Published in final edited form as:

Clin Trials. 2016 October ; 13(5): 545–554. doi:10.1177/1740774516645338.

Response assessment in lymphoma: Concordance between independent central review and local evaluation in a clinical trial setting

Ajay K Gopal¹, Barbara Pro², Joseph M Connors³, Anas Younes⁴, Andreas Engert⁵, Andrei R Shustov⁶, Xuedong Chi⁷, Emily K Larsen⁸, Dana A Kennedy⁸, and Eric L Sievers⁸

¹The University of Washington/Fred Hutchison Cancer Research Center, Seattle Cancer Care Alliance, Seattle, WA, USA

²Thomas Jefferson University, Philadelphia, PA, USA

³The University of British Columbia, Vancouver, BC, Canada

⁴Memorial Sloan Kettering Cancer Center, New York, NY, USA

⁵University Hospital of Cologne, Cologne, Germany

⁶University of Washington Medical Center, Seattle, WA, USA

⁷Millenium Pharmaceuticals, Inc., Cambridge, MA, a wholly owned subsidiary of Takeda Pharmaceuticals Limited, USA

⁸Seattle Genetics, Inc., Bothell, WA, USA

Abstract

Background—Independent central review of clinical imaging remains the standard for oncology clinical trials with registration potential. A limited independent central review strategy has been proposed for solid tumor trials based on concordance between central and local evaluation of response. Concordance between independent central review and local evaluation of response in hematological malignancies is not known.

Methods—We retrospectively evaluated concordance between prospectively performed central and local assessments of response using the Revised Response Criteria for Malignant Lymphoma across two international, open-label, single-arm, registration studies of brentuximab vedotin in patients with relapsed or refractory Hodgkin lymphoma (N = 102) or systemic anaplastic large-cell lymphoma (N = 58).

Reprints and permissions: sagepub.co.uk/journalsPermissions.nav

Corresponding author: Ajay K Gopal, The University of Washington/Fred Hutchison Cancer Research Center, Seattle Cancer Care Alliance, 825 Eastlake Avenue E, G3-200, Seattle, WA 98109, USA. agopal@u.washington.edu.

ClinicalTrials.gov identifiers: NCT00848926 and NCT00866047

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Seattle Genetics, Inc. provided research funding/grants to the institutions of A.K.G., B.P., J.M.C., A.Y., A.E., and A.R.S.

Results—Overall objective response rates were similar between assessors for both the trial in Hodgkin lymphoma (75% independent central review, 72% local evaluation) and the trial in anaplastic large-cell lymphoma (86% independent central review, 83% local evaluation). Patient-specific objective response concordance was also substantial (Hodgkin lymphoma: kappa = 0.68; anaplastic large-cell lymphoma: kappa = 0.74). Median progression-free survival was similar between assessors for patients with anaplastic large-cell lymphoma (14.3 months by independent central review (95% confidence interval: 6.9, –); 14.5 months by local evaluation (95% confidence interval: 9.4, –)), but longer by local evaluation in patients with Hodgkin lymphoma (5.8 months by independent central review (95% confidence interval: 5.0, 9.0); 9.0 months by local evaluation (95% confidence interval: 7.1, 12.0)). Median duration of response was longer by local evaluation in both malignancies, which was primarily attributable to earlier computed tomography and positron emission tomography–based scoring of progression by independent central review.

Conclusion—A limited independent review audit strategy for clinical trials of some lymphomas appears feasible and practical based on substantial concordance in assessments of overall objective response by central and local evaluation in two international, prospective, registration trials in lymphoma. Some variability between assessors in the time-to-event endpoints was observed, which appeared attributable to earlier assignments of progression by independent central review compared with local evaluation.

Keywords

Response criteria; lymphoma; computed tomography/positron emission tomography; independent central review; local evaluation; concordance; Cheson

Introduction

Objective response rate, considered a direct measure of antitumor activity, is one of the endpoints recommended by the Food and Drug Administration (FDA) for approval of anticancer trials based on single-arm trials, and response duration and progression-free survival are considered supportive endpoints.¹ Response and progression are generally assessed according to standardized criteria: the Response Evaluation Criteria in Solid Tumors (RECIST), version 1.1 for solid tumors; and the International Working Group Revised Response Criteria for Malignant Lymphoma for lymphoma.^{2,3} RECIST employs mainly computed tomography (CT) scans, whereas the revised International Working Group criteria typically employ positron emission tomography (PET) scans in addition to CT scans and non-radiographic evaluations, such as of the bone marrow.

