HemaSphere

EHA EUROPEAN HEMATOLOGY ASSOCIATION

Letter
Open Access

Chimeric Antigen Receptor T-cell Therapy: Imaging Response Criteria and Relation to Progression-free and Overall Survival

Michael Winkelmann[1], Kai Rejeski[2], Viktoria Blumenberg[2], Veit L. Bücklein[2], Michael Ruzicka[2], Marcus Unterrainer[1], Christian Schmidt[2], Franziska J. Dekorsy[3], Peter Bartenstein[3,4], Jens Ricke[1,4], Michael von Bergwelt-Baildon[2,4], Marion Subklewe[2,4], Wolfgang G. Kunz[1,4]

Correspondence: Wolfgang G. Kunz (wolfgang.kunz@med.lmu.de).

Chimeric antigen receptor T-cell therapy (CART) uses patient-derived tumor antigen-directed T cells for targeted elimination of cancer cells.[1] The most common form applies modified T cells expressing a chimeric antigen receptor specific for the CD19 antigen to treat relapsed or refractory (rr) lymphoma[2] and leukemia, leading to high rates of durable responses. CART has significantly improved progression-free survival (PFS) and overall survival (OS). Imaging-based response assessment for determination of PFS has most frequently relied on positron emission tomography-computed tomography (PET/CT). The current and ongoing phase III trials are mostly based on the Lugano criteria from 2014.[3,4] Earlier trials have relied on Cheson criteria as published in 2007.[5] In recent years, novel lymphoma imaging response criteria have been proposed, among them the response evaluation criteria in lymphoma (RECIL)[6] and lymphoma response to immunomodulatory therapy criteria (LYRIC).[7] The scientific literature on structured comparisons of these imaging response criteria is scarce for conventional lymphoma treatments and only 2 studies indicate concordance of RECIL and Lugano criteria in previously untreated lymphoma.[8,9] As there are no reports on the prognostic value for lymphoma patients treated with CART, we aimed to assess the different imaging response criteria, their impact on PFS, and their relation to OS.

The study population was based on a prospective registry of all consecutive patients who were treated at the Comprehensive Cancer Center Munich-Ludwig-Maximilian University Munich ($CCCM^{LMU}$) with commercialized CD19-specific CART products. We included patients with refractory or relapsed lymphoma (DLBCL, FL, and MCL), any measurable disease on imaging according to Lugano criteria,[3] and available (PET/CT) imaging studies at baseline and at least 2 follow-up timepoints (FU1 around 30 days and FU2 around 90 days). All medical records and imaging studies were reviewed with the approval of the LMU Munich Institutional Review Board (Ethikkommission der Medizinischen Fakultät der Ludwig-Maximilians-Universität München, Project Number 19-817) and informed patient consent. Patients received lymphodepletion with fludarabine and cyclophosphamide according to the manufacturers' instructions. Overall response was determined based on Lugano criteria with up to 6 target lesions (TLs) that were segmented. The sum of the product of diameters (SPD) was measured to determine tumor burden (TB) for Lugano criteria, Cheson criteria, and LYRIC. Moreover, spleen size was measured with splenomegaly being defined by a vertical length >13.0 cm according to Lugano criteria. To assess response according to RECIL, the sum of longest diameters (SLD) of ≤3 TLs was measured to define TB. All imaging analyses were performed with the dedicated trial reporting software mintLesion 3.8 (mint Medical GmbH; Heidelberg, Germany). For survival analysis, OS was visualized using Kaplan–Meier survival curves with categorization for the patients to the response categories complete response (CR), partial response (PR), stable disease (SD), and progressive disease (PD) for all response criteria. The additional category of minor response (MR) was added for RECIL and indeterminate response (IR) for LYRIC. The overall response rate (ORR) was calculated as the rate of patients with CR and PR. Log-rank (Mantel-Cox) test was performed to examine the significance of the results. *P* values below 0.05 were considered to indicate statistical significance.

Forty-one of 74 patients met the inclusion criteria (median age: 64 years, 41% female). According to Lugano criteria, 23 patients (56.1%) showed a CR, and 5 patients (12.2%) a PR at FU2. Thirteen patients (31.7%) had an overall PD at FU2 by Lugano criteria. Discordance in the classification of overall response was observed when applying other response criteria (Table 1). Notably, Cheson criteria and RECIL classified 4 patients as an SD, whereas there were none by Lugano criteria. In addition, 2 patients had a MR by RECIL, and 6 patients were assigned to the IR category according to LYRIC. ORR was 68% by Lugano criteria, 63% by Cheson criteria, 68% by RECIL, and 68% by LYRIC. There was a significant difference in survival between patients with CR, PR, and PD using Lugano criteria

