Quantitative Imaging Assessment for Clinical Trials in Oncology

Katherine E. Hersberger, PhD^{a,b}; Mishal Mendiratta-Lala, MD^c; Rocky Fischer, MS^b; Ravi K. Kaza, MD^c; Isaac R. Francis, MD^{b,c}; Mirabella S. Olszewski, MS^b; John F. Harju, MBA, PMP^b; Wei Shi, MBA, MS^b; Frank J. Manion, PhD^b; Mahmoud M. Al-Hawary, MD^{b,c}; and Vaibhav Sahai, MBBS, MS^{a,b}

ABSTRACT

Background: Objective radiographic assessment is crucial for accurately evaluating therapeutic efficacy and patient outcomes in oncology clinical trials. Imaging assessment workflow can be complex; can vary with institution; may burden medical oncologists, who are often inadequately trained in radiology and response criteria; and can lead to high interobserver variability and investigator bias. This article reviews the development of a tumor response assessment core (TRAC) at a comprehensive cancer center with the goal of providing standardized, objective, unbiased tumor imaging assessments, and highlights the web-based platform and overall workflow. In addition, quantitative response assessments by the medical oncologists, radiologist, and TRAC are compared in a retrospective cohort of patients to determine concordance. **Patients and Methods:** The TRAC workflow includes an image analyst who pre-reviews scans before review with a board-certified radiologist and then manually uploads annotated data on the proprietary TRAC web portal. Patients previously enrolled in 10 lung cancer clinical trials between January 2005 and December 2015 were identified, and the prospectively collected quantitative response assessments by the medical oncologists were compared with retrospective analysis of the same dataset by a radiologist and TRAC. **Results:** This study enlisted 49 consecutive patients (53% female) with a median age of 60 years (range, 29–78 years); 2 patients did not meet study criteria and were excluded. A linearly weighted kappa test for concordance for TRAC versus radiologist was substantial at 0.65 (95% CI, 0.46–0.85; standard error [SE], 0.10). The kappa value was moderate at 0.42 (95% CI, 0.20–0.64; SE, 0.11) for TRAC versus oncologists and only fair at 0.34 (95% CI, 0.12–0.55; SE, 0.11) for oncologists versus radiologist. **Conclusions:** Medical oncologists burdened with the task of tumor measurements in patients on clinical trials may introduce significant variability and investigator bias, with the potential to affect therapeutic response and clinical trial outcomes. Institutional imaging cores may help bridge the gap by providing unbiased and reproducible measurements and enable a leaner workflow.

J Natl Compr Canc Netw 2019;17(12):1505–1511
doi: 10.6004/jnccn.2019.7331

Background

Evaluation of overall response in oncology clinical trials is largely dependent on quantitative imaging¹ to assess the efficacy of cancer therapies. Patients with cancer may have lesions that require serial anatomic or functional imaging, such as MRI, CT, and/or PET, for response assessment. Objective radiographic assessment is defined per specific response criteria,^{2–13} which are based on the cancer type and location and the therapeutic classification of experimental drugs. The discovery of targeted and immunotherapeutic drugs and advancement in imaging methodology have led to the development of a multitude of objective response criteria.^{5,14–17} It is crucial to use the correct response metrics, because radiographic response is the most commonly used surrogate endpoint in oncology for assessment of therapeutic efficacy and patient outcome. The FDA published guidelines in 1994,¹⁸ with updates in 2004¹⁹ and 2018,²⁰ for standardizing the quantitative imaging assessment. In 2005, the NCI funded imaging response assessment teams across 8 cancer centers to improve the quality of quantitative imaging.²¹

The 2018 FDA document titled “Clinical Trial Imaging Endpoint Process Standards: Guidance for Industry” recommends that in “a randomized controlled trial... clinical trial primary endpoint image readers will be blinded to a subject’s treatment assignment because knowing the assignment would be presumed to create bias.”²⁰ This blinded independent central review (BICR) is usually completed by commercial imaging core laboratories that engage 2 radiologists per study to conduct independent image evaluation and have an additional radiologist available for adjudication of discordant results.²² Clinical trials that use central radiologic assessment²³ mostly perform these post hoc, because real-time review is challenging due to involved logistics and cost.²⁴ Interestingly, these retrospective BICR reviews have been shown to significantly affect both objective tumor response

^aDepartment of Internal Medicine, University of Michigan Medical School;

^bUniversity of Michigan Rogel Cancer Center; and ^cDepartment of Radiology, University of Michigan Medical School, Ann Arbor, Michigan.

