

KEYNOTE LECTURE

Wednesday 18 October 2006, 09:00–09:30

Imaging response assessment in oncology

S D Curran, A U Muellner and L H Schwartz

Department of Radiology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA

Corresponding address: Lawrence H Schwartz, MD, Department of Radiology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10021, USA. E-mail: schwart1@mskcc.org

Abstract

The role of imaging in the clinical setting as well as in the drug development process is expanding rapidly. Imaging technology now exists that is capable of detecting tumor response within hours. In parallel with this advance, a new array of more targeted and specific therapies are being developed. This paradigm shift in turn demands a more sophisticated way of quantifying response. There is a need to update and modify the current response evaluation criteria in solid tumors (RECIST), which rely solely on anatomic size measurement of tumors. In addition, response assessment guidelines will need to be increasingly disease-specific. Response assessment by imaging is now intimately involved with all stages of the drug development process, from exploratory drug discovery through clinical trials, as well as in clinical use. Imaging biomarkers and surrogate endpoints have the potential to speed drug approval significantly. The major funding institutions and the pharmaceutical industry are working more and more with researchers to help maintain progress in this multidisciplinary area involving oncologists, radiologists, molecular imaging specialists, medical physicists, and computer scientists.

Keywords: RECIST; biomarkers; drug discovery.

Introduction

Advances in cross-sectional imaging such as multislice computed tomography (CT) and three-dimensional (3D) techniques have made radiologic measurements more reproducible and accurate. Moreover, functional imaging modalities such as perfusion and diffusion-weighted magnetic resonance imaging (MRI) and positron emission tomography (PET) have the potential to provide new biomarkers reflecting key aspects of tumor biology and growth. The ability to marry increasingly useful imaging information with new clinical endpoints has become important in cancer therapeutic trials assessing a new generation of targeted molecules for cancer treatment. In both human and animal models, response to these new agents is being recorded with imaging at shorter time intervals after treatment. This paves the way for much more rapid drug evaluation and, potentially, clinical decision making. With cancer patients now living longer and receiving more complex neoadjuvant therapies, there is a growing need to develop imaging methods to act as surrogate endpoints to replace the more traditional

endpoints of morbidity or mortality. One of the many challenges ahead will be to correlate early imaging response with survival data.

The capacity and need for imaging to play a greater role in drug development has been recognized by major funding organizations, such as the National Cancer Institute (NCI) and the Food and Drug Administration (FDA) in the United States. With their help, at least ten major cancer centers across the US are building teams to facilitate the role of imaging in therapeutic response assessment of cancers. Image processing and analysis within clinical trials is becoming more and more automated. Computer software is under development to generate detailed quantitative time dependent data on lesions as well as computer aided detection (CAD) of total tumor burden.

WHO and RECIST

Traditional methods of measuring tumor size and assessing response are still in use in clinical oncology

and drug development. With regard to assessing change within clinical oncology trials, four categories of tumor response still exist (Table 1). These are the World Health Organization (WHO) guidelines (two-dimensional) and the more recently agreed upon RECIST guidelines (unidimensional)^[1]. The RECIST guidelines have major limitations^[2].

Table 1 Measuring response in oncology trials as per RECIST criteria

CR (complete response)	Tumors completely disappear
PR (partial response)	Tumors shrink >30%
SD (stable disease)	Tumors stable
PD (progressive disease)	Tumors grow >20%

Several studies have shown that in some cases, response assessment by WHO and RECIST differs significantly, suggesting that better criteria are needed^[3-5]. Several studies have found concordance between WHO and RECIST in assessing response but discordance between them in calculating the time to progression, with RECIST criteria showing a lower rate of disease progression than WHO^[3,6]. With regard to the commonly used slice thicknesses for CT (i.e. 7.5 mm, 5 mm and 3.75 mm), it seems that, for now at least, 5-mm and 7-mm sections are equivalent for the measurement of tumors in two dimensions, but 3.75-mm and 5-mm sections are superior to 7.5-mm sections for measurement of tumor volume^[7].

Difficulties are often encountered in measuring both the primary tumor (e.g. breast or prostate tumors) and metastases. For example, response of bone metastases is not generally reflected by size changes. Osseous metastases are by definition non-target lesions and no definition of what constitutes their response has been established to date. Assessment of response by bone scintigraphy may even be misleading^[8]. The question of whether soft tissue response and osseous metastasis response should be considered separately or together when assessing response remains unanswered.

