affect the classification the same as in the Lugano version [30]. In RECIL target lesions are clearly defined (GTD ≥ 1.5 cm) regardless of their nodal or extranodal position. The minor response category is new. The main differences with respect to the Lugano classification are summarized in ► **Table 8**.

The sum of the longest diameters (SLD) of the target lesions is calculated as an important anatomical measurement parameter. Lymph nodes can be selected as the target lesion when the GTD is ≥ 1.5 cm. Particularly small, elongated lymph nodes should not be selected as the target lesion. Extranodal lesions can be selected as the target lesion when they have a soft-tissue component and a GTD ≥ 1.5 cm that can be reproducibly measured.

The conditions for CR were defined more strictly in comparison to the Lugano classification. Therefore, all defined target lesions must have a post-treatment GTD < 1.0 cm, while a reduction of the GTD to ≤ 1.5 cm was previously sufficient according to Lugano.

PET/CT continues to be recommended for FDG-avid lymphomas. The evaluation is still performed according to Deauville-5-PS. Bone marrow biopsy continues to be important for the corresponding lymphomas. The response evaluation is summarized in ► **Table 9**.

While the LYRIC modification was added to the Lugano classification to address immunomodulatory therapeutic agents, the RECIL classification recommends confirming PD in two consecutive imaging examinations in the interval.

Measured lymph node conglomerates that separate over the course of treatment are treated as described in the Lugano classification. The SLD of the new individual lesions is calculated in the control examinations. The spleen size is measured in the vertical length according to the Lugano classification with the normal size remaining < 13 cm [30].

► **Table 6** LYRIC modification of the Lugano classification.

indeterminate response	criterion
IR 1	increase in the SPD of the up to 6 measurable target lesions $\geq 50\%$ in the first 12 weeks without clinical worsening
IR 2	increase in the SPD of the up to 6 measurable target lesions $< 50\%$ with a) new lesion(s) or b) increase in the PPD of one or more lesion(s) $\geq 50\%$ at any time during treatment
IR 3	increase in ^{18}F -FDG uptake of one or more lesion(s) without fulfilling the PD criteria in the form of an increase in size or new lesion(s)

IR: indeterminate response, SPD: sum of the product of the diameters, FDG: fluorodeoxyglucose.

► **Table 7** Reevaluation of the LYRIC modification in the interval.

	objective true PD in the case of fulfillment of the following criteria
according to IR 1	comparison of SPD in the case of fulfillment of IR 1 with the current SPD: increase in SPD $\geq 10\%$ and increase ≥ 0.5 cm in any axis in the case of one or more lesion(s) ≤ 2.0 cm or increase ≥ 1.0 cm in any axis in the case of one or more lesion(s) > 2.0 cm
according to IR 2	new lesion(s) (maximum of 6) are added to the SPD: increase in the new SPD $\geq 50\%$ above the SPD of these lesions at nadir
according to IR 3	only in the case of the occurrence of new lesion(s) or Increase in size to fulfillment of PD criteria

IR: indeterminate response, PD: progressive disease, SPD: sum of the product of the diameters.

As an example, the RECIL classification is also performed for the case in ► **Fig. 2**. A maximum of 3 target lesions are selected (condition ≥ 1.5 cm). The largest lymph nodes from various anatomical lymph node stations are representatively selected for this purpose: Left supraclavicular lymph node (► **Fig. 2 g, h**) with a GTD of 57 mm, right cervical lymph node (► **Fig. 2i, j**) with a GTD of 46 mm and mediastinal lymph node (► **Fig. 2c, d**) with a GTD of 43 mm. The sum of the GTDs at baseline is 146 mm (= SLD). In the follow-up, the SLD is only 60 mm. This is a decrease in the SLD of 58.9%. Since the criterion for achieving a GTD < 1.0 cm in all target lesions is not met but the SLD reduction is $\geq 30\%$, the patient is also classified as PR according to the RECIL classification.

► **Table 8** Comparison of the response classifications according to Lugano and RECIL.