[1]Department of Radiology, University Hospital, LMU Munich, Germany
[2]Department of Medicine III, University Hospital, LMU Munich, Germany
[3]Department of Nuclear Medicine, University Hospital, LMU Munich, Germany
[4]Comprehensive Cancer Center München-LMU ($CCCM^{LMU}$), LMU Munich, Germany

HemaSphere (2022) 6:10(e781).
http://dx.doi.org/10.1097/HS9.0000000000000781.
Received: July 29, 2022 / Accepted: August 28, 2022

Table 1

Overview of Different Imaging Response Criteria at 90 Days

	ORR	Median PFS, d	CR	PR	MR	SD	PD	IR
Lugano	28 (68%)	153	23 (56%)	5 (12%)	—	0	13 (32%)	—
Cheson	26 (63%)	169	19 (46%)	7 (17%)	—	4 (10%)	11 (27%)	—
RECIL	28 (68%)	198	21 (51%)	7 (17%)	2 (5%)	4 (10%)	7 (17%)	—
LYRIC	28 (68%)	200	23 (56%)	5 (12%)	—	0	7 (17%)	6 (15%)

Comparison of different classifications for overall response in lymphoma according to Lugano, Cheson, RECIL, and LYRIC at 90 days after CART. Shown are the color-coded response categories: CR (green), PR (yellow), MR (orange), SD (gray), PD (red), and IR (blue). ORR was calculated as rate of patients with CR and PR. Median PFS is shown for every classification.

CART = chimeric antigen receptor T-cell therapy; CR = complete response; IR = indeterminate response; LYRIC = lymphoma response to immunomodulatory therapy criteria; MR = minor response; ORR = overall response rate; PD = progressive disease; PFS = progression-free survival; PR = partial response; RECIL = response evaluation criteria in lymphoma; SD = stable disease.

(Fig. 1A) at FU2 ($P < 0.001$). Patients with CR had longer OS than patients with PD. Patients with PR were between CR and PD ($P = 0.046$). Dichotomization into responding versus nonresponding patients allowed stratification of OS (HR = 18.2, $P < 0.001$). Using Cheson criteria (Fig. 1B), similar results were observed ($P < 0.001$), with the addition of Kaplan–Meier curve of SD patients falling between PR and PD. Survival curves by LYRIC (Fig. 1C) showed essentially the same results as by Lugano criteria with a significant difference between the categories CR, PR, and PD ($P < 0.001$). There was a small yet statistically nonsignificant difference between PD and IR patients. When categorized by RECIL (Fig. 1D), a difference between response categories was also demonstrated ($P < 0.001$); notably, the 2 patients with MR showed longer OS.

The most widely established lymphoma response criteria showed considerable differences in imaging endpoints in our study on patients that received CART. While the ORR was least affected, rates of CR and PR differed slightly. Notably, the response categories SD and PD were more strongly affected and thus the endpoint PFS varied (PFS_{Lugano} 153 d versus PFS_{Cheson} 169 d versus PFS_{RECIL} 198 d versus PFS_{LYRIC} 200 d). Interestingly, the immune-adapted LYRIC criteria yielded the strongest association of PFS_{LYRIC} to OS in this patient population.

The Cheson and Lugano criteria have evolved from the unidimensional Response Evaluation Criteria in Solid Tumors (RECIST1.1) criteria[10] and capture bidimensional extension for this typically nodal-dominant tumor phenotype.[3,5] As treated lymphomas are often characterized by morphologic residues, lymphoma criteria incorporate the metabolic status as visualized by ^{18}F-FDG PET/CT to identify CR of the initially vital tumor tissue. RECIL was developed to facilitate the study assessment by reducing the number of TL to capture the same response.[6] A reduction to 3 representative lesions yielded robust response classification as in the other criteria relying on 6 manifestations. Furthermore, RECIL incorporates the depth or response of the morphologic extension and is thereby not only based on metabolic changes.

Yet, in the course of the disease with relapsed of refractory lymphomas, phenotypic and metabolic changes of the lymphoma manifestations occur. Typically, more widespread nodal locations are involved and extranodal lesions are more frequently encountered. Moreover, preexisting morphologic residues can be mistaken for active lymphoma if prior imaging exams are not reviewed. Notably, response criteria have not been adapted to such changes in the disease course and data on association of PFS with OS in lymphoma is scarce.[3,10]

Data characterizing response to CART is also limited and the optimally discriminating method remains ill-defined. In a single-arm, prospective study of 7 patients, early PET/CT data from lymphoma patients treated with CD19 CART were evaluated according to Lugano criteria. In this study, all patients with less than CR at 1-month subsequently relapsed.[11] A recent multicenter study of 171 patients showed similar results. Patients with Deauville Score 1+2 at 1-month FU had an excellent long-term outcome, whereas 31% of patients with Deauville Score 3+4 were at risk for early relapse, and all patients with Deauville Score 5 relapsed by month 3.[12]