See JNCCN.org for supplemental online content.

rates and survival compared with institutional or investigator assessments.^{25,26}

The workflow for completing the standardized imaging assessments for trials is different at each institution, and the onus has been shouldered by radiologists, oncologists, clinical trial staff, and/or imaging cores.^{27,28} In the absence of radiologists or imaging cores performing the assessment, the treating oncologist may either transcribe measurements from the subjective clinical radiology report or independently perform the measurements.^{27,29,30} However, medical oncologists are often inadequately trained for both imaging assessment and the application of study-specific response evaluation criteria, and prior studies have shown high interobserver variability.³¹ In addition, there is a potential for introduction of evaluation or experimenter bias in the assessment due to patient-provider relationship and knowledge of the treatment arm and clinical course.^{32,33}

At the University of Michigan Rogel Cancer Center, medical oncologists were previously burdened with response assessment. In an effort to improve data quality of imaging assessment and decrease investigator and coordinator/regulatory workload, we established a tumor response assessment core (TRAC) to provide standardized, objective, unbiased imaging assessment and consultative services for clinical trials. This article discusses our experience in establishing this core and highlights the workflow and the web-based platform, as well as compares the independent response assessments by oncologists, radiologist, and TRAC for patients enrolled in clinical trials at a major comprehensive cancer center to assess the potential benefit of an imaging core.

Patients and Methods

Tumor Response Assessment Core

TRAC is a collaborative effort between the departments of internal medicine and radiology at University of Michigan Rogel Cancer Center, and its development was supported by an internal grant from the cancer center from September 2015 through August 2017. The grant provided partial salary support for the informatics team, image analysts (IAs), and codirectors. The core was established in 2016 and draws on the clinical trial and radiologic expertise of its codirectors and several other board-certified radiologists and nuclear medicine physicians. The central mission of the imaging core is to provide independent, unbiased, and verifiable measurements of treatment response for patients enrolled in clinical trials and to serve as a centralized, web-based data resource to enable efficient internal and external auditing. The sustainability of TRAC is based on a fee-for-service model and largely depends on the revenue

generated from sponsored trials. The Rogel Cancer Center substantially subsidizes the costs for providing tumor response assessments for National Clinical Trials Network (NCTN) and investigator-initiated trials. The fee per scan was determined by modeling the current and expected number of patient accruals, type of funding of trials, and core budget. The fee is also determined by the type of response criteria and scan modality. A nominal remuneration per scan is transferred to the reading radiologist for reviewing tumor measurements for the imaging core. TRAC also provides consultative services for optimal use of quantitative imaging biomarkers for investigator-initiated trials at the University of Michigan. Since its inception in 2016, TRAC has provided service for >175 clinical trials with review of >1,500 scan time points. The Oncology Clinical Trials Support Unit (O-CTSUS) lean workflow assessment showed that, on average, the turnaround time for tumor measurements was reduced from 33 to 3 days.³⁴

IA and TRAC Workflow

We established the role of an IA as one who is responsible for reviewing orders, performing preassessment of scans followed by review with a radiologist, and then uploading finalized tumor response assessment data on the TRAC web portal (Figure 1). An IA is assigned to each clinical trial to review the research protocol and provide study-specific, role-based access to web portal-trained study investigators, coordinators, data managers, and sponsors.