To ensure uniform application of response criteria in oncology, the use of fully independent central review of tumor measurements has been recommended by regulatory agencies.¹ A limited auditing strategy has started to gain consideration among regulators based on results from meta-analyses showing reasonable concordance between independent central review assessments and local evaluation of response in patients with solid tumors.^{4–10}

The overall concordance between lymphoma response assessments by local evaluation and independent central review has not been previously evaluated. Systematic investigator bias, heterogeneity of lymphoma disease burden among patients, and multiple variables that

comprise the response evaluation each might be anticipated to lead to significant discordance between assessors. To investigate the degree of local and central concordance in assessment of response in hematological malignancies, we retrospectively analyzed the concordance of local and central evaluation of response across two phase-2, open-label, single-arm, registration studies of brentuximab vedotin in patients with relapsed or refractory Hodgkin lymphoma or systemic anaplastic large-cell lymphoma.

Patients and methods

Study design

The two registration studies that provided data for this analysis were conducted in patients with relapsed or refractory Hodgkin lymphoma (N = 102) and patients with relapsed or refractory systemic anaplastic large-cell lymphoma (N = 58).^{11,12} Both studies were single-arm, multinational, open-label, phase-2 studies conducted in the United States, Canada, and Europe. The primary efficacy endpoint was objective response rate per independent central review, and progression-free survival and duration of response were secondary endpoints. Local evaluation of efficacy endpoints was used for clinical management of study patients and was collected as a protocol-defined exploratory analysis. The studies were approved by each investigational site's institutional review board or ethics committee.

Patients and treatment

The Hodgkin lymphoma study enrolled patients with relapsed or refractory Hodgkin lymphoma after autologous stem-cell transplantation from February to August 2009, and the anaplastic large-cell lymphoma study enrolled patients with relapsed or refractory systemic anaplastic large-cell lymphoma from June 2009 until May 2010. Patients provided informed consent in accordance with the Declaration of Helsinki. Both studies required documented CD30-positive disease by central pathology review. Brentuximab vedotin 1.8 mg/kg was administered by intravenous infusion every 3 weeks for up to 16 cycles.

Study assessments

Disease assessments included baseline documentation of disease-related signs and symptoms, a physical examination, bone marrow biopsy, and radiographic studies, including CT of the neck, chest, abdomen, and pelvis and fluorodeoxyglucose-PET. Subsequent clinical and radiological (CT) assessments were performed at Cycles 2, 4, 7, 10, 13, and 16; PET assessments were performed at Cycles 4 and 7. Patients with stable disease or better at the end of treatment continued to be followed by CT every 12 weeks until progressive disease was assessed by the investigator. Response was to be assessed according to the revised International Working Group criteria.² Up to six of the largest dominant nodes or nodal masses were to be followed for response. The lesions were to be selected at baseline and were required to be PET fluorodeoxyglucose-avid at baseline.

The methodology for the independent central review facility (BioClinica, formerly known as CoreLab Partners and as RadPharm; Princeton, NJ) was detailed under a charter. Briefly, each patient was assessed by radiological review of CT/PET images performed independently by two radiologists; in the case of discordant results, a third radiological

reviewer adjudicated. Radiological review was followed by a clinical data review performed by an oncologist. Final determination of response was made by an additional oncologist using the radiological and clinical data reviews. All reviews were conducted in a blinded fashion such that readers did not have access to sites' assessments of response and were restricted from communicating with the sites. Review was performed on a rolling basis throughout the course of the study according to a sequential locked read paradigm.

Local evaluation of clinical response followed guidelines in the study protocol. All response assessments were reviewed per the revised International Working Group criteria and source-verified by the sponsor.

Statistical analysis

This retrospective analysis is based primarily on the data cuts used to support the Hodgkin lymphoma and anaplastic large-cell lymphoma studies' final primary analyses performed after the last patients' end-of-treatment visit, which occurred in August 2010 and June 2011, respectively. A subsequent data cut (June 2013) was used to examine long-term data in a subset of patients.

Objective response rate was defined as the proportion of patients with complete remission or partial remission. Kappa coefficients were calculated to characterize the agreement in response assessments between local evaluation and independent central review (kappa = 1 implies perfect agreement).¹¹

Progression-free survival was defined as the time from start of study treatment to first documentation of objective tumor progression or death due to any cause, whichever came first. Progression-free survival was censored at the last radiological assessment documenting absence of progressive disease for patients who were removed from the study or who started another treatment other than stem-cell transplant. Duration of response was defined as the time from start of the first documentation of objective tumor response (complete or partial remission) and was censored in the same manner as progression-free survival.

