Everything you need to know:

Overview of Imaging Endpoints and FDA Guidelines for Use of Imaging in Oncology Clinical Trials

The Median Technologies Innovation Team

Contents

The Median Technologies Innovation Team

Fredrik Brag
CEO, Chairman and Co-founder
fredrik.brag@mediantechnologies.com

Jeanne Hecht
Chief Operating Officer
jeanne.hecht@mediantechnologies.com

Nicholas Campbell
Chief Commercial Officer
nicholas.campbell@mediantechnologies.com

Conducting a clinical trial in oncology is a major investment of both time and money, with pharmaceutical companies spending nearly 10 years and $1.4 billion ($2.6 billion if capital costs are taken into account) to develop a new prescription drug. Given these expenditures, it's not surprising that sponsors are driven to streamline and optimize the clinical trial process, as early identification of failed therapies can save significant time and money.

In oncology, therapeutic benefit is typically assessed by measuring tumor size and growth in response to treatment, and is monitored through various medical imaging modalities. Imaging is a powerful tool for visualizing cancer, but it must be performed quantitatively using appropriate response criteria in order to generate data that are comparable among trials. In addition, imaging's complex and variable nature can produce unreliable results if standardization practices are not implemented and carefully followed. A thorough understanding of quantitative imaging, its potential for variability, and the standardization guidelines necessary to produce data that reliability detects drug treatment effects are key to running a successful imaging trial in oncology.

Introduction

Medical imaging is an integral component of oncology clinical trial design. In late phase trials, imaging is used to assess trial endpoints that indicate clinical benefit, such as progression-free survival (PFS) and overall response rate (ORR), particularly when an overall survival (OS) endpoint is not practical due to limited patient numbers or the need for prolonged monitoring. Imaging in late phase trials measures changes in tumor size and burden as indicators of disease progression (i.e., new or enlarging tumor lesions) or response to treatment, (i.e., tumor shrinkage). These endpoints are valid surrogates for clinical benefit because direct correlations have been observed between tumor shrinkage and patient survival. [Park 2015; Jain 2012]

In contrast, early phase trials incorporate imaging to evaluate the extent of disease, detect the presence of biomarkers that represent early surrogate endpoints, determine trial eligibility, identify patients that are most likely to respond to treatment, and make go/no go decisions on the continuation of a given trial. The use of imaging data in the early phases is critical for producing faster, safer, and more efficient later phase trials. [Heertum 2015]

Qualitative Versus Quantitative Imaging

Imaging can produce both qualitative and quantitative assessment of disease. Qualitative imaging is a visually based, subjective evaluation of the region of interest that produces text notes (e.g., mild worsening of disease), while quantitative imaging is an objective, numbers-based measurement. Although qualitative measurements may be satisfactory in the clinic when evaluating an individual, only quantitative measurements can be used when reporting clinical trial results and comparing results across trials. The World Health Organization (WHO) was the first medical organization to promote the use of quantitative imaging in clinical trials by designing and implementing a set of objective criteria for evaluating tumor response to treatment. The WHO criteria were adopted in 1979 and have since formed the framework for the standardization of reporting response for patients with cancer (see Table 1).

Choosing the Best Imaging Modality For Your Study

1. Determine which modality will be best suited to evaluate the drug's therapeutic effects by considering its molecular target, mechanism of action, and pharmacokinetic characteristics.
2. Consider the modality's accessibility, ease-of-use, reliability, spatial coverage, image reproducibility, and cost.
3. Balance benefits with the assessment of risk to the patient, such as potential adverse effects of a radiotracer or contrast agent, radiation dose, or general discomfort.

[Korn 2013, Lui 2015, Morgan 2011]

Quantitative Imaging Techniques

The most commonly used quantitative imaging techniques in oncology trials are standard CT and MRI, which can be used to measure lesion size, detect the appearance of new lesions, and provide anatomical information (i.e., tumor localization and structural evaluation). These modalities offer ease of use and are widely available at clinical trial sites.