The IAs recruited have a doctoral degree with anatomic and/or radiologic experience and are trained by board-certified faculty radiologists and by prior IAs across different specialties in anatomy, scan modalities, image assessment, and software, and undergo interobserver and intraobserver reliability tests across multiple modalities and response criteria to assure accuracy and consistency in image analysis. Imaging assessment by IA includes transferring the Digital Imaging and Communications in Medicine (DICOM) images into the patient archiving and communication system research server and then importing the images in the FDA-approved McKesson software (McKesson Corp) for MRI and CT scans or in MIM (MIM Software, Inc) for nuclear imaging. The IA evaluates clinic notes prior to review of the baseline scan for accurate identification of target and nontarget lesions as defined per specific criteria. A preliminary assessment of the scan is then completed, followed by review with an expertise-specific radiologist (eg, abdominopelvic, thoracic, head and neck) or nuclear medicine physician. The final measurement data with corresponding annotated images are then manually uploaded to the TRAC web portal for research staff and investigators to review.

Figure 1. TRAC workflow.

Abbreviations: IA, image analyst; RC, research coordinator; SI, study investigator; TRAC, tumor response assessment core.

The aforementioned workflow is duly completed within 72 business hours of a scan request. The treating investigator can then either e-sign if they agree with the report or trigger the discrepancy management workflow via the web portal if they disagree (see Figure 1).

TRAC Web Portal

The TRAC web portal was designed in collaboration with the cancer center informatics team. The portal supports the workflow of the core, which includes order request, work list management, annotated image upload, study team notifications, investigator e-signature, and protocol-specific result reporting (see supplemental eTable 1, available with this article at JNCCN.org). The key workflows and functions of the TRAC are shown in Figure 1. The software also provides administrative features, such as protocol registration (including type and number of response criteria), research personnel training documentation and type of access, and generation of automated billing reports.

The application is built in Java SE version 1.8 software (Oracle) running on SUSE Linux Enterprise Server 11, and the data are stored within the Oracle version 12c database. Due to institutional architectural standards, Java SE and Oracle were used; however, the core application can be implemented entirely using open-source components, such as Java-OpenJDK and PostgreSQL. Other open-source components, including Spring MVC and React/Redux, were incorporated for handling complex user interactions and user interface development. The application maintains a complete audit log for all

interactions per Code of Federal Regulations part 11 compliance guidelines. User authentication is controlled through the university identity management services and implemented based on lightweight directory access protocol. Authorization is controlled by the Spring Security framework and uses a role-based model consisting of separate roles for investigator, coordinator/data manager, analyst, sponsor/monitor, and core director. User access is granted only after completion of a mandatory web-based training module. Transport Layer Security is provided by Secure Socket Layer-coupled X.509 certificates. Secure backups of the system are performed hourly by infrastructure services provided by the information technology services group. Currently, interfaces with the electronic medical record and the clinical trial management system are under development.

Comparative Analysis Between Investigators, TRAC, and Radiologist

Study Population

A total of 49 consecutive patients aged ≥ 18 years who enrolled in a clinical trial at the University of Michigan Rogel Cancer Center for treatment of primary lung cancer were included in this comparative analysis study. Included subjects had at least one target lesion and underwent serial helical CT examinations at least 4 weeks apart between January 1, 2005, and December 31, 2015, using either unidimensional RECIST version 1.0, RECIST version 1.1, or bidimensional immune-related response

Table 1. Patient and Clinical Trial Characteristics

Characteristics	n
Patients, N	47
Median age (range), y	60 (29–78)
Sex	
Male	22
Female	25
Type of lung cancer	
Bronchogenic	4
Small cell	8
Non–small cell	31
Squamous cell	4
Type of response criteria	
RECIST version 1.0	25
RECIST version 1.1	17
irRC	5
Clinical trials, N	10
Trials based on type of response criteria	
RECIST version 1.0	3
RECIST version 1.1	6
Immune-related response criteria	1

criteria (see Table 1). Two patients were excluded because of missing images and having a follow-up scan date outside the 10-year study period. The cohort included 25 women (53%), and the median age of the overall cohort was 60 years (range, 29–78 years). Primary lung cancer morphologies included were non–small cell (n=31), small cell (n=8), squamous cell (n=4), and bronchogenic (n=4). This study was approved by the University of Michigan Institutional Review Board and was HIPAA-compliant. The study population was based on the trial database of the O-CTSU and the office of the lead academic participating site grant under the NCTN.