	Lugano classification	RECIL classification
anatomical measurement	two-dimensional	one-dimensional
definition of target lesion	nodal: GTD > 1.5 cm extranodal: GTD > 1 cm	nodal/extranodal: GTD ≥ 1.5 cm
number of target lesions	max. 6	max. 3
number of response categories	4	5
SD range	-50 % to +50 % (calculation of area)	-10 % to +20 % (calculation of diameter)

SD: stable disease; RECIL: Response Evaluation Criteria in Lymphoma; GTD: great transverse diameter.

► **Table 9** Response classification according to RECIL.

	CR	PR	MR	SD	PD
CT	complete disappearance of all target lesions and lymph nodes with GTD < 1.0 cm	≥ 30 % reduction of SLD of the target lesions, but CR criteria not met	≥ 10 % reduction of SLD of the target lesions, but PR criteria not met	< 10 % reduction or ≤ 20 % increase in SLD of target lesions	> 20 % increase in SLD of target lesions (for small lymph nodes with baseline value < 1.5 cm, increase of at least 0.5 cm and GTD > 1.5 cm) or
PET/CT	≥ 30 % decrease in SLD of target lesions with a score of 1/2/3 according to 5-PS	score 4/5 according to 5-PS	every score	every score	every score
new lesions	none	none	none	none	none or present

RECIL: Response Evaluation Criteria in Lymphoma, CR: complete response, PR: partial response, MR: minor response, SD: stable disease, PD: progressive disease, GTD: great transverse diameter, SLD: sum of longest diameters, PET: positron emission tomography, 5-PS: Deauville 5-point scale.

Discussion

While only anatomical measurement parameters were used in the first classification according to Cheson (IWG criteria), functional-metabolic imaging in the form of PET/CT was included in the revised Cheson classification. However, the evaluation was only positive or negative. In the subsequent Lugano classification, the dichotomous PET evaluation was replaced by the Deauville 5-point scale. The addition of the LYRIC modification takes into consideration the use of new therapeutic agents and their possible transient effects, such as pseudoprogression and increased metabolic activity on PET/CT. The latest RECIL classification is based on the Lugano classification but attempts to greatly simplify its use and to adapt to the RECIST evaluation for solid tumors. Therefore, studies show that despite the use of simplified parameters, such as the one-dimensional measurement of target lesions, a comparable response classification as in the Lugano classification can be achieved [30]. The response category “minor response” was also added.

A central problem regarding the imaging evaluation of lymphomas is the handling of residual disease after treatment. Lymph nodes do not disappear completely even in CR. In addition, fibrosis, inflammation, and necrosis can result in an increase in the size of the original lymph node despite tumor-free status [34]. Moreover, it must be taken into consideration that the response of the tumor stroma is greatly delayed and is thus not yet concluded in the case of an early response evaluation, especially since desmoid segments can often persist. These changes cannot always be clearly differentiated from lymph nodes with vital tumor cells based on morphological imaging data [35]. Therefore, an increasing number of non-morphological imaging parameters are also used for prognosis evaluation. In an analysis of patients of the German CLL Study Group, the role of the MRD status was able to be examined more closely. In the case of a negative MRD status, no difference regarding progression-free survival could be shown between CR patients and PR patients with residual splenomegaly. However, MRD-nega-

tive PR patients with residual lymphadenopathy had a shorter progression-free survival [36]. It must be assumed that additional non-morphological imaging parameters will be applied to the response evaluation thereby increasing its accuracy.

Although the use of standardized criteria has many advantages, there are also a number of sources of error. Examinations being performed by different examiners at various times in the course of treatment can cause measurement errors that can result in incorrect categorization. Studies were able to show, for example for RECIST, that the interobserver variability is significantly greater than the intraobserver variability [37]. This means that multiple evaluations by one radiologist have fewer inaccuracies than evaluations performed by various radiologists. As a result of technical progress and particularly better image resolution, measurement differences have decreased [38]. Nonetheless, if possible, one examiner should be assigned to each patient and kept consistent to at least avoid interobserver variability in studies. Studies have since been able to show that semiautomatic measurement was significantly more accurate for response evaluation in lymphoma than the manual measurement method [39, 40].

Even if a return to one-dimensional measurement of lesions is recommended in the latest RECIL classification for simplification, the role of the volumetric measurement of target lesions for lymphomas should also be examined in the future. This is particularly true in light of the increasingly better functioning of automatic volumetry [39, 40]. Volumetric limits have already been calculated for RECIST 1.1. A primary advantage of volumetry is that it is not affected by asymmetry and other shape-related anomalies of target lesions [41].