Results

Patient characteristics, treatment exposure, and the primary safety and efficacy findings of these studies have been previously reported.¹²⁻¹⁴

Comparison of assessments of objective response by independent central review and local evaluation

The objective response rate by independent central review and local evaluation was nearly identical for both Hodgkin lymphoma (75% by independent central review and 72% by local evaluation) and anaplastic large-cell lymphoma (86% by independent central review and 83% by local evaluation) (Table 1). Concordance between independent central review and local evaluation for objective response at the individual patient level was substantial (kappa coefficients = 0.68, 95% confidence interval (CI): 0.51, 0.84 (Hodgkin lymphoma) and 0.74, 95% CI: 0.49, 0.99 (anaplastic large-cell lymphoma)).

Most discordance at the level of best response was between the complete and partial remission categories (Table 2). Although smaller than those for objective response, kappa coefficients for best response were still considered moderate (Hodgkin lymphoma) to substantial (anaplastic large-cell lymphoma). Different PET interpretations contributed to the majority of discordance between complete remission and partial remission (9 of the 15 discordant complete/partial remission assessments in Hodgkin lymphoma and 5 of the 6 discordant complete/partial remission assessments in anaplastic large-cell lymphoma).

In total, 17 patients had discordant objective response assessments. In 12 patients, different CT interpretations led to discordant assessments of objective response, primarily at the level of partial remission and stable disease (10 patients, Figure 1). Achievement of partial remission requires a 50% decrease in the sum of the products of diameters of up to 6 of the largest dominant masses.² The different CT assessments were primarily due to variability in reduction of the sum of the products of diameters, which was attributable to selection of different index lesions and variations in their measurements.

In two patients, the different CT interpretations led to discordance at the level of complete remission and stable disease (Figure 1). For one patient (“Other”), the site did not record lesion measurements from the patient’s end-of-treatment CT/PET scan, and the patient was assessed by local evaluation as having stable disease despite all lesions being PET-negative. The CT/PET scan was used by independent central review to calculate the sum of the products of diameters, and a complete remission was declared based upon resolution of lesions and a negative PET scan. In the other patient, progressive disease was declared by independent central review based on the identification of a new lesion. Although the patient was subsequently assessed by independent central review as having partial remission, per the charter, a patient’s best overall response could not be classified as complete or partial remission if a determination of progressive disease preceded the assessment.

Four patients had discordant objective response assessments due to variability in reduction of the sum of the products of diameters, as well as residual masses with qualitatively different PET interpretations (Figure 1). Post-treatment residual masses can meet criteria for complete remission if they are PET-negative.² Three patients were discordant at the level of complete remission and stable disease. In one patient (“Other”), the discordance was at the level of partial remission and stable disease. In this patient, although 100% reduction in the sum of the products of diameters and PET-negativity status would confer a status of complete remission per the revised International Working Group criteria, a non-index lesion was assessed as stable disease, and per the charter, non-index lesions assessed as stable disease prevented declaration of complete remission.

One patient (not shown) was assessed as not evaluable by local evaluation because neck and chest CT scans were not obtained by the investigator at Cycle 2. Independent central review was able to make an assessment in the absence of disease involvement in the neck and chest at baseline; the independent central review assessment was partial remission.

Comparison of assessments of time to progression by independent central review and local evaluation

The median progression-free survival was longer by local evaluation by approximately 3 months in the Hodgkin lymphoma study (Figure 2(a)). Particularly noteworthy was a relatively frequent declaration of progression by independent central review at the Cycle 4 PET evaluation (approximately 3 months) that was subsequently echoed by local evaluation at the Cycle 7 PET evaluation (approximately 5 months) where the two progression-free survival curves come together. A relatively higher rate of declared progression by independent central review compared with local evaluation is again observed subsequent to this timepoint. Although the progression-free survival curve corresponding to local evaluation is also above the progression-free survival curve for independent central review in the anaplastic large-cell lymphoma study, medians were similar (Figure 2(b)).