Image Assessment

Patients had a baseline scan and at least one follow-up scan. The assessments completed by 5 thoracic medical oncologists were retrospectively collected from paper charts and electronic case report forms. The TRAC assessment was completed prospectively by an IA in conjunction with 2 board-certified radiologists as per the TRAC workflow described earlier. Another board-certified radiologist prospectively performed an independent review of the dataset. All assessments were blinded to each other. Target lesions were chosen per trial-specific criteria with respect to size and reproducibility based on baseline CT examinations.

Statistics

Assessments were conducted using the GraphPad QuickCalcs (GraphPad Software) statistical package, and results were compared using the linearly weighted kappa test for concordance of responses. The kappa test uses the Landis and Koch scale to evaluate degree of concordance: 0.21 to 0.40 is fair, 0.41 to 0.60 is moderate, 0.61 to 0.80 is substantial, and 0.81 to 1.00 is almost perfect.³⁵

Results

Patient and clinical trial characteristics are detailed in Table 1. All 47 patients were included in the comparison of the response assessment. Patient responses as seen on the follow-up scans were defined as complete response, partial response, stable disease, or progressive disease per protocol-specific response criteria (ie, RECIST version 1.0, RECIST version 1.1, immune-related response criteria) and then directly compared for interreader variability. A linearly weighted kappa test for concordance for TRAC versus radiologist was classified as substantial at 0.65 (95% CI, 0.46–0.85; standard error [SE], 0.10), whereas it was moderate at 0.42 (95% CI, 0.20–0.64; SE, 0.11) for TRAC versus oncologists and fair at 0.34 (95% CI, 0.12–0.55; SE, 0.11) for oncologists versus radiologist (Table 2). The counts of observed agreements were 37 (78.7%), 31 (66.0%), and 28 (59.6%) between the comparison groups noted above, respectively. A detailed review of the target lesions as assessed by the oncologists revealed assignment of nonmeasurable/nonreproducible lesions as target on baseline scans in 3 patients (6.4%) and inaccurate interpretation or application of the response criteria in 12 (25.5%) (Table 3).

Discussion

Objective radiographic assessment is essential for accurate evaluation of patient outcomes in clinical trials dependent on imaging as a surrogate endpoint for therapeutic efficacy. Imaging assessment workflows can be complex, can vary with institution, and may increase the workload of medical oncologists, who are often inadequately trained in radiology and response criteria; this can therefore lead to high interobserver variability together with investigator bias. We reviewed the development and use of an imaging core to provide unbiased, reproducible, standardized response assessment of imaging scans for patients enrolled in clinical trials at our institution. We noted only fair and moderate concordance between the quantitative imaging assessment completed by the treating medical oncologists compared with the radiologist and imaging core, respectively. There were also multiple discrepancies noted in the application of protocol criteria and selection of target lesions. In comparison, the blinded independent assessments by TRAC and the radiologist had a

Table 2. Comparison of Response Assessment in Per-Patient Analysis (N=47)

Response Assessment ^a	Imaging Core n	Medical Oncologists n	Radiologist n
Complete response	0	0	0
Partial response	15	17	13
Stable disease	25	26	25
Progressive disease	2	3	4
No target lesions identified on baseline scan ^b	5	1	5
Response Comparison	Imaging Core vs Radiologist	Imaging Core vs Medical Oncologists	Medical Oncologists vs Radiologist
Linearly weighted kappa value	0.65 (95% CI, 0.46–0.85)	0.42 (95% CI, 0.20–0.64)	0.34 (95% CI, 0.12–0.55)
Kappa scale	Substantial	Moderate	Fair
Observed agreements, n (%)	37 (78.72)	31 (65.96)	28 (59.57)

^aPer specific criteria (outlined in Table 1).

^bPatients included in retrospective analysis despite no target lesion at baseline, because patients were enrolled in trials per oncologist.

kappa concordance value of 0.65 (substantial; 95% CI, 0.46–0.85) (Table 2). A significant level of discordance has the potential to affect patient treatment and trial outcome. These discrepant data underscore the need for improved imaging criteria training for medical oncologists, consideration for radiologist interpretation, and/or development of an imaging core for response assessment.