The evaluation of overall response and response patterns, including the impact of pseudoprogression associated with CAR T-cell therapy, has not yet been studied in detail.[13] Some studies described pseudoprogression after CART analogous to solid tumors under immunotherapy.[14] To face the challenge of pseudoprogression, LYRIC introduced the category of IR, with 3 subcategories: IR1, increase in overall tumor burden within the first 12 weeks of therapy, without clinical deterioration; IR2, appearance of new lesions, or growth of one or more existing lesions ≥50% at any time during treatment in the absence of overall progression; IR3, increase in FDG uptake of one or more lesions without a concomitant increase in lesion size or number.[7] LYRIC encouraged biopsy for IR1 and IR2 and advised to evaluate these intermediate features by follow-up in all cases after 12 weeks. In contrast to LYRIC, Lugano or RECIL do not provide recommendations for lesion follow-up.[6] Therefore, patients with assigned PD solely based on newly appearing lesions should be further investigated with regard to clinical benefit and may represent a new response category. Novel imaging endpoints and response criteria in lymphoma will likely evolve from selected lesion-based assessments to whole tumor burden quantification. In the first-line setting, the recently published International Metabolic Prognostic Index (IMPI) additionally integrates metabolic tumor volume and has outperformed the conventional IPI in estimating outcome of DLBCL patients.[15]

We investigated overall response by Lugano criteria, Cheson criteria, RECIL, and LYRIC. While the ORR was comparable between the different criteria, we found striking differences between the SD and PD response categories and thus discrepancies in the surrogate endpoint PFS. Response assessment by LYRIC exhibited superior association between PFS and OS. The response assessment method must therefore be considered when interpreting the impact of imaging endpoints on outcomes in clinical trials. Our study has limitations which need to be considered when interpreting the results. First, this is a single-center study with a limited number of subjects. Second, there were a few patients that were missed to follow up or had no measurable disease. Considering the heterogeneity, our results argue for standardization and harmonization across centers.

AUTHOR CONTRIBUTIONS

MW and WGK conceived and design the study; VLB, VB, KR, MR, MU, and CS collected the data; MW, VLB, VB, KR, and WGK analyzed and interpreted the data; and MW and WGK drafted the manuscript; and KR, FJD, PB, JR, MvB-B, and MS revised the manuscript.

DISCLOSURES

VB has received industry research support from BMS/Celgene, Kite/Gilead, Janssen, Novartis, Roche, and Takeda. KR declares having received research funding and travel support from Kite/Gilead and honoraria from Novartis. CS received travel support from Kite/Gilead. MvB-B received