The use of size measurement to assess response in lymph nodes also has potential pitfalls. First, the smaller the nodes, the more likely there is to be measurement error^[9]. Secondly, lymph nodes do not disappear even with a complete response, and it is well known that normal-sized lymph nodes can contain viable tumor, whereas enlarged nodes may be free of tumor. In addition, RECIST does not take into account 3D changes in tumor size and shape, which may not necessarily correlate with one- or two-dimensional measurements^[5,10]. The assessment of malignant pleural mesothelioma (MPM) is a difficult problem. MPM is an especially difficult to measure tumor. At present, the currently accepted modified RECIST criteria are used involving a series of unidimensional measurements^[11]. Further work is necessary to achieve the best method of quantifying reproducible changes in this tumor^[11,12]. It may be that volume measurements of tumors is a better predictor of response than

traditional two-dimensional measurements. Automated lung nodule volume measurement on CT shows less inter-observer variability over manual measurement^[13]. The technology is available to analyze CT and MR images in three dimensions. One prospective study of esophago-gastric cancer showed tumor volumetry based on CT correlates with histologic tumor regression whereas two-dimensional measurements did not^[10]. Further studies are needed to validate the robustness of volume measurements before it is accepted as a new biomarker.

Drug discovery

The term molecular imaging encompasses not only clinical nuclear medicine but a broad area of pre-clinical research involving the discovery and development of molecular probes to better define disease and help quantify treatment targets and responses in both animals and humans. The ability of molecular imaging to accurately depict targets in vivo has revolutionized pre-clinical testing of novel anti-tumor agents.

Advances in proteomics and genomics have opened the door for the development of many compounds that have the potential to be effective drugs in oncology^[14]. Imaging probes have been developed already to image all the main processes driving neoplasia such as proliferation, angiogenesis, apoptosis, and hypoxia. The role of imaging in clinical oncology has vast potential and is only at the beginning in terms of its ability to influence treatment regimens and predict response^[15] (Fig. 1).

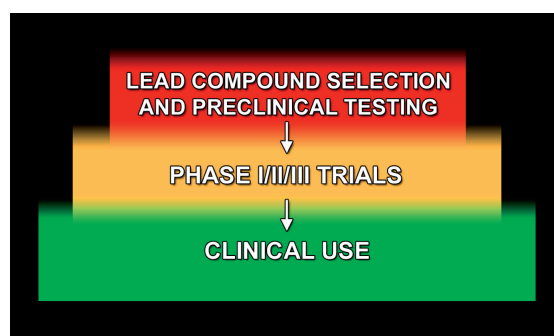


Figure 1 Stages of the drug discovery process—all of which benefit from imaging biomarkers and imaging surrogate endpoints.

Biomarkers are anatomic, physiologic, biochemical, or molecular parameters associated with the presence and severity of specific disease states. They are imaging ‘signals’ or ‘findings’ that reflect and thus act as a measure of a pathologic process (e.g. tumor growth, angiogenesis). Biomarkers are detectable and measurable by a variety of methods including physical examination, laboratory assays and medical imaging. The simplest anatomic example is tumor size, but there is an array of molecular targets which can be imaged to help predict early response to treatment. Other examples of

biomarkers include blood pressure, heart rhythm, blood glucose and prostate-specific antigen (PSA). Imaging markers are currently under intense investigation. The ultimate aim is to identify imaging findings that can provide a true clinical benefit for the patient. There are, however, examples of failures or limitations of imaging biomarkers. Reasons for the failure of imaging markers include a lack of true biological relevance and a lack of accuracy and reproducibility. A potential biomarker should be statistically validated with rigorous criteria, such as Prentice's criteria^[16]. The FDA in conjunction with the NCI announced in February 2006 a memorandum of agreement for the Oncology Biomarker Qualification Initiative (OBQI)^[17].

Being able to show an early response is useful in many tumors, as early response has been shown to correlate with better prognosis^[18]. Clinical trials are using more and more imaging response data. For biotechnology and pharmaceutical companies, the potential savings involved are vast, since imaging endpoints have the potential to help get a drug approved more rapidly than conventional endpoints (e.g. mortality and morbidity). Examples of exciting new targeted molecules that have been successfully used in oncology are trastuzumab, the HER-2/*neu* antibody for breast cancer and the kinase inhibitor Imatinib for chronic myeloid leukemia and gastrointestinal stromal tumors (GIST). It is well known that a responding GIST may actually increase in size and that a progressing tumor may show an enhancing nodule on contrast-enhanced CT, criteria which are not included in the RECIST guidelines^[19]. The actions of the new targeted agents (e.g. inhibition of tyrosine or protein kinases), even when effective, do not necessarily trigger changes in tumor size. This presents a difficulty in assessing these new agents, as no established reference standard exists by which to judge their effectiveness^[20,21].