The workflow for obtaining quantitative imaging assessment for patients enrolled in clinical trials at each institution is different. The responsibility lies with either radiologists, oncologists, clinical trial support staff, or imaging cores.^{27,28} The lack of uniformity is multifactorial, and includes budgetary constraints, lack of standardized reporting,³⁶ and perhaps inadequately written clinical trial protocols. The subjective clinical radiologic report is not a substitute for the standardized assessment^{27,37} and should not be used to transcribe measurements from the report by the oncologist or study coordinator.^{29,30} In addition, medical oncologists with limited training in imaging or application of imaging criteria may have high interobserver variability in their

assessment compared with other oncologists and radiologists.³¹ These alternate workflows for response assessment may lead to inefficiency, high variability, and low interreader agreement. The role of a single IA per trial can serve as a catalyst in the relationship between radiologists, oncologists, and clinical staff, enabling improved reliability, decreased interreader variability, and faster turnaround time without experimenter bias.³² The IA also eliminates investigator bias that may be introduced by the oncologists due to their relationship with their patients and clinical trials. More robust data in turn provide greater confidence in determining therapeutic tumor response in clinical trials, which is crucial in this era of precision medicine with smaller cohorts.

BICRs offer a structured review process to eliminate evaluation or experimenter bias²² and reduce measurement variability. However, most clinical trials usually do not use the BICRs for real-time response assessment due to the involved cost and logistics. BICRs have shown major discrepancy rates in progression-free survival²⁵ and response rates^{22,26} compared with site evaluations, and have the potential to introduce informative censoring because the review is often retrospective and lacks confirmation of subject eligibility and progression.²⁵ In contrast, variation has also been observed for radiologist reviewers in the selection, classification, and measurement of nodal and nonnodal lesions, which can lead to significant rates of discordance.^{22,38–40} These issues have led some to believe that a single radiologist trained in radiologic tumor measurements and blinded to patient outcome and treatment plan would be best suited to evaluate tumor response.⁴¹ However, at our cancer center, the O-CTSU and medical oncologists experienced challenges in identifying a single specialty-specific radiologist per trial. With development of the imaging core, we established study-specific IAs to decrease

Table 3. Discrepancies at Baseline Scan Evaluation in Oncologist Assessment

Response Criteria	Reason for Discrepancy	n
RECIST version 1.0	Patient enrolled without target lesions	1
RECIST version 1.1	Lymph node <15-mm short axis included as target	3
RECIST version 1.0	Target lesion not considered reproducible	3
RECIST version 1.0	Lesion <10 mm included as target	1
RECIST version 1.1	Lesion <10 mm included as target	4
RECIST version 1.1	>2 target lesions included per organ	3
RECIST version 1.1	>5 target lesions included per scan	1

this interreader variability. In addition, it has led to a more nuanced understanding of the response criteria within each assigned clinical trial protocol and improved communication with the study teams. A few other cancer centers have developed similar cores and web-based platforms, such as the Quantitative Imaging Analysis Core at the University of Texas MD Anderson Cancer Center, OncoRad at the University of Washington, and the Tumor Imaging Metrics Core at Dana-Farber/Harvard Cancer Center.

Our investigation has a few limitations, including comparison of a prospectively compiled dataset by several medical oncologists for patients enrolled in 10 different clinical trials versus retrospective analyses of the same data performed by an imaging core. In addition, there is inherent variability in the prospectively collected data that include measurements from several medical oncologists. However, the premise of this study was to compare the quality of data collected across multiple investigators with that gathered by a study-specific IA in conjunction with radiologists. The dataset did include consecutive patients, and the radiologist and imaging core measurements were performed in a blinded manner to reduce bias. As noted, our dataset included only patients with lung cancer, and therefore we are unable to generalize our findings across all cancers. Additionally, due to the small dataset, we were unable to evaluate whether the difference in response assessment may have affected trial outcome. Lastly, the independent radiologist assessment was performed by a single radiologist, and data were not reviewed by another radiologist for accuracy. Interobserver variability exists even among radiologists, and therefore it is even more crucial to engage radiologists who keep abreast of the nuances of the many imaging criteria in oncology and their continually changing modifications.