Advanced imaging techniques, such as PET, SPECT, dynamic contrast-enhanced MRI, and perfusion CT, provide both anatomical and functional information, including metabolic activity, expression of specific molecular targets, cell proliferation, and hypoxia. [Heertum 2015; Morgan 2011] These modalities often provide additional insights into tumor physiology or microenvironment. For example, in patients with colorectal cancer, perfusion CT can identify areas of hypoxia, which has been shown to be an indicator of poor prognosis. [Goh 2014] Quantitative imaging can also be used to stratify patients into treatment groups that are most likely to receive therapeutic benefit. Use of the PET tracer 18F-fluoro-17β-estradiol (FES) in patients with breast cancer can detect ER-positive tumors to determine a patient's ER status, thus guiding subsequent treatment or trial enrollment decisions. [Heertum 2015] Modalities are often combined (i.e., PET/CT and PET/MRI) to simultaneously provide anatomical and physiological data in a single imaging session. For example, PET/CT imaging with the radiotracer 18F-fluoro-2-deoxy-D-glucose (FDG) measures glucose metabolism, which is frequently upregulated in malignant cells, and is used to assess treatment response in lymphoma. [Cheson 2014].

How Do We Assess Response?

As mentioned previously, the WHO criteria were the first set of guidelines for quantitative evaluation of tumor burden. WHO criteria assess tumor response by calculating the sum of bidimensional measurements (longest diameter multiplied by the longest perpendicular diameter) of target lesions. However, an updated and simpler set of guidelines became necessary as imaging technologies evolved. In 2000, the Response Evaluation Criteria in Solid Tumors (RECIST) were developed, which assessed response using unidimensional measurements (longest diameter) of tumor size in a defined number of target lesions. RECIST was revised in 2009 (RECIST1.1) to include updated recommendations on the number of lesions to assess, measurement of lymph nodes, and parameters for defining disease progression. [Eisenhauer 2009] Today, RECIST1.1 is the most widely used set of criteria for assessing response in solid tumors, and is an acceptable imaging marker for Phase III approval by the FDA.

RECIST1.1 at a glance

Lesion Measurement	Unidimensional
Baseline Lesion Size	= >10 mm
Baseline Lesion Number	5 lesions total, 2 per organ
Appearance of New Lesions	Always represents PD
Response	CR = disappearance of all lesions
	PR = >30% decrease in SoD
	SD = when neither PR nor PD can be established
	PD = >20% increase in SoD (minimum 5 mm)

CR = complete response; PR = partial response; SD = stable disease; PD = progressive disease; SoD = sum of diameters

Alternate Response Criteria

For certain drug classes or types of cancers, evaluating response with RECIST1.1 is not practical or meaningful. This is particularly true for some modern drug targets (e.g., immunotherapies, tyrosine kinase inhibitors, and angiogenesis blockers) that tend to stabilize disease but not cause tumor shrinkage, as well as for tumors that do not grow in easily measurable spheres, or their invasive natures make identification of tumor margins difficult. [Sharma 2012; Yankeelov 2016; Villaruz 2013] In these cases, alternative evaluation criteria should be used to accurately assess drug benefit.

