Published in final edited form as:

Semin Oncol. 2011 February ; 38(1): 16–25. doi:10.1053/j.seminoncol.2010.11.007.

The Role of Magnetic Resonance Imaging Biomarkers in Clinical Trials of Treatment Response in Cancer

Thomas E. Yankeelov, Ph.D.^{1,2,3,4,5,7}, Lori A. Arlinghaus, Ph.D.^{1,2}, Xia Li, Ph.D.^{1,2}, and John C. Gore, Ph.D.^{1,2,3,4,6,7}

¹Institute of Imaging Science, Vanderbilt University, Nashville, Tennessee 37232-2310

²Department of Radiology and Radiological Sciences, Vanderbilt University, Nashville, Tennessee 37232-2310

³Department of Biomedical Engineering, Vanderbilt University, Nashville, Tennessee 37232-2310

⁴Department of Physics and Astronomy, Vanderbilt University, Nashville, Tennessee 37232-2310

⁵Department of Cancer Biology, Vanderbilt University, Nashville, Tennessee 37232-2310

⁶Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee 37232-2310

⁷Program in Chemical and Physical Biology, Vanderbilt University, Nashville, Tennessee 37232-2310

Abstract

Current standard-of-care radiological methods for assessing the response of solid tumors to treatment are based on measuring changes in lesion size in a single dimension using high-resolution x-ray computed tomography or magnetic resonance imaging (MRI). Even if size measurements are adapted to record true volume changes more accurately, the effects of therapeutic drugs on tumor size may not occur for several cycles of treatment. Furthermore, current and future generations of anti-cancer drugs will be designed to affect highly specific cancer characteristics, and their effects may not be immediately cytotoxic. More sensitive and specific measures are required that can report on tumor status and treatment response early in the course of therapy. Several MRI techniques have matured to the point where they can offer quantitative information on tissue status and greater insight into specific biophysical and physiological characteristics of tumors. Here we review and provide illustrative examples of two MRI methods that have already been incorporated into clinical trials of treatment response in solid tumors: diffusion imaging and dynamic contrast enhanced MRI. We also discuss the limitations and future research directions required for these techniques to gain greater acceptance and to have their maximum impact.

© 2010 Elsevier Inc. All rights reserved.

Please address correspondence to: Thomas Yankeelov, Ph.D. 1161 21st Avenue South AA 1105 MCN Nashville, Tennessee 37232-2310 615-322-8354 (phone) 615-322-0734 (fax) thomas.yankeelov@vanderbilt.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

The effective and reliable assessment of the response of tumors to treatment is critical for the clinical management of cancer. During a course of treatment, biopsies may be obtained only infrequently (due to their invasive nature) to assess treatment response. Additionally, the reliability of their spatial sampling may be poor and the results misleading, so this method is not acceptable for routine patient care. Furthermore, certain disease locations (e.g., brain and lung) preclude the ability to acquire repeat biopsies. The net result of these limitations is that clinical judgments of the effectiveness of treatments are subjective and prone to error. Repeatable, quantitative, non-invasive imaging methods that can reliably assess tumor response would greatly improve cancer care. This is especially true in clinical trials employing investigational drugs or drug combinations with unknown or unproven efficacies, or when such treatments are effective for specific tumor phenotypes only.

The number of imaging techniques that can potentially be applied for the quantitative characterization of cancer has increased significantly in recent years.¹ They include magnetic resonance imaging (MRI), optical imaging, computed tomography (CT), positron emission tomography (PET), single photon emission computed tomography (SPECT), and ultrasound. Such techniques have increasingly tried to assess specific cellular and molecular features of tumors as biomarkers of response. Relevant techniques that have been developed report on, for example, tumor cellularity, blood vessel perfusion and permeability, metabolism and hypoxia, as well as other specific cellular and molecular processes. It is a reasonable hypothesis that using appropriate noninvasive imaging methods to more comprehensively assess tumor status will offer increased sensitivity and specificity when assessing treatment response. Furthermore, as many current anti-cancer drugs are designed to alter specific tumor characteristics, imaging methods designed to report on those same phenomena promise to offer improved methods of planning treatment as well as assessing the response of tumors to treatment. To date, however, despite substantial progress in imaging science and the incorporation of such advances into pre-clinical research, very few of these techniques are regularly used clinically or in clinical trials.