Conclusions

Among patients with lung cancer previously enrolled in clinical trials, quantitative response assessments prospectively collected by medical oncologists were compared with retrospective analysis of the same dataset by a radiologist and an imaging core. Results show substantial concordance between the study radiologist and imaging core but only fair and moderate concordance between them and the medical oncologists, respectively. These findings indicate that it is crucial to either engage radiologists or develop institutional imaging cores such as TRAC to provide unbiased, reproducible, and longitudinal records of lesion measurements for oncology clinical trials with a surrogate imaging endpoint.

Acknowledgments

The authors acknowledge the investigators/medical oncologists whose data were used retrospectively in this study. In addition, the authors acknowledge Dr. Bin Nan, Professor of Biostatistics and Statistics, who provided consultation on the statistics for this project.

Submitted June 12, 2018; accepted for publication June 18, 2019.

Author contributions: *Study design:* Hersberger, Sahai. *Core design:* Hersberger, Fischer, Francis, Olszewski, Harju, Shi, Manion, Sahai. *Data acquisition:* Hersberger, Mendiratta-Lala, Kaza, Francis, Al-Hawary, Sahai. *Statistics:* Sahai. *Manuscript preparation:* Hersberger, Fischer, Harju, Sahai. *Manuscript editing:* All authors.

Disclosures: The authors have disclosed that they have not received any financial considerations from any person or organization to support the preparation, analysis, results, or discussion of this article.

Funding: Research reported in this publication was supported by the NCI of the NIH under award number P30CA046592. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Correspondence: Vaibhav Sahai, MBBS, MS, Department of Internal Medicine, University of Michigan Medical School, 1500 East Medical Center Drive, C412 Med Inn Building, Ann Arbor, MI 48109. Email: vsahai@med.umich.edu

References

- Rosenkrantz AB, Mendiratta-Lala M, Bartholmai BJ, et al. Clinical utility of quantitative imaging. *Acad Radiol* 2015;22:33–49.
- Miller AB, Hoogstraten B, Staquet M, et al. Reporting results of cancer treatment. *Cancer* 1981;47:207–214.
- Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. *J Natl Cancer Inst* 2000; 92:205–216.
- 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.
- Choi H, Chamsangavej C, Faria SC, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol* 2007;25:1753–1759.
- Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. *J Clin Oncol* 1999;17:1244.
- Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;25:579–586.
- 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: 3059–3067.
- Wahl RL, Jacene H, Kasamon Y, et al. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med* 2009;50(Suppl 1):122S–150S.
- Wen PY, Macdonald DR, Reardon DA, et al. Updated response assessment criteria for high-grade gliomas: Response Assessment in Neuro-Oncology Working Group. *J Clin Oncol* 2010;28: 1963–1972.
- Lin NU, Lee EQ, Aoyama H, et al. Response assessment criteria for brain metastases: proposal from the RANO group. *Lancet Oncol* 2015;16: e270–278.
- Scher HI, Halabi S, Tannock I, et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. *J Clin Oncol* 2008;26:1148–1159.
- Scher HI, Morris MJ, Stadler WM, et al. Trial design and objectives for castration-resistant prostate cancer: updated recommendations from the