Computer aided detection (CAD) of lesions in the lung is a good example of advanced image processing that can enable superior and more efficient analysis of images for clinical trials^[22]. The automation of lesion detection is an area that has received a lot of attention recently. Work is in progress to enable total tumor burden assessment by CAD. Automating imaging assessment and analysis of lesions is essential to obtain objective as well as more accurate and reproducible results^[22].

FDG-PET and molecular imaging

The last few years have seen the incorporation of PET, and subsequently, PET-CT into routine practice at many centers. The use of FDG-PET as a functional biomarker in clinical care needs to be standardized if it is to provide a robust surrogate endpoint for clinical trials. Recent NCI guidelines spell out specific protocols for patient preparation and all aspects of image acquisition for any trials involving this modality^[20]. False positive FDG

uptake after treatment with radiation especially needs to be considered. The hope is that FDG-PET along with other emerging functional imaging probes will serve as biomarkers, resulting in clinical benefits for patients. Compared to conventional imaging, functional imaging can allow earlier detection of response—sometimes after just one treatment—and is being used in some studies as a marker for whether a patient is responding to a particular therapy. [¹⁸F]FDG uptake has been shown to predict prognosis in lung, esophageal and thyroid cancers^[23–25]. The International Working Group held a recent workshop to update treatment response criteria for lymphoma to include FDG-PET, and these guidelines are due out soon^[26]. Correlation between SUV uptake reduction and response just one week after treatment with chemotherapy has been demonstrated^[27–30]. One study showed a correlation between a greater than 35% [¹⁸F]FDG uptake reduction and progression-free survival^[27]. Larger prospective studies are required to determine if FDG-PET positivity correlates with histologic grade in each cancer type^[31–34]. Each cancer type needs to be studied in detail in well-designed research protocols. Looking ahead, it is likely that evidence will emerge to show an increasing role for FDG-PET to assess response to treatment. Further studies will help stratify patients into separate subgroups with different treatment plans avoiding unnecessary treatments and toxicities and enable adjustment of treatment regimens.

Functional MRI (diffusion-weighted and perfusion)

Perfusion MRI holds great promise because contrast uptake correlates with the neoangiogenesis of certain tumors^[35]. The rationale is that decreased tumor perfusion correlates with inhibition of angiogenesis and VEGF pathways. This is the topic of intense research. Guidelines for the use of MRI in clinical trials to assess new therapies inhibiting angiogenesis pathways have been published^[36].

Some of the functional properties of tumors can be quantified by dynamic contrast-enhanced (DCE) MRI. Not only the intensity or density of contrast enhancement but the contrast kinetics is quantifiable on a pixel-by-pixel basis. This enables more robust data on changes in tumor characteristics and aggressiveness after treatment^[37]. Functional imaging is non-invasive and has the potential to avoid biopsy which can be misleading and may have sampling error. Improvements in the temporal and spatial resolution of CT and MRI have increased the accuracy of these modalities, but the addition of perfusion and diffusion-weighted techniques provides information about tumor activity that has the potential to serve as a biomarker for treatment response in tumors^[35].

The generation of an apparent diffusion coefficient (ADC) map is done on a voxel-by-voxel basis, making

this technique quantifiable and objective. A correlation between contrast uptake by tumors and angiogenesis has been shown^[38].

Future challenges

Imaging needs to keep up with the ‘microscopic’ revolution that represents the future of cancer therapeutics and diagnostics^[39]. With the advent of molecular medicine, the ultimate goal of research in oncology is to be able to tailor treatments to both the specific type of cancer and the individual patient. To achieve this, it is crucial that new imaging biomarkers be developed and validated to keep pace with the testing of novel therapeutic agents in clinical trials. The decision of Medicare and Medicaid to provide reimbursement of PET studies for any prospective cancer trial in the US should encourage more validation studies (<http://www.cancerpetregistry.org>). Validation of functional biomarkers including but not limited to FDG-PET and DCE-MRI is essential to ensure imaging continues to improve response assessment. For existing biomarkers, improvement of image robustness and image processing and analysis will help speed the development of new therapies for cancer.

Conclusion

The current guidelines for assessment of tumor response as outlined by RECIST address anatomical tumor size. These may not be sufficient to encompass the assessment of response to newer targeted chemotherapeutic agents, which are cytostatic and do not necessarily cause tumor shrinkage when they are effective. FDG-PET as well as perfusion and diffusion weighted imaging all show great promise as biomarkers and surrogate endpoints for clinical trials. Validation studies are required to ensure these techniques are useful in monitoring tumor response and translate into real clinical benefits. The use of imaging biomarkers in all stages of the drug discovery process has increased efficiency by helping to speed up the development of more effective therapies.