The current radiological methods of monitoring treatment response in solid tumors rely on frank changes in tumor morphology as measured by x-ray CT or MRI. While two-dimensional and three-dimensional methods for assessing changes in tumor volumes in response to treatment have been considered, the currently accepted method is based on measuring changes in a single dimension only of a tumor. Formal guidelines were provided in a seminal 2000 publication which introduced RECIST, the Response Evaluation Criteria in Solid Tumors.² RECIST criteria offer a simplified, convenient, practical but highly imperfect method for extracting salient features from anatomical images. Following a pre-treatment scan, “target lesions are identified” and their longest dimension is measured. The “baseline sum longest dimension” is then defined by summing the longest dimension of all the target lesions. The degree to which the sum of longest tumor dimensions changes during treatment determines the response category. While many clinical trials have used these criteria for cancers in many disease sites, it is well recognized that this approach needs to be significantly improved to address a number of practical, technical, and scientific issues.

Clearly, improved methods are needed for characterizing the underlying pathophysiological changes induced by specifically targeted agents. Such methods may be more likely to offer earlier and more specific indications of response to treatment than changes in tumor dimensions. There are several MRI techniques that have matured to the point where they can offer quantitative and specific information on tumor status. Here we review two techniques that provide complementary information on tissue biophysical properties that have seen substantial application in cancer clinical trials: diffusion weighted MRI (DWI), and dynamic

contrast-enhanced MRI (DCE-MRI). Although these do not directly indicate specific molecular changes or events, they have proven reliable indicators of important sequelae to the direct consequences of targeted therapies. We survey a selection of investigations from the literature that have used these techniques to assess the response of tumors to treatment within the confines of a clinical trial. Rather than provide an exhaustive review of the literature, our goal is to provide illustrative examples on how these techniques can be incorporated into clinical trials. We then conclude by pointing out a series of research directions appropriate for making these techniques more readily available for clinical trials and incorporation into tumor response guidelines like RECIST.

DEFINITION OF “QUANTITATIVE” MRI

A MR image is composed of a series of volume elements (voxels) arranged in a matrix. The relative signal intensities within each voxel determine the contrast within an image. In a typical MRI study the voxel intensities are determined by a number of factors including selected acquisition parameters (such as the repetition time, TR , and echo time, TE), and the tissue's inherent magnetic resonance characteristics (such as relaxation times and proton density). Consequently, the signal intensity that is displayed in conventional MR images presented to radiologists represents a complicated combination of these parameters and, importantly, has no physical units. The usual primary design consideration for such images is to maximize contrast between particular tissue types and structures in order to optimize detection and diagnosis. However, these images do not portray intrinsic properties of tissues or unambiguously allow particular properties of tissues to be assessed. For example, regions of high proton density (e.g., cerebrospinal fluid or edema) may appear brighter or darker than tissues with less water depending on the choice of pulse sequence, so even simple measures such as the amount of water in tissue cannot be made. Furthermore, signal intensities in typical images are difficult to compare across patients and across institutions, thereby complicating multi-site studies. For example, panels A through C in Fig 1 depict sagittal views of a breast acquired with different flip angle values. In each of these three images the contrast reflects variations in signal intensity that do not have physical units associated with them and are difficult to interpret in terms of individual physiological or physical parameters; rather they are a confluence of many parameters. Quantitative MRI connotes those techniques that provide measurements of intrinsic tissue properties and reflect particular physiological phenomena or biophysical properties and which are specified in appropriate units. By combining and analyzing the data from multiple images, and by assuming particular mathematical relationships between images (i.e., a model), such intrinsic properties can be computed. In the case of Fig 1, these three data sets can be used to develop a T_1 map, wherein every voxel in the image corresponds to the tissue's T_1 value and the units of this image can be expressed in absolute terms (for T_1 , the units are seconds).