- Prostate Cancer Clinical Trials Working Group 3. *J Clin Oncol* 2016; 34:1402–1418.
14. Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin Cancer Res* 2009;15:7412–7420.
 15. Nishino M, Giobbie-Hurder A, Gargano M, et al. Developing a common language for tumor response to immunotherapy: immune-related response criteria using unidimensional measurements. *Clin Cancer Res* 2013;19:3936–3943.
 16. Seymour L, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* 2017;18:e143–152.
 17. Okada H, Weller M, Huang R, et al. Immunotherapy response assessment in neuro-oncology: a report of the RANO working group. *Lancet Oncol* 2015;16:e534–542.
 18. On and offsite image reads: is basing drug efficacy on the site read risky business? Available at: <http://www.appliedclinicaltrials.com/and-offsite-image-reads>. Accessed October 14, 2019.
 19. U.S. Food and Drug Administration. Guidance Document. Developing medical imaging drug and biological products part 3: design, analysis, and interpretation of clinical studies. Available at: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/developing-medical-imaging-drug-and-biological-products-part-3-design-analysis-and-interpretation>. Accessed October 14, 2019.
 20. U.S. Food and Drug Administration. Clinical trial imaging endpoint process standards: guidance for industry. Available at: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-imaging-endpoint-process-standards-guidance-industry>. Accessed October 14, 2019.
 21. Graham MM, Badawi RD, Wahl RL. Variations in PET/CT methodology for oncologic imaging at U.S. academic medical centers: an imaging response assessment team survey. *J Nucl Med* 2011;52:311–317.
 22. Ford R, Schwartz L, Dancey J, et al. Lessons learned from independent central review. *Eur J Cancer* 2009;45:268–274.
 23. Koshkin VS, Bolejack V, Schwartz LH, et al. Assessment of imaging modalities and response metrics in Ewing sarcoma: correlation with survival. *J Clin Oncol* 2016;34:3680–3685.
 24. Nygren P, Blomqvist L, Bergh J, et al. Radiological assessment of tumour response to anti-cancer drugs: time to reappraise. *Acta Oncol* 2008; 47:316–318.
 25. 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.
 26. Von Hoff DD, Ervin T, Arena FP, et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *N Engl J Med* 2013;369:1691–1703.
 27. Sullivan DC. Imaging as a quantitative science. *Radiology* 2008;248: 328–332.
 28. Yankeelov TE, Mankoff DA, Schwartz LH, et al. Quantitative imaging in cancer clinical trials. *Clin Cancer Res* 2016;22:284–290.
 29. Folio LR, Nelson CJ, Benjamin M, et al. Quantitative radiology reporting in oncology: survey of oncologists and radiologists. *AJR Am J Roentgenol* 2015;205:W233–243.
 30. Sevenster M, Chang P, Bozeman J, et al. Radiologic measurement dictation and transcription error rates in RECIST (Response Evaluation Criteria in Solid Tumors) clinical trials: a limitation of the radiology narrative report to accurately communicate quantitative data. Presented at the Radiological Society of North America 2013 Scientific Assembly and Annual Meeting; December 1–6, 2013; Chicago, IL.
 31. Shao T, Wang L, Templeton AJ, et al. Use and misuse of waterfall plots. *J Natl Cancer Inst* 2014;106:dju331.
 32. Urban T, Zondervan RL, Hanlon WB, et al. Imaging analysts: how bringing onboard multimodality trained personnel can impact oncology trials. Available at: <http://www.appliedclinicaltrials.com/imaging-analysts?id=&sk=&date=&pageID=5>. Accessed October 14, 2019.
 33. Tang PA, Pond GR, Chen EX. Influence of an independent review committee on assessment of response rate and progression-free survival in phase III clinical trials. *Ann Oncol* 2010;21:19–26.
 34. Hersberger KE, Fischer R, Bebee PA, et al. On TRAC at the Rogel Cancer Center: centralized trial imaging metrics system. Presented at the Association of American Cancer Institutes Clinical Research Innovation Meeting; July 11–12, 2018; Chicago, Illinois. Available at: <https://www.aaci-cancer.org/Files/Admin/CRI/2018-Abstracts-and-Posters.pdf>.
 35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–174.
 36. Jaffe TA, Wickersham NW, Sullivan DC. Quantitative imaging in oncology patients: part 2, oncologists' opinions and expectations at major U.S. cancer centers. *AJR Am J Roentgenol* 2010;195:W19–30.
 37. Rubin DL, Willrett D, O'Connor MJ, et al. Automated tracking of quantitative assessments of tumor burden in clinical trials. *Transl Oncol* 2014;7:23–35.
 38. Keil S, Barabasch A, Dirrichs T, et al. Target lesion selection: an important factor causing variability of response classification in the Response Evaluation Criteria for Solid Tumors 1.1. *Invest Radiol* 2014; 49:509–517.
 39. Yoon SH, Kim KW, Goo JM, et al. Observer variability in RECIST-based tumour burden measurements: a meta-analysis. *Eur J Cancer* 2016;53: 5–15.
 40. Gierada DS, Pilgram TK, Ford M, et al. Lung cancer: interobserver agreement on interpretation of pulmonary findings at low-dose CT screening. *Radiology* 2008;246:265–272.
 41. 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.