The steady technical progress has an impact not only on the automated measurement of lesions but also on examination methods. In the recommendations for the Lugano classification, MRI is only mentioned with respect to CNS involvement [10], while in RECIL MRI is mentioned as an alternative examination modality to CT [30]. The importance of the hybrid method PET/MRI is the subject of current studies [42, 43]. The potential of PET/MRI to reduce ionizing radiation and provide better morphological imaging conditions for certain lymphoma entities is increasingly important.

There are also points for discussion on a statistical basis. For example, the classification of response into categories results in a loss of information since the data was originally continuous. Although categories can simplify study results and also contribute to a better representation of data, the thresholds are ultimately arbitrary. A major advantage of continuous data is the possibility to reduce the size of study populations while maintaining the same level of significance of the results [44]. The more accurate the statistical analysis of continuous data, the smaller the loss of information [45]. Although the categorization of data contributes to the simplification of study results and thus to a better representation, the power gain of the study design for examining rare results (like lymphomas) with the resulting challenge regarding recruiting sufficient patient numbers should be given priority. In future updates of the Lugano or RECIL classification, the use of adequate statistical tests with continuous data should be considered and the maintenance of categorized criteria should be examined. It would be more useful, for example, to calculate response

rates as a percentage without subsequent categorization and to examine their statistical effect on survival. This idea is admittedly not completely new. In 1981, Lavin et al. suggested using the numerical ratio of tumor measurements for further examinations rather than classification into categories [44].

Standardized criteria for evaluating treatment response in lymphomas are essential for the comparability and standardization of study data and for facilitating multicenter studies. Technical progress, e. g. in the field of hybrid imaging and semiautomatic evaluation, necessitates regular updating of criteria-based imaging. In the further development of criteria-based evaluation methods, it is necessary to constantly weigh possibilities for optimizing measurement accuracy, the complexity of the criteria, correlation to the actual tumor burden, and finally practicability.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

- [1] Stiefelhagen P. Die Therapie der Lymphome ist im Wandel. *InFo Hämatologie und Onkologie* 2017; 2017: 62–63. doi:10.1007/s15004-017-5763-z
- [2] Cheson BD, Horning SJ, Coiffier B et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. NCI Sponsored International Working Group. *J Clin Oncol* 1999; 17: 1244. doi:10.1200/JCO.1999.17.4.1244
- [3] Lister TA, Crowther D, Sutcliffe SB et al. Report of a committee convened to discuss the evaluation and staging of patients with Hodgkin's disease: Cotswolds meeting. *J Clin Oncol* 1989; 7: 1630–1636. doi:10.1200/JCO.1989.7.11.1630
- [4] Dorfman RE, Alpern MB, Gross BH et al. Upper abdominal lymph nodes: criteria for normal size determined with CT. *Radiology* 1991; 180: 319–322. doi:10.1148/radiology.180.2.2068292
- [5] Einstein DM, Singer AA, Chilcote WA et al. Abdominal lymphadenopathy: spectrum of CT findings. *Radiographics* 1991; 11: 457–472. doi:10.1148/radiographics.11.3.1852937
- [6] Glazer GM, Gross BH, Quint LE et al. Normal mediastinal lymph nodes: number and size according to American Thoracic Society mapping. *Am J Roentgenol* 1985; 144: 261–265. doi:10.2214/ajr.144.2.261
- [7] Kiyono K, Sone S, Sakai F et al. The number and size of normal mediastinal lymph nodes: a postmortem study. *Am J Roentgenol* 1988; 150: 771–776. doi:10.2214/ajr.150.4.771
- [8] Steinkamp HJ, Hosten N, Richter C et al. Enlarged cervical lymph nodes at helical CT. *Radiology* 1994; 191: 795–798. doi:10.1148/radiology.191.3.8184067
- [9] Cheson BD, Pfistner B, Juweid ME et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007; 25: 579–586. doi:10.1200/JCO.2006.09.2403
- [10] Barrington SF, Mikhael 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. doi:10.1200/JCO.2013.53.5229
- [11] 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: 3059–3068. doi:10.1200/JCO.2013.54.8800
- [12] Meignan M, Gallamini A, Meignan M et al. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma* 2009; 50: 1257–1260. doi:10.1080/10428190903040048