Response Criteria in Oncology

Response Criteria	For Use With	Date Established	Notes
WHO	Solid tumors	1979	Original response criteria using bidimensional measurements, the framework for standardization of reporting of response for patients with cancer
RECIST	Solid tumors	2000	Gold standard for evaluating response in solid tumors; Revised guidelines (version1.1) lowered the number of measurable lesions and accounts for measurement of lymph nodes
RECIST1.1	Solid tumors	2009	for measurement of lymph nodes Modified WHO criteria that account for delayed response and pseudoprogression with immunotherapies
irRC	Immunotherapies	2009	Modified WHO criteria that account for delayed response and pseudoprogression with immunotherapies
irRECIST	Immunotherapies	2013	Modified irRC and RECIST criteria that account for delayed response and pseudoprogression; directly comparable to RECIST1.1
MacDonald	Glioma, astrocytoma	1990	Modified WHO for malignant glioma, later modified to be RANO
RANO	Glioma, astrocytoma	2010	Accounts for irregular growth patterns, the presence of cystic cavities, pseudoprogression; Uses bidimensional measurements with contrast-enhanced MRI
RANO-BM	Parenchymal brain metastasis	2013	Modified RECIST to evaluate brain metastases
mRECIST	Hepatocellular carcinoma	2010	Accounts for tumor necrosis by not including necrotic portion of lesions in measurement
Cheson	Lymphoma (non-Hodgkin's)	1999	Original response criteria for non-Hodgkin's lymphoma using bidimensional measurements
IWG-Cheson	Lymphoma (non-Hodgkin's)	2007	Includes FDG-PET to characterize response
Lugano Classification	Lymphoma (non-Hodgkin's)	2014	Uses FDG-PET to define response with or without CT/MRI
NCI-WG	Chronic lymphocytic leukemia	1996	Defines role of imaging in clinical trials and research
IWCLL	Chronic lymphocytic leukemia	2008	Updates the 1996 NCI-WG criteria
PERCIST	18F-FDG-PET	2009	Criteria for assessing metabolic tumor response using FDG-PET
Choi	GIST	2007	Incorporates tumor attenuation measurement to account for tumor necrosis
Byrne	Mesothelioma	2004	Defines measurement criteria of mesothelioma to account for patterns of appearance on cross-sectional imaging
PCWG-2	Prostate cancer	2011	Defines criteria to assess visceral and bone disease

The Benefits of Imaging

Medical imaging offers many benefits as a tool to evaluate therapeutic efficacy. First, images provide a visual assessment of disease by identifying the exact anatomical location(s) and extent of the target lesion(s).

This helps clinicians obtain biopsies and guides subsequent treatment decisions. In addition, the noninvasive nature of imaging permits continual assessment of disease staging and response throughout the treatment period, unlike invasive blood- or tissue-based assays that may create unnecessary trauma for the patient or be restricted to specific time points. Lastly, the digital format allows for image quantitation, automated measurements, and real-time transmission from the clinical site to the analysis site or trial sponsor. The quantitative measurements increase interpretation accuracy by eliminating subjective assessments, while the digital format streamlines the analytical process through faster and more efficient transfer of digital data versus analog data. [Korn 2013]

The Challenge: Variability

Imaging platforms can contribute powerful data to a clinical trial, but decisions on therapeutic efficacy can only be made based on data that is consistent, reliable, and accurate. The complex and inherently tomographic nature of CT, MRI, and other imaging modalities makes them prone to variability, and controlling for variability is one of the greatest challenges clinical researchers will face when conducting imaging-based trials.

Variability can occur during both image capture and interpretation. Variability introduced in the image capture stage can be due to differences across clinical trial sites in the type and working order of the scanner used; the type, volume, and administration of contrast agent; the choice of slice thickness or other measurement parameters (e.g., unidimensional versus volumetric measurements), or the way patients are handled or positioned. [Zhao 2013, Sullivan 2013]

However, one of the greatest sources of variability stems from reader interpretation, which includes inter- and intra-reader variability, reader bias, or reader fatigue.

Intra-reader variation	read differences from a single reader, usually stems from the inherent difficulty of a case
Inter-reader variation	interpretation differences among multiple readers, usually due to reader skill
Reader bias	reader knowledge of treatment conditions influences interpretation of images
Reader fatigue	a reduced capacity to correctly interpret images after long work hours or reading a large number of scans

As expected, inter-reader variability has been found to be higher than intra-reader variability. [Zhao 2013] One of the primary causes of discordance among readers is the selection of target lesions. Multiple readers may not always select the same lesions for measurement, and choice of lesion or lesion location can impact a reader's ability to consistently interpret an image. [Krajewski 2014] Studies have shown that lesions with unclear boundaries produce more measurement variability than those with well-defined margins. [Muenzel 2012] Irregularly shaped tumors are similarly difficult to measure consistently. Variability in lesion measurement can also be attributed to tumor size, particularly when lesions are small, as measurement variability increases as lesion size decreases. [Oxnard 2011] Although RECIST1.1 recommends choosing lesions greater than 10 mm, lesion size can sometimes drop below this threshold after being subjected to beneficial treatment.