DIFFUSION IMAGING

Like conventional MRI, DWI records signals from the mobile water molecules within tissues, but in DWI the contrast reflects how far water molecules can migrate from one place to another in a specified time interval. The random, thermally-induced motions of water molecules in a solution is referred to as Brownian motion, and the integrated effect of multiple microscopic translations and intermolecular collisions results in the gradual migration of molecules away from their origin, a process termed self-diffusion. Einstein (1905)³ showed that, in an unrestricted medium, the mean distance L traveled by a water molecule in a time Δ is given by $L = \sqrt{2D\Delta}$, where D is the self-diffusion coefficient, an intrinsic property of the material that is measured in cm^2/sec . For pure water at room temperature, D is approximately $2 \times 10^{-5} \text{ cm}^2/\text{sec}$; so L is about 10 microns when L is 50 milliseconds. In MRI, in the presence of applied gradients of the main magnetic field, this

spatial migration gives rise to small changes (decreases) in the acquired MRI signals, which thus provides the basis of the contrast depicted in DWI. By acquiring two or more images with different degrees of “diffusion weighting” (obtained by applying the diffusion sensitizing gradients with different amplitudes on successive image acquisitions), the value of the water self-diffusion coefficient can be calculated at each voxel, so a map of D is obtained. An example is presented in Fig 2. The experimentally measured rate of self-diffusion of water is found to be lower in organized tissues than in free solution because various structures restrict or hinder the free movement of water. For example, cell membranes and intracellular organelles may confine and redirect the paths followed by water molecules diffusing through biological tissue. Then the distance L travelled in time Δ is reduced, so D appears to be smaller. The resultant map is then of the *apparent diffusion coefficient* or ADC. Diffusion imaging methods indirectly provide information about the density of restrictive structures in the local cellular environment. The ADC thus provides information about tissue microstructure and composition, and is reduced from the intrinsic diffusion coefficient of free water to a degree that is dependent upon the density and spacing of these restricting structures.^{4,5}

Cancerous tissues often show significantly different ADC values from healthy tissues, providing motivation for the use of DWI techniques to study tumor proliferation and response to therapy.⁶ As tumor cells proliferate, the ADC of water within the tumor often decreases as cell density rises, whereas after treatment the ADC often increases, presumably because of decreases in cell density consequent to apoptosis and cell death, with concomitant disruption of cell membranes that allows water to move around more freely. Such changes occur quite early after treatments, and thus are suitable for assessing early responses. This increase in ADC following treatment can often be directly linked to a decrease in cellularity. It has recently been shown that exposure of tumors to both chemotherapy and radiotherapy consistently leads to measurable increases in water diffusion in cases of favorable treatment response and several examples are discussed below.

DYNAMIC CONTRAST ENHANCED MRI

In approximately 35% of all clinical MRI exams, a gadolinium-based contrast agent (CA) is injected into the patient to enhance differences between pathologic and healthy appearing tissues.⁷ The enhancement is due to the CA's ability to change a tissue's native T_1 and/or T_2 relaxation time. MRI techniques that are designed to record and analyze the changes in T_1 that occur during the passage of a CA through tissue are termed DCE-MRI. During a DCE-MRI session, T_1 -weighted images are acquired before, during, and after the injection of the CA into (usually) a peripheral vein of a patient. As the agent perfuses into the field-of-view, it changes the T_1 relaxation time to a degree determined by the concentration of the CA in that voxel. As the concentration of CA changes due to influx and efflux, the measured signal intensity of that voxel will change; first increasing as CA accumulates, and then decreasing as the CA concentration decreases. Thus, if multiple images are acquired during this process, a signal intensity time course will result for each voxel and this time course can be analyzed to report certain physiological parameters related to tissue status. If we let C_t and C_p denote the concentration of CA in the tissue space and the blood plasma, respectively, then the differential equation describing the transfer of CA from the blood space to the tissue space is:

$$\frac{d}{dt}C_t(t) = K^{trans}C_p(t) - (K^{trans}/v_e)C_t(t), \quad [1]$$

where K^{trans} is the CA transfer rate constant (in units of ml(blood)/ml(tissue)/(min)) that reports on blood vessel perfusion and permeability, and v_e is the extravascular, extracellular volume fraction. Solving Eq. [1] yields

$$C_t(T) = K^{trans} \int_0^T C_p(t) \exp\left(-\left(K^{trans}/v_e\right)(T-t)\right) dt. \quad [2]$$

Many investigators have noted that Eq. [2] neglects the contribution of the blood plasma to imaging voxel and have amended it as follows:

$$C_t(T) = K^{trans} \int_0^T C_p(t) \exp\left(-\left(K^{trans}/v_e\right)(T-t)\right) dt + v_p C_p(T), \quad [3]$$

where v_p is the plasma volume fraction.