- [13] Barrington SF, Qian W, Somer EJ et al. Concordance between four European centres of PET reporting criteria designed for use in multicentre trials in Hodgkin lymphoma. *Eur J Nucl Med Mol Imaging* 2010; 37: 1824–1833. doi:10.1007/s00259-010-1490-5
- [14] 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. doi:10.1200/JCO.2017.75.2543
- [15] Spaepen K, Stroobants S, Dupont P et al. Early restaging positron emission tomography with (18)F-fluorodeoxyglucose predicts outcome in patients with aggressive non-Hodgkin's lymphoma. *Ann Oncol* 2002; 13: 1356–1363. doi:10.1093/annonc/mdf256
- [16] Hutchings M, Mikhaeel NG, Fields PA et al. Prognostic value of interim FDG-PET after two or three cycles of chemotherapy in Hodgkin lymphoma. *Ann Oncol* 2005; 16: 1160–1168. doi:10.1093/annonc/mdi200
- [17] Hutchings M, Loft A, Hansen M et al. FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. *Blood* 2006; 107: 52–59. doi:10.1182/blood-2005-06-2252
- [18] Cerci JJ, Pracchia LF, Linardi CC et al. 18F-FDG PET after 2 cycles of ABVD predicts event-free survival in early and advanced Hodgkin lymphoma. *J Nucl Med* 2010; 51: 1337–1343. doi:10.2967/jnumed.109.073197
- [19] Borchmann P, Goergen H, Kobe C et al. PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. *Lancet* 2018; 390: 2790–2802. doi:10.1016/S0140-6736(17)32134-7
- [20] Duhren U, Muller S, Hertenstein B et al. Positron Emission Tomography-Guided Therapy of Aggressive Non-Hodgkin Lymphomas (PETAL): A Multicenter, Randomized Phase III Trial. *J Clin Oncol* 2018; 36: 2024–2034. doi:10.1200/JCO.2017.76.8093
- [21] Le Roux PY, Gastinne T, Le Gouill S et al. Prognostic value of interim FDG PET/CT in Hodgkin's lymphoma patients treated with interim response-adapted strategy: comparison of International Harmonization Project (IHP), Gallamini and London criteria. *Eur J Nucl Med Mol Imaging* 2011; 38: 1064–1071. doi:10.1007/s00259-011-1741-0
- [22] Kajary K, Molnar Z, Gyorke T et al. Comparison of the International Harmonization Project, London and Gallamini criteria in the interpretation of 18F-FDG PET/CT examinations after first-line treatment in Hodgkin's lymphoma. *Nucl Med Commun* 2014; 35: 169–175. doi:10.1097/MNM.0000000000000024
- [23] Fallanca F, Alongi P, Incerti E et al. Diagnostic accuracy of FDG PET/CT for clinical evaluation at the end of treatment of HL and NHL: a comparison of the Deauville Criteria (DC) and the International Harmonization Project Criteria (IHPC). *Eur J Nucl Med Mol Imaging* 2016; 43: 1837–1848. doi:10.1007/s00259-016-3390-9
- [24] Abrey LE, Batchelor TT, Ferreri AJ et al. Report of an international workshop to standardize baseline evaluation and response criteria for primary CNS lymphoma. *J Clin Oncol* 2005; 23: 5034–5043. doi:10.1200/JCO.2005.13.524
- [25] Zucca E, Copie-Bergman C, Ricardi U et al. Gastric marginal zone lymphoma of MALT type: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013; 24 (Suppl. 6): vi144–vi148. doi:10.1093/annonc/mdt343
- [26] Olsen EA, Whittaker S, Kim YH et al. Clinical end points and response criteria in mycosis fungoides and Sezary syndrome: a consensus statement of the International Society for Cutaneous Lymphomas, the United States Cutaneous Lymphoma Consortium, and the Cutaneous Lymphoma Task Force of the European Organisation for Research and Treatment of Cancer. *J Clin Oncol* 2011; 29: 2598–2607. doi:10.1200/JCO.2010.32.0630