Inter-reader variation is a serious issue that if left unchecked can lead to misclassification of RECIST response, either through inaccurate declaration of progressive disease that can prematurely remove a patient from a trial or the inability to recognize true therapeutic benefit. [Yoon 2016, Oxnard 2011] Technology that offers automated lesion selection, measurement, and tracking can overcome many of these issues, greatly reducing variability. One study found that using computer-assisted measurements reduced inter-reader variability by half as compared with manual measurements. [Dinkel 2013] Measurement software can also decrease the time required for image assessment and creation of a study report. [Muenzel 2012]

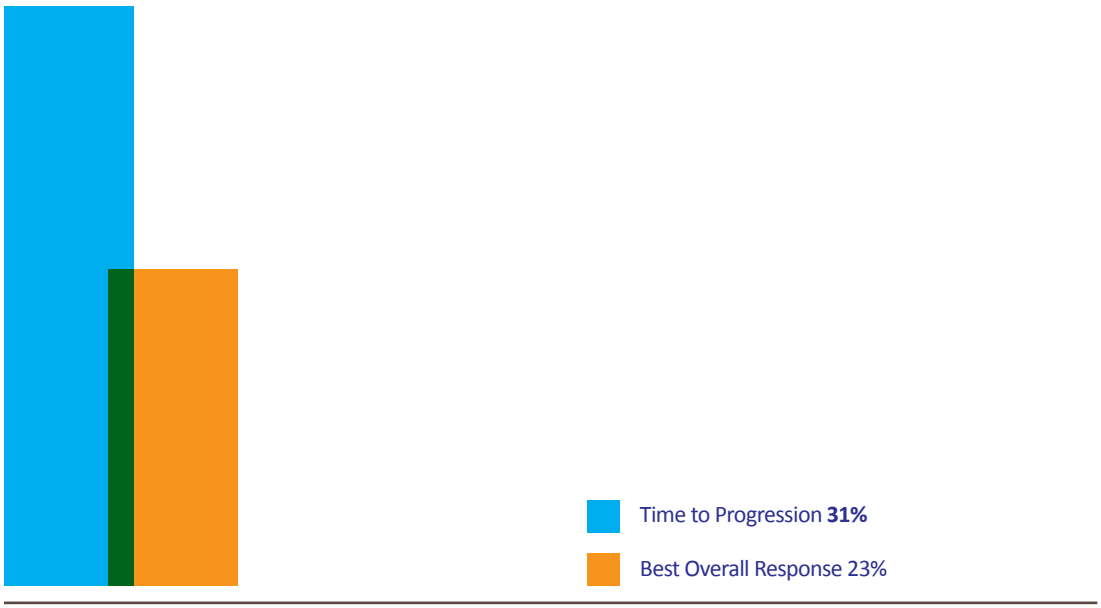
Inter-reader discordance

Data compiled from 40 oncology clinical trials (12 indications; 12,299 patients) examining discordance among independent central reviewers.