In order to generate a map of K^{trans} , v_e , or v_p , estimates of C_t and C_p as a function of time are required. Then Eq. [2] or [3] is used to fit the C_t and C_p time courses by varying K^{trans} , v_e and v_p as free parameters; the value of the parameters that minimizes the error between the model and the measured data are returned. Since each voxel generates a time course, each voxel will have these parameters associated with it and an image can be formed that displays a “map” of each parameter as shown in Fig 3. In this way, a quantitative DCE-MRI study is conducted that yields parameter maps where each voxel in the image corresponds to a particular physiological characteristic. However, since C_t and C_p are not measured directly in an MRI study, a conversion from the measured signal intensities to the concentration time courses must also be employed. In order to make this transformation, a pre-contrast T_1 map (i.e., an image in which the value of each voxel is the T_1 value of the tissue it encompasses) and T_1 -weighted images during and after the injection of the CA are required. The pre-contrast T_1 map (denoted by T_{10}) is used to calibrate the post-contrast T_1 -weighted images to estimate a T_1 time course for each voxel. The T_1 time course is then converted to a concentration of CA time course *via*, e.g., Eq. [4]:

$$R_1 = r_1 C_x + R_{10}, \quad [4]$$

where $R_1 = 1/T_1$, r_1 is the relaxivity of the CA (which measures the efficacy of the CA at enhancing relaxation in units of $\text{mM}^{-1}\text{s}^{-1}$), C_x is C_t or C_p , and $R_{10} = 1/T_{10}$.

It should be noted that Eqs. [2] and [3] may be considered the “starting point” for DCE-MRI and there is a mature literature that exists on how to make these measurements and how to develop more elaborate models and data fitting techniques to more accurately and precisely reflect tissue properties of interest.⁸

DIFFUSION IMAGING IN CLINICAL TRIALS

Much effort has been expended in assessing the response of brain tumors *via* ADC mapping. Pope *et al.*⁹ used ADC histogram analysis to predict the response of glioblastoma multiforme to bevacizumab (humanized anti-vascular endothelial growth factor antibody) in a multi-center trial. They studied 82 patients and found that, for the bevacizumab treated patients, the pre-treatment ADC value outperformed enhancing tumor volume at first follow-up in stratifying six-month progression-free survival. The ADC values had 73% accuracy, versus 58% accuracy ($P < 0.05$) for the volume measurements. The authors concluded that pre-treatment ADC histogram analysis can stratify progression-free survival in bevacizumab treated patients. Since bevacizumab is an anti-angiogenic agent, the fact that

ADC, which reports on cell density, can stratify response to such treatment is worthy of note.

A recent development has been the application of “functional diffusion maps” to predict treatment response.¹⁰ The functional diffusion map (fDM) attempts to summarize the changes in ADC values over time by co-registering ADC maps obtained before and then after a period of treatment. Hamstra *et al.*¹¹ applied this technique in a study of 60 patients with high-grade glioma and compared it to the standard radiologic assessment of brain tumors using the Macdonald criteria, an approach based on size changes that is somewhat similar to RECIST. This study showed that the volume of tumor with increased diffusion by fDM at three weeks was the strongest predictor of patient survival at one year (with larger fDM predicting longer median survival), whereas the Macdonald criteria had a similar prognostic value at 10 weeks. The authors concluded that fDM provided an earlier assessment of equal predictive value than the Macdonald criteria, and that the combination of the two methods provided the most accurate prediction of patient survival.

The ability to predict treatment response earlier than morphological imaging has also been investigated in breast cancers. Sharma *et al.*¹² investigated the ability of ADC, tumor diameter and volume to predict response of 56 patients with locally advanced breast cancer receiving neoadjuvant chemotherapy. Measurements were made at four time points (before treatment and after each of three cycles of therapy). The authors found that after the third cycle of therapy, the sensitivity for differentiating responders from non-responders was 89% each for volume and diameter and 68% for ADC, and the specificities were 50%, 70%, and 100%, respectively. Thus ADC had a lower sensitivity but a higher specificity than anatomical imaging.

DCE-MRI IN CLINICAL TRIALS

There are a number of other excellent reports of DCE-MRI in clinical trials, and one particularly exciting result involved 37 patients with primary breast cancer who received six cycles of neoadjuvant 5-fluorouracil, epirubicin, and cyclophosphamide chemotherapy.¹³ The patients received a DCE-MRI exam before and after two cycles of treatment, with the goal being to assess whether changes in DCE-MRI after two cycles of therapy could predict final clinical and pathologic response. The authors found that a change in K^{trans} was the best predictor of pathologic nonresponse (area under the receiver operating characteristic curve, 0.93; sensitivity, 94%; specificity, 82%), correctly identifying 94% of nonresponders and 73% of responders. Interestingly, changes in MRI-derived tumor size did not predict pathologic response. The authors concluded that changes in DCE-MRI after two cycles of anthracycline-based neoadjuvant chemotherapy can predict final clinical and pathologic response. If such a study can be replicated in additional patient sets, this represents a significant step forward in imaging treatment response in breast cancer.