Just as the combination of specific targeted therapies such as Trastuzumab in combination with standard chemotherapy has been shown to be more effective than standard chemotherapy alone, a combination of both anatomic and functional imaging biomarkers will more accurately reflect the true response of tumors.

Acknowledgment

We are grateful to Carolina Montalvo for help with the artwork.

References

- [1] Therasse P, Arbuck SG, Eisenhauer EA *et al*. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000; 92: 205–16.
- [2] Husband JE, Schwartz LH, Spencer J *et al*. Evaluation of the response to treatment of solid tumours—a consensus statement of the International Cancer Imaging Society. *Br J Cancer* 2004; 90: 2256–60.
- [3] Schwartz LH, Mazumdar M, Wang L *et al*. Response assessment classification in patients with advanced renal cell carcinoma treated on clinical trials. *Cancer* 2003; 98: 1611–19.
- [4] Scher HI, Morris MJ, Kelly WK, Schwartz LH, Heller G. Prostate cancer clinical trial end points: RECISTing a step backwards. *Clin Cancer Res* 2005; 11: 5223–32.
- [5] Schwartz LH, Colville JA, Ginsberg MS *et al*. Measuring tumor response and shape change on CT: esophageal cancer as a paradigm. *Ann Oncol* 2006; 17: 1018–23.
- [6] Therasse P, Eisenhauer EA, Verweij J. RECIST revisited: a review of validation studies on tumour assessment. *Eur J Cancer* 2006; 42: 1031–9.
- [7] Zhao B, Schwartz LH, Moskowitz CS *et al*. Pulmonary metastases: effect of CT section thickness on measurement—initial experience. *Radiology* 2005; 234: 934–9.
- [8] Smith PH, Debruyne F, Armitage TG *et al*. The value of the bone scan. *Prog Clin Biol Res* 1990; 357: 53–7.
- [9] 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: 2574–82.
- [10] Beer AJ, Wieder HA, Lordick F. Adenocarcinomas of esophagogastric junction: multi-detector row CT to evaluate early response to neoadjuvant chemotherapy. *Radiology* 2006; 239: 472–80.
- [11] Byrne MJ, Nowak AK. Modified RECIST criteria for assessment of response in malignant pleural mesothelioma. *Ann Oncol* 2004; 15: 257–60.
- [12] Oxnard GR, Armato 3rd SG, Kindler HL. Modeling of mesothelioma growth demonstrates weakness of current response criteria. *Lung Cancer* 2006; 52: 141–8.
- [13] Marten K, Auer F, Schmidt S, Kohl G, Rummeny EJ, Engelke C. Inadequacy of manual measurements compared to automated CT volumetry in assessment of treatment response of pulmonary metastases using RECIST criteria. *Eur Radiol* 2006; 16: 781–90.
- [14] Kelloff GJ, Krohn KA, Larson SM *et al*. The progress and promise of molecular imaging probes in oncologic drug development. *Clin Cancer Res* 2005; 11: 7967–85.
- [15] Weissleder R. Molecular imaging in cancer. *Science* 2006; 312: 1168–71.
- [16] Alonso A, Molenberghs G, Geys H, Buyse M, Vangeneugden T. A unifying approach for surrogate marker validation based on Prentice’s criteria. *Stat Med* 2006; 25: 205–21.
- [17] Woodcock J, Barker AD, Straube B. Memorandum of Understanding. Oncology Biomarker Qualification