The median duration of response was longer by local evaluation by approximately 4 to 5 months across studies (Figure 2(c) and (d)). Differences in response durations for all patients with an objective response by both local and central evaluation are shown in Figure 3. For the majority of patients (94/115, 82%; indicated by bars in the middle section of Figure 3, the difference in duration of response was no more than 3 months or approximately one additional response evaluation. For 21 patients, the difference between assessors in duration of response was greater than 3 months; these patients are outlined in Figure 3 and shown in Figure 4 by reason for the difference. For the majority of patients (17/21), response duration was longer by local evaluation, which was primarily due to earlier assessments of disease progression by independent central review. Independent central review assessed progression according to CT-based assessments of new lesions and progression of existing lesions, as well as PET-based assessments of newly positive lesions and recurrence of PET activity in previously active lesions (Figure 4).

Per protocol, patients who had progressed in the opinion of the investigator were to be discontinued from treatment. To explore whether declarations of progression between the investigator and independent central review occurred closer together when progression did not require early treatment discontinuation (i.e. prior to Cycle 16), the difference between assessors in time to progression was compared separately for Cycles 1–15 and the long-term follow-up period. Progression dates were the same for the majority of patients with progression declared by both independent central review and local evaluation: 45% (27/60) of patients had progression called at the same time during Cycles 1–15 and 55% (6/11) of patients had progression called at the same during the follow-up period. However, in the subset of patients for whom progression dates were different, the median difference was approximately 2 months longer by local evaluation relative to independent central review for Cycles 1–15 (median difference = 2.1 months; range = –1.1 to 8.5 months) compared with the follow-up period (median difference = 0.2 months; range = –0.1 to 1.6 months).

A small number of progression events declared by independent central review were never documented by the local investigator despite extended follow-up. Based on the June 2013 data cut-off, which represents a median of approximately 3 years of observation time from first dose, 4 of 160 (2.5%) patients remained in remission per local evaluation more than 1

year (range = 17.2 to 27.1 months), and several additional response assessments (range = 5–7) after progression was declared by independent central review.

Discussion

We observed substantial concordance between independent central review and local evaluation in assessment of objective response across two international, open-label, single-arm, registration studies of brentuximab vedotin in patients with Hodgkin lymphoma or systemic anaplastic large-cell lymphoma. An assessment of objective response was made for 72% and 75% of Hodgkin lymphoma patients per local evaluation and independent central review, respectively. In patients with anaplastic large-cell lymphoma, objective response was reported for 83% and 86% of patients per local evaluation and independent central review, respectively. Differences in objective response assessments were primarily due to variability in the reduction of the sum of the products of diameters, which was attributable to the selection of different index lesions at baseline and variations in their measurements. The observation that nearly identical objective response rates between local and central evaluation were observed in both trials despite 11% (17/160) of patients having had different assessments of objective response argues against evaluation bias; nonetheless, a blinded trial showing agreement in the treatment effect estimate between local and independent assessments would be required to confirm the lack of a systematic investigator bias.

Variability was observed between local and central assessments in the time-to-event endpoints, which is not unexpected in advanced diseases where many lesions need to be followed. The progression-free survival curves for local evaluation stayed above those of independent central review in both the Hodgkin lymphoma and the anaplastic large-cell lymphoma studies; however, median progression-free survival in the anaplastic large-cell lymphoma study was almost identical between local and central evaluation. The decreased variability observed in the anaplastic large-cell lymphoma study could potentially be due to the smaller sample size relative to the Hodgkin lymphoma study or the slightly lower proportion of events in the anaplastic large-cell lymphoma study relative to the Hodgkin lymphoma study.

The most variability was observed in the duration of response endpoint, which, by incorporating the measurement of time to objective response as well as the measurement of time to progression, has additional capacity for differences among assessors. Although different measurements of time to response contributed to variability in the duration of response for a few patients, earlier CT- and PET-based declarations of progression by independent central review relative to local evaluation contributed to the majority of patients with significantly different (> 3 months) response durations.