Another study of DCE-MRI in breast cancer attempted to use the technique to assess treatment response after only a single cycle of neoadjuvant therapy to predict clinical response and five-year survival in 24 patients with locally advanced breast cancer.¹⁴ This study is of particular note in that quantitative modeling of tumor kinetics was not assessed; rather the relative signal intensity (RSI, defined as the mean signal intensity of images two and three after contrast agent injection relative to the pre-contrast image) was used to assess response. The authors found that the mean and 10th percentile RSI values before neoadjuvant chemotherapy are significantly lower in patients surviving more than five years compared to nonsurvivors ($P = 0.05$ and 0.02 , respectively). The prospect of using such a straightforward method of analyzing contrast agent kinetics to assess long-term survival is quite compelling since such an approach could readily be implemented at many facilities.

Jarnagin *et al.*¹⁵ employed semi-quantitative and quantitative DCE-MRI to assess survival in a phase II clinical trial studying the use of regional chemotherapy for unresectable primary liver cancer. Thirty-four unresectable patients received hepatic arterial infusion with floxuridine and dexamethasone in patients with unresectable intrahepatic cholangiocarcinoma or hepatocellular carcinoma. The group showed that patients with pre-treatment area under the kinetic curve over 180 s (AUC_{180}) larger than 34.2 mM·s had a longer median survival than those with AUC_{180} less than 34 mM·s ($P = 0.002$). Quantitative analysis of DCE-MRI data showed a decrease in K^{trans} and k_{ep} ($=K^{trans}/v_e$) after the first post-treatment scan predicted survival.

ISSUES TO ADDRESS IN DWI

There are several practical limitations that have restricted the general applications of quantitative measurements of ADC in cancer clinical trials. Some of these are discussed below.

Data Acquisition in DWI

First, the data acquisition methods that are used for generating ADC maps generally produce images of inferior quality to conventional anatomic images. Thus in part because of the use of single “snapshot” imaging methods such as echo planar imaging for the acquisition of the base images from which ADC maps are calculated. Images that are sensitive to diffusion are by definition sensitive to the effects of very small movements, and gross image artifacts arise in diffusion-weighted images that make use of conventional modes of signal acquisition. Echo planar imaging is a technique in which the entire image data set can be acquired following a single excitation of the nuclei, and it thereby avoids the major “ghost” artifacts produced by small coherent movements during imaging. However, echo planar images generally have lower spatial resolution (fewer and larger voxels) and poorer signal-to-noise ratio than images that take much longer to record, and they are prone to suffer from distortions in the presence of major tissue inhomogeneities such as air spaces and bone. Higher quality diffusion weighted images can in principle be obtained using more conventional multi-shot approaches, when corrections for the signal changes induced by motion may be attempted, but these do not work well in the abdomen or thorax, when the body region changes shape and size substantially on the time scale of the imaging process (i.e., through respiratory and cardiac motion). Gross motion between scans with different degrees of weighting also precludes calculation of accurate values of ADC because voxels are not properly co-registered.¹⁶

Further technical problems may have to be addressed for wider spread use of ADC maps. The optimal degrees of diffusion weighting (and thus scanner settings) for calculating the ADC of any single voxel can be predicted from theory,¹⁷ but in any section there occurs a widespread and unknown range of ADC values, so some selections have to be made that are sub-optimal for estimating some tissue values. If tumors are very different from the majority of other tissues, their values may be the most compromised. In general it is important to introduce a significant signal change caused by diffusion during and after the diffusion-sensitizing gradients are applied. These demands the gradients be of sufficient amplitude and duration to cause a measurable signal decrease. If the gradients are weak then they have to be applied for longer, and as a result the overall signal decreases because of T_2 relaxation. Using stronger gradients has a major advantage for estimates of ADC.

ISSUES TO ADDRESS IN DCE-MRI

A number of decisions must be made when developing a DCE-MRI protocol for a clinical trial and the optimal set of acquisition or analysis parameters has not been definitively settled.