- [27] Hallek M, Cheson BD, Catovsky D et al. iwCLL guidelines for diagnosis, indications for treatment, response assessment, and supportive management of CLL. *Blood* 2018; 131: 2745–2760. doi:10.1182/blood-2017-09-806398
- [28] 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. doi:10.1182/blood-2016-05-718528
- [29] Cheson BD. Rethinking clinical response and outcome assessment in a biologic age. *Curr Oncol Rep* 2015; 17: 27. doi:10.1007/s11912-015-0452-2
- [30] 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. doi:10.1093/annonc/mdx097
- [31] Eisenhauer EA, Therasse P, Bogaerts J et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009; 45: 228–247. doi:10.1016/j.ejca.2008.10.026
- [32] Fournier L, Ammari S, Thiam R et al. Imaging criteria for assessing tumour response: RECIST, mRECIST, Cheson. *Diagn Interv Imaging* 2014; 95: 689–703. doi:10.1016/j.diii.2014.05.002
- [33] Assouline S, Meyer RM, Infante-Rivard C et al. Development of adapted RECIST criteria to assess response in lymphoma and their comparison to the International Workshop Criteria. *Leuk Lymphoma* 2007; 48: 513–520. doi:10.1080/10428190601078126
- [34] Lewis E, Bernardino ME, Salvador PG et al. Post-therapy CT-detected mass in lymphoma patients: is it viable tissue? *J Comput Assist Tomogr* 1982; 6: 792–795
- [35] Ganten MK, Weber MA, Ganten TM. Cellular mechanisms of tumor response: clinical demands. *Radiologe* 2008; 48: 820–831. doi:10.1007/s00117-008-1739-1
- [36] Kovacs G, Robrecht S, Fink AM et al. Minimal Residual Disease Assessment Improves Prediction of Outcome in Patients With Chronic Lymphocytic Leukemia (CLL) Who Achieve Partial Response: Comprehensive Analysis of Two Phase III Studies of the German CLL Study Group. *J Clin Oncol* 2016; 34: 3758–3765. doi:10.1200/JCO.2016.67.1305
- [37] Erasmus JJ, Gladish GW, Broemeling L et al. Interobserver and intraobserver variability in measurement of non-small-cell carcinoma lung lesions: implications for assessment of tumor response. *J Clin Oncol* 2003; 21: 2574–2582. doi:10.1200/JCO.2003.01.144
- [38] Hopper KD, Kasales CJ, Van Slyke MA et al. Analysis of interobserver and intraobserver variability in CT tumor measurements. *Am J Roentgenol* 1996; 167: 851–854. doi:10.2214/ajr.167.4.8819370
- [39] Wessling J, Puesken M, Koch R et al. MSCT follow-up in malignant lymphoma: comparison of manual linear measurements with semi-automated lymph node analysis for therapy response classification. *Röfo* 2012; 184: 795–804. doi:10.1055/s-0032-1312751
- [40] Wessling J, Schulke C, Koch R et al. Therapy response evaluation of malignant lymphoma in a multicenter study: comparison of manual and semiautomatic measurements in CT. *Röfo* 2014; 186: 768–779. doi:10.1055/s-0033-1356424
- [41] Winter KS, Hofmann FO, Thierfelder KM et al. Towards volumetric thresholds in RECIST 1.1: Therapeutic response assessment in hepatic metastases. *Eur Radiol* 2018; 28: 4839–4848. doi:10.1007/s00330-018-5424-0
- [42] Afaq A, Fraioli F, Sidhu H et al. Comparison of PET/MRI With PET/CT in the Evaluation of Disease Status in Lymphoma. *Clin Nucl Med* 2017; 42: e1–e7. doi:10.1097/RLU.0000000000001344
- [43] Atkinson W, Catana C, Abramson JS et al. Hybrid FDG-PET/MR compared to FDG-PET/CT in adult lymphoma patients. *Abdom Radiol (NY)* 2016; 41: 1338–1348. doi:10.1007/s00261-016-0638-6
- [44] Lavin PT. An alternative model for the evaluation of antitumor activity. *Cancer Clin Trials* 1981; 4: 451–457
- [45] Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006; 25: 127–141. doi:10.1002/sim.2331