A certain level of variability between central and local assessments of progression can be expected. Independent central review must rely on a predefined set of data and strictly adhere to a charter when determining progression. For example, the independent central review facility in this study was to make an assessment of progression if a fluorodeoxyglucose–PET positive site returned to being positive if it was previously negative, or if a new extranodal lesion, regardless of the size, appeared by CT. In contrast,

the treating oncologist may have been more likely to continue study treatment for one or two additional cycles if a responding patient remained asymptomatic and had new, but equivocal CT and/or PET findings, knowing that a score of progression would require discontinuation of the study drug. Although small sample sizes and different assessment schedules during the treatment and long-term follow-up period preclude strong conclusions, the observation that progression was assessed by local evaluation and independent central review more closely in time during the long-term follow-up period relative to the treatment period suggests that a concern over jeopardizing treatment may have partially contributed to the observed differences in response durations. Nonetheless, we observed that the majority of response durations per local and central evaluation fell within approximately one response evaluation of each other and that progression events declared by independent review, but never identified by local evaluation despite extended follow-up with continued assessments, were reassuringly rare. In these cases, independent review may have declared non-malignant, reversible inflammatory processes as progressions. Alternatively, lesions identified by independent review might have represented early tumor growth that subsequently diminished with continued brentuximab vedotin treatment. Regardless of the underlying reasons, this particular categorical difference between local and central evaluation only occurred in 2.5% of all enrolled patients. The quantitative differences observed in the time-to-event endpoints in this study, therefore, are likely due to the physician's ability to adopt a "wait-and-see" approach rather than a systematic qualitative difference that might be more indicative of an evaluation bias on the part of the investigator. However, a double-blinded trial showing a similar degree of variability between local and central assessments of progression across treatment arms would be needed to confirm the lack of differential bias.

Updated recommendations (the Lugano classification) for initial evaluation, staging, and assessment of response in patients with Hodgkin lymphoma and non-Hodgkin lymphoma have recently been published.¹⁵ Although the definitions of objective response are consistent with the earlier criteria, the recommendations include several clarifications on how response should be interpreted. For example, rather than using the mediastinal blood pool as the comparator for PET-based assessments as was recommended by the International Working Group criteria, the new guidelines recommend the use of a five-point Deauville scale. Given the observation that different interpretations of PET contributed to some of the variability observed in this retrospective analysis, it will be of interest to see whether adoption of the new guidelines promotes additional concordance in future trials of lymphoma.

These findings of a substantial degree of concordance of objective response across two different types of lymphoma provide preliminary support for local assessment of response with a limited independent review audit strategy in future trials in Hodgkin lymphoma and systemic anaplastic large-cell lymphoma. These findings require replication with additional and larger studies, blinded and unblinded, that could potentially provide the basis for the type of meta-analyses that were used to support the limited independent review audit strategy that has been proposed for solid tumors.⁴⁻¹⁰ In addition, similar analyses in more common subtypes of lymphoma, particularly those known to be less fluorodeoxyglucose-avid, would be required to evaluate a more general recommendation for this assessment strategy in hematological malignancies.

Acknowledgments

Financial support was provided to A.K.G. by the Hutchinson Center/University of Washington Cancer Consortium Cancer Center (support grant P30 CA015704), K24CA184039 (A.K.G.), and by philanthropic gifts from Frank and Betty Vandermeer. A.K.G. is a Scholar in Clinical Research for the Leukemia and Lymphoma Society. The authors wish to acknowledge Lisa Thomson for assistance in manuscript preparation, as an employee of Seattle Genetics, Inc.

A.K.G. has received research funding/grants from BMS, Gilead, Spectrum, Pfizer, Cephalon/Teva, Emergent/Abbott, Gilead, Janssen, Merck, Takeda Pharmaceuticals International Co., Piramal, Biogen Idec, and BioMarin. A.K.G. has acted as a consultant for Seattle Genetics, Inc., Pfizer, Gilead, Janssen, and Sanofi-Aventis. A.K.G. has received honoraria from Seattle Genetics, Inc., Takeda Pharmaceuticals International Co., and Sanofi-Aventis. A.E. has received honoraria from Seattle Genetics, Inc. and Takeda Pharmaceuticals International Co. and has received research funding/grants from Takeda Pharmaceuticals International Co. A.Y. has received honoraria from Seattle Genetics, Inc., Takeda Pharmaceuticals International Co., Celgene, Sanofi-Aventis, and Bayer. A.Y. has received research funding/grants from Curis, Janssen, and Novartis. B.P. has received travel expenses from Seattle Genetics, Inc. J.M.C. has received research funding/grants from Roche. E.L.S., D.A.K., and E.K.L. are, or have been, employed by and have equity ownership (including stock options) in Seattle Genetics, Inc. X.C. is employed by and has equity ownership (including stock options) in Takeda Pharmaceuticals International Co.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Direct funding for this research was issued by Seattle Genetics, Inc. through the joint financial support of Seattle Genetics, Inc. and Takeda Pharmaceuticals International Co.