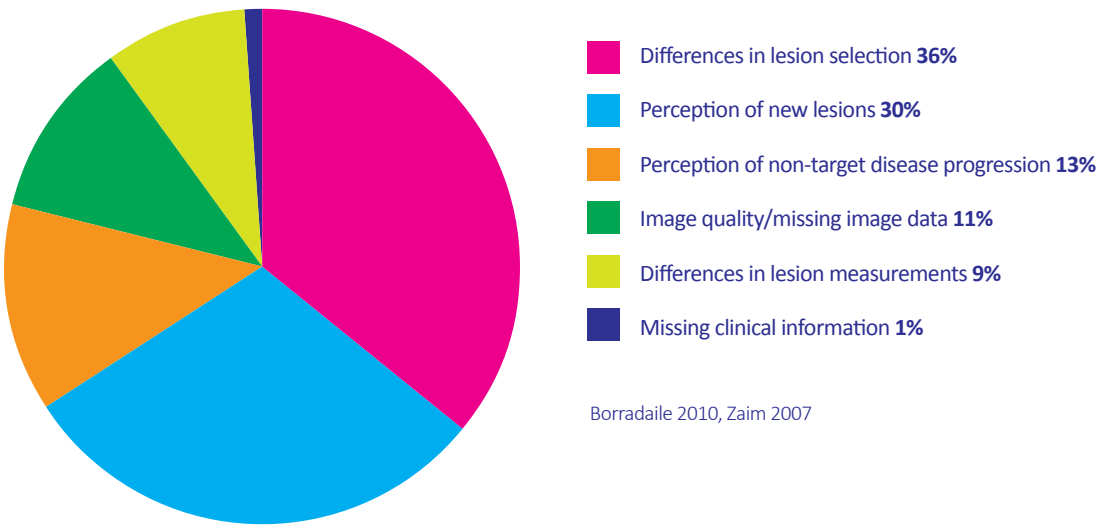
A Percent disagreement when determining best overall response and date of progression.

B Reasons for disagreement among readers.

A. Inter-reader Discordances



B. Reasons for Discordance



Borradaile 2010, Zaim 2007

Minimize Variability Through Standardization

The results obtained from imaging should never depend on the equipment manufacturer, imaging site, type of analysis software used, or the reader interpretation issues described above. The key to reducing all this variability is standardization.

Rigorous standardization provides a set of technical guidelines or rules that minimize variability, enhance reproducibility, and maintain objectivity.

Common standardization strategies include:

- Clarifying image acquisition techniques such as choice of modality and target lesions and calibrate the scanner regularly at each clinical site
- Properly training technologists and radiologists in image acquisition and interpretation methods
- Using computer-assisted interpretation software to minimize inter- and intra-reader variability
- Creating standard tumor measurement methods and parameters
- Assessing images in the order in which they were obtained using the same single reader for any patient
- Optimizing protocols and workflows to increase efficiency and prevent data loss [Sullivan 2013]

Therefore, in order to optimize your chances for submission success, it is important for all clinical investigators to be up to date with current FDA guidelines for the use of imaging in clinical trials.

FDA Guidelines for Use of Imaging Endpoints in Clinical Trials

In March 2015, the FDA released a new draft guidance containing updated process standards for the use of imaging endpoints in clinical trials. [http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm268555.pdf] These standards were created to help trial sponsors maximize the quality and reproducibility of imaging data that will be used as part of the drug approval process. The recommendations cover issues to consider when 1) designing a clinical trial and 2) creating trial-specific process standards.

Logistic and Technical Considerations for Clinical Trial Design

Blinded Interpretation. The FDA recommends that image readers should be blinded to a patient's treatment assignment in a randomized trial in order to minimize reader bias. In addition, readers should also be blinded to the date of image acquisition and previous imaging results if the potential exists for this information to influence a reader's assessment.

Timing

The frequency of imaging should be the same in all arms of a study if the primary endpoint is a time point-based assessment. When making time-to-event measurements, imaging should be performed at baseline and at frequencies sufficient for the disease under investigation. Timing must also be taken into account when considering the turnaround time for image analysis in that it must be appropriate for the trial design. In some trials, immediate interpretation may be necessary if the results will guide subsequent trial decisions or patient care, while batch analysis may be sufficient for other trial designs.

Image Interpretation Quality Control.

Monitoring and controlling the quality of image interpretation is critical for creating the reliable data needed for drug approval. The FDA recommends reader training and performance evaluations to maximize interpretation quality, particularly for highly specialized imaging methods or those prone to variability. When designing trials, it is also necessary to account for and overcome issues related to reader fatigue. If sufficient reader interpretation quality control measures cannot be undertaken at the clinical or sponsor site, the use of a centralized image interpretation service is recommended.

It is critical that all imaging-based clinical trials incorporate standardization practices that are compliant with federal guidelines in order to obtain subsequent FDA drug approval.