See JNCCN.org for supplemental online content.

Supplemental online content for:

Quantitative Imaging Assessment for Clinical Trials in Oncology

Katherine E. Hersberger, PhD; Mishal Mendiratta-Lala, MD; Rocky Fischer, MS; Ravi K. Kaza, MD;
Isaac R. Francis, MD; Mirabella S. Olszewski, MS; John F. Harju, MBA, PMP; Wei Shi, MBA, MS;
Frank J. Manion, PhD; Mahmoud M. Al-Hawary, MD; and Vaibhav Sahai, MBBS, MS

J Natl Compr Canc Netw 2019;17(12):1505–1511

eTable 1: Clinical Trial Response Criteria Supported by TRAC

eTable 1. Clinical Trial Response Criteria Supported by TRAC

Response Evaluation Criterion	Trials, n
Cheson et al ¹	11
Choi et al ²	1
Hairy cell leukemia ^{3,a}	1
International Myeloma Working Group ^{4,a}	1
Immune RECIST ⁵	7
Immune-related RECIST ⁶	25
Immune-related response criteria ⁷	4
International Working Group – chronic lymphocytic leukemia ^{8,a}	1
Lugano ^{9,a}	19
Modified RECIST for hepatocellular carcinoma ¹⁰	1
Prostate Cancer Clinical Trials Working Group 2 ¹¹	1
Prostate Cancer Clinical Trials Working Group 3 ¹²	5
Response Assessment in neuro-oncology ¹³	5
Response Assessment in neuro-oncology – brain metastases ¹⁴	2
Response evaluation criteria in lymphoma ¹⁵	1
RECIST version 1.1 ¹⁶	86
Volumetric analysis (spleen and liver) for myeloproliferative neoplasms ^{17,a}	3
Waldenström's macroglobulinemia response criteria ^{18,a}	1
WHO ¹⁹	1

Abbreviation: TRAC, tumor response assessment core.

^aImaging criteria subset of global response.

References

- Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;25:579–586.
- Choi H, Charnsangavej C, Faria SC, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol* 2007;25:1753–1759.
- Robak T, Matutes E, Catovsky D, et al. Hairy cell leukaemia: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2015;26(Suppl 5):v100–107.
- Kumar S, Paiva B, Anderson KC, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol* 2016;17:e328–346.
- Seymour L, Bogaerts J, Perrone A, et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* 2017;18:e143–152.
- Nishino M, Giobbie-Hurder A, Gargano M, et al. Developing a common language for tumor response to immunotherapy: immune-related response criteria using unidimensional measurements. *Clin Cancer Res* 2013;19:3936–3943.
- Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. *Clin Cancer Res* 2009;15:7412–7420.
- 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.
- 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;32:3059–3067.
- Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. *Semin Liver Dis* 2010;30:52–60.
- Scher HI, Halabi S, Tannock I, et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. *J Clin Oncol* 2008;26:1148–1159.
- Scher HI, Morris MJ, Stadler WM, et al. Trial design and objectives for castration-resistant prostate cancer: updated recommendations from the Prostate Cancer Clinical Trials Working Group 3. *J Clin Oncol* 2016;34:1402–1418.
- Okada H, Weller M, Huang R, et al. Immunotherapy response assessment in neuro-oncology: a report of the RANO working group. *Lancet Oncol* 2015;16:e534–542.
- Lin NU, Lee EQ, Aoyama H, et al. Response assessment criteria for brain metastases: proposal from the RANO group. *Lancet Oncol* 2015;16:e270–278.
- 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.
- 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.
- Tefferi A, Cervantes F, Mesa R, et al. Revised response criteria for myelofibrosis: International Working Group-Myeloproliferative Neoplasms Research and Treatment (IWG-MRT) and European LeukemiaNet (ELN) consensus report. *Blood* 2013;122:1395–1398.
- Owen RG, Kyle RA, Stone MJ, et al. Response assessment in Waldenström macroglobulinaemia: update from the VIth International Workshop. *Br J Haematol* 2013;160:171–176.
- Miller AB, Hoogstraten B, Staquet M, et al. Reporting results of cancer treatment. *Cancer* 1981;47:207–214.