References

1. Food and Drug Administration. Clinical trial endpoints for the approval of cancer drugs and biologics (FDA guidance for industry). Silver Spring, MD: Food and Drug Administration; 2007. p. 1-19.
2. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol*. 2007; 25:579–586. [PubMed: 17242396]
3. 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. [PubMed: 19097774]
4. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. *Eur J Cancer*. 2011; 47:1772–1778. [PubMed: 21429737]
5. Dodd LE, Korn EL, Freidlin B, et al. An audit strategy for progression-free survival. *Biometrics*. 2011; 67:1092–1099. [PubMed: 21210772]
6. Dodd LE, Korn EL, Freidlin B, et al. Blinded independent central review of progression-free survival in phase III clinical trials: important design element or unnecessary expense? *J Clin Oncol*. 2008; 26:3791–3796. [PubMed: 18669467]
7. European Medicines Agency. Appendix 1 to the guideline on the evaluation of anticancer medicinal products in man: methodological consideration for using progression-free survival (PFS) or disease-free survival (DFS) in confirmatory trials. 2011. http://www.ema.europa.eu/docs/en_GB/document_library/Other/2009/12/WC500017749.pdf
8. European Medicines Agency. Guideline on the evaluation of anticancer medicinal products in man. 2012. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/01/WC500137128.pdf
9. Food and Drug Administration. Summary minutes of the Oncologic Drugs Advisory Committee meeting. 2012. <http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/OncologicDrugsAdvisoryCommittee/UCM316460.pdf>
10. Zhang JJ, Chen H, He K, et al. Evaluation of blinded independent central review of tumor progression in oncology clinical trials: a meta-analysis. *Ther Innov Regul Sci*. 2013; 47:167–174.
11. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]

12. Pro B, Advani R, Brice P, et al. Brentuximab vedotin (SGN-35) in patients with relapsed or refractory systemic anaplastic large-cell lymphoma: results of a phase II study. *J Clin Oncol.* 2012; 30:2190–2196. [PubMed: 22614995]
13. Younes A, Gopal AK, Smith SE, et al. Results of a pivotal phase II study of brentuximab vedotin for patients with relapsed or refractory Hodgkin's lymphoma. *J Clin Oncol.* 2012; 30:2183–2189. [PubMed: 22454421]
14. Gopal AK, Chen R, Smith SE, et al. Three-year follow-up data and characterization of long-term remissions from an ongoing phase 2 study of brentuximab vedotin in patients with relapsed or refractory Hodgkin lymphoma. *Blood.* 2013; 122:4382.
15. 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. [PubMed: 25113753]

Figure 1. Comparison of independent central review and local evaluation discordant objective response assessments

CT: computed tomography; PET: positron emission tomography.

Individual patients are represented by bars shaded according to their best response (CR = complete remission, PR = partial remission, and SD = stable disease) per independent central review and local evaluation. Patients with Hodgkin lymphoma are designated as "HL" and patients with systemic anaplastic large-cell lymphoma are designated as "ALCL." Patients are sorted left to right (x axis) within a given category of discordance by the degree of difference in response (complete remission versus stable disease or partial remission versus stable disease); the direction of discordance (better by independent central review or local evaluation); and subsequently, by the degree of difference between independent central review and local evaluation in the percentage decrease in the sum of the products of diameters (SPD) (y axis).

Figure 2. Progression-free survival and duration of response

CI: confidence interval; PD: progressive disease.

Progression-free survival was defined as the time from start of study treatment to first documentation of objective tumor progression or to death due to any cause, whichever came first. Duration of response was defined as the time from start of the first documentation of objective tumor response (complete remission or partial remission) to the first subsequent documentation of objective tumor progression or to death due to any cause, whichever came first. Kaplan-Meier curves are presented for assessments by independent central review (ICR) (black line) and local evaluation (LE) (gray line). Symbols on the plot indicate

censored patients. Panels are presented separately for patients with Hodgkin lymphoma (HL—left panels; (a): progression-free survival, (c): duration of response) and patients with systemic anaplastic large-cell lymphoma (ALCL—right panels; (b): progression-free survival, (d): duration of response).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Difference in duration of response between independent central review and local evaluation

The 115 patients with an objective response by both independent central review and local evaluation are represented by bars sorted by the difference in response duration between assessors (local evaluation minus independent central review). Bars to the right of 0 represent longer durations by local evaluation and bars to the left of 0 represent longer durations by independent central review. Patients with Hodgkin lymphoma are indicated by solid bars and squares on 0 and patients with systemic anaplastic large-cell lymphoma are indicated by open bars and squares. Boxes at the top and bottom of the figure indicate the 21 patients with a difference in duration of response between independent central review and local evaluation of more than 3 months. Patients with no difference in duration of response between independent central review and local evaluation are shown as squares on 0.