- Initiative (OBQI). 2006. <http://www.fda.gov/oc/mous/domestic/FDA-NCI-CMS.html>.
- [18] Carde P, Koscielny S, Franklin J *et al*. Early response to chemotherapy: a surrogate for final outcome of Hodgkin's disease patients that should influence initial treatment length and intensity? *Ann Oncol* 2002; 13 (Suppl 1): 86–91.
- [19] Choi H. Critical issues in response evaluation on computed tomography: lessons from the gastrointestinal stromal tumor model. *Curr Oncol Rep* 2005; 7: 307–11.
- [20] Larson SM, Schwartz LH. ¹⁸F-FDG PET as a candidate for qualified biomarker: functional assessment of treatment response in oncology. *J Nucl Med* 2006; 47: 901–3.
- [21] Michaelis LC, Ratain MJ. Measuring response in a post-RECIST world: from black and white to shades of grey. *Nat Rev Cancer* 2006; 6: 409–14.
- [22] Zhao B, Gamsu G, Ginsberg MS *et al*. Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. *J Appl Clin Med Phys* 2003; 4: 248–60.
- [23] Downey RJ, Akhurst T, Gonen M *et al*. Preoperative F-18 fluorodeoxyglucose-positron emission tomography maximal standardized uptake value predicts survival after lung cancer resection. *J Clin Oncol* 2004; 22: 3255–60.
- [24] Rizk N, Downey RJ, Akhurst T *et al*. Preoperative [¹⁸F]-fluorodeoxyglucose positron emission tomography standardized uptake values predict survival after esophageal adenocarcinoma resection. *Ann Thorac Surg* 2006; 81: 1076–81.
- [25] Robbins RJ, Wan Q, Grewal RK *et al*. Real-time prognosis for metastatic thyroid carcinoma based on 2-[¹⁸F]fluoro-2-deoxy-D-glucose-positron emission tomography scanning. *J Clin Endocrinol Metab* 2006; 91: 498–505.
- [26] Guermazi A, Juweid ME. Commentary: PET poised to alter the current paradigm for response assessment of non-Hodgkin's lymphoma. *Br J Radiol* 2006; 79: 365–7.
- [27] Weber WA, Ott K, Becker K *et al*. Prediction of response to preoperative chemotherapy in adenocarcinomas of the esophagogastric junction by metabolic imaging. *J Clin Oncol* 2001; 19: 3058–65.
- [28] Kostakoglu L, Leonard JP, Kuji I *et al*. Comparison of fluorine-18 fluorodeoxyglucose positron emission tomography and Ga-67 scintigraphy in evaluation of lymphoma. *Cancer* 2002; 94: 879–88.
- [29] Downey RJ, Akhurst T, Ilson D *et al*. Whole body ¹⁸FDG-PET and the response of esophageal cancer to induction therapy: results of a prospective trial. *J Clin Oncol* 2003; 21: 428–32.
- [30] Ott K, Fink U, Becker K *et al*. Prediction of response to preoperative chemotherapy in gastric carcinoma by metabolic imaging: results of a prospective trial. *J Clin Oncol* 2003; 21: 4604–10.
- [31] Folpe AL, Lyles RH, Sprouse JT *et al*. (F-18) fluorodeoxyglucose positron emission tomography as a predictor of pathologic grade and other prognostic variables in bone and soft tissue sarcoma. *Clin Cancer Res* 2000; 6: 1279–87.
- [32] Kurokawa T, Yoshida Y, Kawahara K *et al*. Expression of GLUT-1 glucose transfer, cellular proliferation activity and grade of tumor correlate with [F-18]-fluorodeoxyglucose uptake by positron emission tomography in epithelial tumors of the ovary. *Int J Cancer* 2004; 109: 926–32.
- [33] Cho SM, Park YG, Lee JM *et al*. ¹⁸F-fluorodeoxyglucose positron emission tomography in patients with recurrent ovarian cancer: in comparison with vascularity, Ki-67, p53, and histologic grade. *Eur Radiol* 2006; Jun 20 [Epub ahead of print].
- [34] Guo J, Higashi K, Ueda Y *et al*. Microvessel density: correlation with ¹⁸F-FDG uptake and prognostic impact in lung adenocarcinomas. *J Nucl Med* 2006; 47: 419–25.
- [35] Provenzale JM, Mukundan S, Barboriak DP. Diffusion-weighted and perfusion MR imaging for brain tumor characterization and assessment of treatment response. *Radiology* 2006; 239: 632–49.
- [36] Leach MO, Brindle KM, Evelhoch JL *et al*. The assessment of antiangiogenic and antivascular therapies in early-stage clinical trials using magnetic resonance imaging: issues and recommendations. *Br J Cancer* 2005; 92: 1599–610.
- [37] d'Arcy JA, Collins DJ, Padhani AR *et al*. Informatics in Radiology (infoRAD): Magnetic Resonance Imaging Workbench: analysis and visualization of dynamic contrast-enhanced MR imaging data. *Radiographics* 26: 621–32.
- [38] Morgan B, Thomas AL, Dreys J *et al*. Dynamic contrast-enhanced magnetic resonance imaging as a biomarker for the pharmacological response of PTK787/ZK 222584, an inhibitor of the vascular endothelial growth factor receptor tyrosine kinases, in patients with advanced colorectal cancer and liver metastases: results from two phase I studies. *J Clin Oncol* 2003; 21: 3955–64.
- [39] Cristofanilli M. The microscopic revolution in breast carcinoma. *Cancer* 2005; 103: 877–80.