Factors to Consider When Standardizing.

Clinical trials vary widely in their design and objectives, and therefore, no single set of process standards apply to all trials. Instead, sponsors are encouraged to carefully consider the sources of variation specific to their unique trial design and imaging modality and how best to overcome them. Potential areas of standardization can include:

- Imaging Modality
  Consider the modality's availability, variations in performance across trial sites, and the potential for necessary equipment upgrades or failures
- Imaging Acquisition
  Consider the need for calibration standards or the use of phantoms to minimize variability and make technical parameters (imaging timing, patient positioning, use of contrast agents) consistent across sites
- Reader Interpretation
  Consider qualifications and training standards for technologists and radiologists, particularly for trial-specific quantification methods; consider ways to reduce variability, such as computer-assisted interpretation tools.
- Archiving
  Consider how and where data will be archived for subsequent monitoring and auditing

Creation of Trial-Specific Imaging Process Standards

Existing medical practice imaging process standards are sufficient for use in clinical trials if the inherent variability associated with a modality is low and if interpretation of the resulting data will be readily apparent. However, oncology trials often rely on complex quantification methods, and in such cases, the FDA recommends that trial sponsors create process standards specific to their trial. When a trial requires trial-specific standards that are lengthy or extensive, it's a good idea to develop an imaging charter, which is a document that provides detailed information on the methodology. FDA recommendations on imaging charters can be broken down into three key areas: image acquisition, image interpretation, and data transfer/archiving.

Image Acquisition: Detect and Track Lesions Consistently and Reliably

- Equipment

  All vendor-specific equipment and platforms should be FDA approved or cleared. This includes all software, which must also be FDA compliant. Trial-specific standards should also clearly outline how software upgrades and equipment substitutions will be handled, and specify technical standards such as slice thickness and contrast injection time.

- Human Parameters

  The role, qualifications, and training standards for imaging technicians should be identified and detailed. This includes the technician's role in the operation of the equipment and the initial assessment of image quality, as well as any necessary technologist training. Procedures to remove any site-to-site variation in how patients are imaged should be in place. This includes how a patient is prepared and positioned for imaging and any given comfort measures.

- Image Quality Controls

  How will the clinical trial site be qualified? Address how the equipment performance will be verified, the site will be inspected, and whether phantoms will be used. It is the site's responsibility to regularly check scanners for optimal performance and quality. The imaging schedule should also be outlined and followed to ensure consistent frequency (e.g., once per month) and timing (e.g., early morning). Finally, contrast agents, preparative drugs, and radiopharmaceutical agents should be used consistently from site to site, and their dose, method of administration, and potential risk to patients should be identified.

Image Interpretation: Increase Accuracy, Objectivity, and Reliability

- ### Computer-Assisted Interpretation Tools

  Computer-assisted interpretation tools can greatly reduce variability by overcoming many issues associated with manual interpretation, such as reader bias, reader fatigue, and inter-reader variability. Most importantly, all computer-assisted interpretation tools must be FDA approved. These software programs should be available to all readers within a trial and subjected to periodic quality control checks.

- ### Reader Qualifications

  The education and qualifications necessary to be selected as a reader should be clearly outlined. In addition, readers should undergo sufficient training in all computer programs, analysis tools, endpoint criteria (e.g., RECIST), and measurement parameters related to the trial. If necessary, define reader retraining and proficiency testing that will be conducted throughout the trial duration, and address the issue of reader fatigue by determining the maximum number of images that a reader will be allowed to interpret in a single day.

- ### Analysis

  Effective analysis begins with establishing selection criteria to determine which images will be analyzed and how images will be excluded. Who will select images for interpretation? What makes an image uninterpretable, and how is missing data handled? The turnaround time for image analysis (i.e., how soon after acquisition is the image evaluated by the reader) is also an important consideration and should reflect the specific trial endpoints. Throughout analysis, the FDA recommends the use of data locks to prevent readers from changing their final assessment. Lastly, trial-specific standards should explain how reader interpretational drift will be prevented if images within a set are obtained and assessed over time.