Figure 4. Patients with 3 month difference between independent central review and local evaluation in duration of response

Duration of response shaded according to independent central review or local evaluation is shown for the 21 patients who had response durations that differed by more than 3 months. Patients with Hodgkin lymphoma are designated as “HL” and patients with systemic anaplastic large-cell lymphoma are designated as “ALCL.” The difference in response duration (local evaluation minus independent central review) is indicated in parentheses for each patient (positive numbers indicate response durations longer by local evaluation and negative numbers represent response durations longer by independent central review). Patients were censored (indicated by “+”) after the last radiological assessment that

documented the absence of progressive disease. Patients are sorted on the y axis according to the degree of difference between assessors within each subcategory of reasons for the difference (“Earlier Progression,” “Earlier Response,” or “Other”). The patient categorized as “Other” had an additional set of long-term follow-up scans available to independent central review for assessment of response, whereas the investigator’s response duration was censored back to the last on-treatment response evaluation.

Table 1

Overall objective response to treatment by independent central review and local evaluation.

	HL (N = 102)		ALCL (N = 58)	
	ICR	LE	ICR	LE
ORR, n (%)	76 (75)	73 (72)	50 (86)	48 (83)
95% CI	64.9, 82.6	61.8, 80.1	74.6, 93.9	70.6, 91.4
Best response, n (%)				
CR	35 (34)	34 (33)	34 (59)	35 (60)
95% CI	25.2, 44.4	24.3, 43.4	44.9, 71.4	46.6, 73.0
PR	41 (40)	39 (38)	16 (28)	13 (22)
SD	22 (22)	28 (27)	2 (3)	4 (7)
PD	3 (3)	0 (0)	3 (5)	2 (3)
HI	–	–	2 (3)	2 (3)
NE	1 (1)	1 (1)	1 (2)	2 (3)
Median time to objective response	5.7 weeks	5.7 weeks	5.9 weeks	5.7 weeks
Median time to CR	12.0 weeks	12.2 weeks	11.9 weeks	11.9 weeks

ALCL: anaplastic large-cell lymphoma; CI: confidence interval; CR: complete remission; HI: histologically ineligible, defined as patients determined to have anaplastic large-cell lymphoma by local assessment, but not by central pathology review; scored as non-responders per protocol; HL: Hodgkin lymphoma; ICR: independent central review; LE: local evaluation; NE: not evaluable; ORR: objective response rate; PD: progressive disease; PR: partial remission; SD: stable disease.

Table 2
Concordance in best response per independent central review and local evaluation.

Hodgkin lymphoma (N = 102)	Investigator					
	Independent Central Review	CR	PR	SD	PD	HI NE
CR		25 (25)	7 (7)	3 (3)	0	0
PR		8 (8)	28 (27)	5 (5)	0	0
SD		1 (1)	4 (4)	17 (17)	0	0
PD		0	0	3 (3)	0	0
NE		0	0	0	0	1 (1)
Kappa coefficient (95% CI)		0.55 (0.42, 0.68)				

Anaplastic large-cell lymphoma (N = 58)	Investigator					
	Independent Central Review	CR	PR	SD	PD	HI NE
CR		31 (53)	2 (3)	1 (2)	0	0
PR		4 (7)	10 (17)	1 (2)	0	1 (2)
SD		0	1 (2)	1 (2)	0	0
PD		0	0	1 (2)	2 (3)	0
HI		0	0	0	0	2 (3)
NE		0	0	0	0	1 (2)
Kappa coefficient (95% CI)		0.67 (0.50, 0.85)				

CI: confidence interval; CR: complete remission; HI: histologically ineligible, defined as patients determined to have anaplastic large-cell lymphoma by local assessment, but not by central pathology review; scored as non-responders per protocol; NE: not evaluable; PD: progressive disease; PR: partial remission; SD: stable disease.
The values given in boldface indicate patients who were discordant at the objective response level (13 Hodgkin lymphoma patients and 4 anaplastic large-cell lymphoma patients).