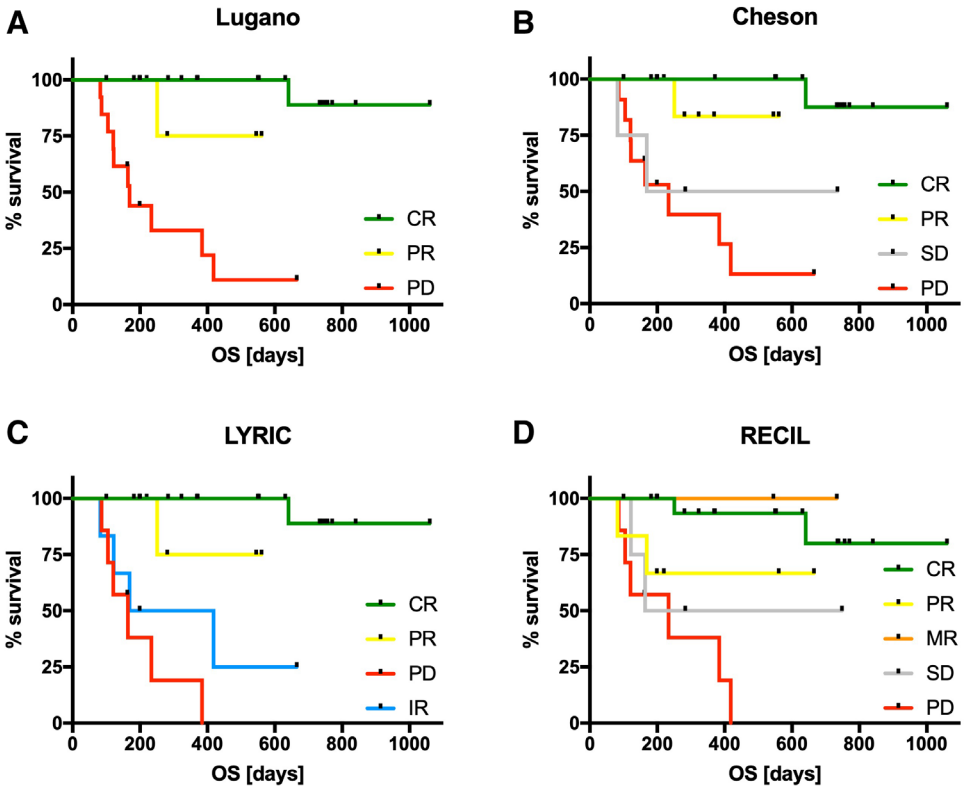

Figure 1. Overall survival stratification. Analysis of OS according to the response criteria Lugano (A), Cheson (B), LYRIC (C), and RECIL (D). The same color coding as in Table 3 was used to label the different response categories: CR (green), PR (yellow), MR (orange), SD (gray), PD (red), and IR (blue). CR = complete response; IR = indeterminate response; LYRIC = lymphoma response to immunomodulatory therapy criteria; MR = minor response; OS = overall survival; PD = progressive disease; PFS = progression-free survival; PR = partial response; RECIL = response evaluation criteria in lymphoma; SD = stable disease.

research funding and honoraria from Novartis, Kite Pharma, Miltenyi Biotec, Mologen, MSD, Astellas, and Roche. MS received industry research support from Amgen, Gilead, Miltenyi Biotec, MorphoSys, Roche, and Seattle Genetics; served as a consultant or advisor to Amgen, Bristol Myers Squibb, Celgene, Gilead, Pfizer, Novartis, and Roche; is on the advisory boards of Amgen, Celgene, Gilead, Janssen, Novartis, Pfizer, and Seattle Genetics; and serves on the speaker's bureau at Amgen, Celgene, Gilead, Janssen, and Pfizer. All the other authors have no conflicts of interest to disclose.

SOURCES OF FUNDING

The work was supported by funding from the research program "Förderung für Forschung und Lehre (FöFoLe) project number 1147" of the Medical Faculty of Ludwig Maximilian University (LMU) Munich and the Bavarian Cancer Research Center (BZKF) to MW. The work was further supported by the Else-Kröner-Fresenius Stiftung (to VB) and the German Cancer Consortium DKTK (to VB).

REFERENCES

1. June CH, Sadelain M. Chimeric antigen receptor therapy. *N Engl J Med*. 2018;379:64–73.
2. Neelapu SS, Locke FL, Bartlett NL, et al. Axicabtagene ciloleucel CAR T-Cell therapy in refractory large B-cell lymphoma. *N Engl J Med*. 2017;377:2531–2544.
3. 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.
4. 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.
5. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007;25:579–586.
6. 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.
7. 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.
8. Berzaczy D, Haug A, Staber PB, et al. RECIL versus Lugano for treatment response assessment in FDG-avid non-Hodgkin lymphomas: a head-to-head comparison in 54 patients. *Cancers (Basel)*. 2019;12:E9.
9. Kostakoglu L, Martelli M, Sehn LH, et al. Complete response status according to RECIL 2017 criteria shows high concordance with Lugano 2014 criteria and is highly prognostic for outcome in previously untreated patients with CD20-positive diffuse large B-cell lymphoma (DLBCL). *Blood*. 2019;134:489–489
10. 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.
11. Shah NN, Nagle SJ, Torigian DA, et al. Early positron emission tomography/computed tomography as a predictor of response after CTL019 chimeric antigen receptor -T-cell therapy in B-cell non-Hodgkin lymphomas. *Cytotherapy*. 2018;20:1415–1418.
12. Kuhnl A, Roddie C, Kirkwood AA, et al. Early FDG-PET response predicts CAR-T failure in large B-cell lymphoma. *Blood Adv*. 2022;6:321–326.
13. Vercellino L, de Jong D, di Blasi R, et al. Current and future role of medical imaging in guiding the management of patients with relapsed and refractory Non-Hodgkin lymphoma treated with CAR T-cell therapy. *Front Oncol*. 2021;11:664688.
14. Wang J, Hu Y, Yang S, et al. Role of fluorodeoxyglucose positron emission tomography/computed tomography in predicting the adverse effects of chimeric antigen receptor T cell therapy in patients with non-Hodgkin lymphoma. *Biol Blood Marrow Transplant*. 2019;25:1092–1098.
15. Mikhaeel NG, Heymans MW, Eertink JJ, et al. Proposed new dynamic prognostic index for diffuse large B-cell lymphoma: international metabolic prognostic index. *J Clin Oncol*. 2022;40:2352–2360.