Data Transfer and Archiving: Enable Efficient FDA Review and Verification

The goal of any clinical trial is to generate reliable data that can be used to assess therapeutic efficacy and support a drug application. However, proper transfer and archiving of data can be instrumental in enabling efficient FDA review. A secure documentation and tracking system is necessary for all stages of image acquisition, digitization, and interpretation, and particularly for data transfer from the trial site to the sponsor or central interpretation facility. Data should be archived in a way that creates a sufficient audit trail for FDA verification.

References

Borradaile K, Ford R, O'Neal M et al. Applied Clinical Trials Online. Nov 2010

Cheson BD, Fisher RI, Barrington SF et al. (2014) J Clin Oncol. 32, 3059-3068.

Dinkel J, Khalilzadeh O, Hintze C et al. (2013) Lung Cancer 82, 76-82.

Eisenhauer EA, Therasse P, Bogaerts J et al. (2009) Eur J Cancer. 45, 228-247.

Goh V and Glynne-Jones R. (2014) Br J Radiol. 87, 20130811

Heertum RL, Scarimbolo R, Ford R, Berdougo E, and O'Neal M. (2015) Drug Design Devel Ther. 9, 5215-5223.

http://csdd.tufts.edu/news/complete_story/pr_tufts_csdd_2014_cost_study

http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm268555.pdf

Jain RK, Lee JJ, Ng C et al. (2012) J Clin Oncol. 30, 2684-2690.

Korn RL and Crowley JJ. (2013) Clin Cancer Res. 19, 2607-2612.

Krajewski KM, Nishino M, Franchetti Y et al. (2014) Cancer 120, 711-721.

Lui Y, deSouza NM, Shankar LK et al. (2015) Lancet Oncol. 16, e622-628.

Morgan B. (2011) Nat Rev Clin Oncol. 8, 517-527.

Muenzel D, Engels HP, Bruegel M, Kehl V, Rummeny EJ, and Metz S. (2012) Radiol Oncol. 46, 8-18.

Oxnard GR, Zhao B, Sima CS et al. (2011) J Clin Oncol. 29, 3114-3119.

Park DI, Kim SY, Kim JO et al. (2015) Tuberc Respir Dis. 78, 315-320.

Sharma MR, Maitland ML, and Ratain MJ. (2012) Cancer Res. 72, 5145-5149.

Sullivan DC, Schwartz LH, and Zhao B. (2013) Clin Cancer Res. 19, 2621-2628.

Villaruz LC and Socinski MA (2013) Clin Cancer Res. 19, 2629-2636.

Yankeelov TE, Mankoff DA, Schwartz LH et al. (2016) Clin Cancer Res. 22, 284-290.

Yoon SH, Kim KW, Goo JM, Kim DW, and Hahn S. (2016) Eur J Cancer. 53, 5-15.

Zaim S, Ounadjela S, Bouzegaou N et al. (2007) J Clin Oncol. ASCO Annual Meeting Proceedings 25 (June 20 Supplement), 6602.

Zhao B, Tan Y, Bell DJ et al. (2013) Eur J Radiol. 82, 959-968.

Design: www.thinkerdoer.co.uk

median

The Imaging Phenomics™ Company

About Median Technologies

Median Technologies provides innovative imaging solutions and services to advance healthcare for everyone. We leverage the power of Imaging Phenomics™ to provide insights into novel therapies and treatment strategies. Our unique solutions, LMS for lesion management and our imaging phenotyping system iBiopsy™, together with our global team of experts, are advancing the development of new drugs and diagnostic tools to monitor disease and assess response to therapy.

Median Technologies supports biopharmaceutical sponsors and healthcare professionals around the world to bring new treatments to patients in need more precisely, quicker and with an eye on reducing overall care costs. This is how we are helping to create a healthier world.

ALMDT
LISTED
ALTERNEXT

Median Technologies

inforequest@mediantechnologies.com
mediantechnologies.com

Version 1.0