Spatial versus Temporal Resolution

The two most common applications of DCE-MRI in cancer are to: (1) report quantitatively on tissue properties including blood vessel perfusion and permeability as well as volume fractions within heterogeneous lesions; and (2) provide clinically relevant insights into lesion enhancement. While these two are related, the requirements of the data to perform these applications can be substantially different. In order to study perfusion, vessel permeability, and tissue volume fractions, kinetic modeling as described above must be performed. This requires that the CA time courses be sampled sufficiently fast; if the temporal sampling is too slow, then the curves will not be described completely and this results in significant errors in the estimation of the pharmacokinetic parameters during the curve-fitting procedure.¹⁸ Unfortunately, high temporal resolution data is incompatible with the acquisition of the high spatial resolution data required for clinical needs. This difficulty may have been actually compounded by the adoption of RECIST; to carry out a RECIST analysis investigators need high spatial resolution to track the changes in longest dimension thereby frequently precluding the use of high temporal resolution data. This makes it very difficult to perform a direct comparison between RECIST and more quantitative DCE-MRI techniques.

There are MR acquisition methods under development that may be able to deliver both high spatial and temporal resolution in the same study. Such “keyhole” methods have found some application in the DCE-MRI community and remain an active area of research.¹⁹ Additionally, some DCE-MRI data analysis techniques may relax the requirement for high temporal resolution data and we discuss those approaches below.

Arterial input functions—The rate of change of concentration of CA in the blood plasma, the so-called arterial input function (AIF) denoted by C_p above is needed for analysis with most current pharmacokinetic models and must be acquired with high temporal resolution because it changes so rapidly. Much effort has gone into developing reliable and reproducible methods for measuring the AIF. Unfortunately, imaging methods that accurately capture the AIF first-pass kinetics can result in poor signal-to-noise and/or poor spatial resolution. Rather than measuring the AIF in individual patients, some have explored using an average AIF obtained from a sub-population²⁰, assuming a model AIF²¹, or using a reference region model.²²⁻²⁴ Each method has its own strengths and weaknesses.

Several investigators have developed reference region models in which the signal intensity in a well-characterized reference tissue (e.g., muscle) is compared to the signal intensity changes in the tissue of interest (typically the tumor). These methods have the benefit of not requiring explicit measurement of the AIF and do not require as high temporal resolution because the CA kinetics in the tissue site are substantially slower than in the blood plasma. There have been several initial studies to assess reference region reproducibility and stability.^{25,26} If reference region models can be shown to be accurate and reproducible, they may provide a way to circumvent the spatial versus temporal resolution compromise.

From signal intensity to contrast agent concentration

Pharmacokinetic analysis cannot be performed directly on the signal intensity time courses measured in a DCE-MRI study; the intensity time courses must be converted to

concentration of contrast agent time courses. In most models, the tissue extravascular space is assumed to be a well-mixed, homogeneous compartment or, equivalently, that the system remains in what is typically called the fast exchange limit with respect to the water exchange between two regions. In most tissues, most water is intracellular, and the common gadolinium chelates cannot access this intracellular water directly, so water exchange between the extravascular, extracellular space and the intracellular space must be incorporated into analytic models under certain circumstances. The practical result of accounting for slower water exchange is that the mathematical formalism becomes more cumbersome and introduces additional parameters that affect the curve-fitting analysis. The pharmacokinetic parameters extracted by Eqs. [2], [3], and [4] and those extracted by the extended models may differ by many fold, with a significant effect on determining if a treatment is working or not or separating benign from malignant lesions.²⁷ However, some careful studies have indicated that this effect is not common in many DCE-MRI applications and may needlessly complicate the analysis.²⁸ Thus, there is some disagreement over which technique(s) is(are) most appropriate and this is one aspect that prevents quantitative DCE-MRI from gaining more widespread acceptance.

OVERALL NEEDS—In addition to the issues specific to DWI or DCE-MRI described above, there are a number of more general issues that are common to both techniques.

Methods of data summary

DWI and DCE-MRI return a large quantity of data that must be statistically summarized in an appropriate manner. The majority of studies report the average of a particular parameter value (e.g., ADC, K^{trans} , etc.) from a region of interest (ROI) drawn around the tumor. While this is a succinct and straightforward way to summarize the data, it clearly is not optimal since information on tumor heterogeneity and spatial location is discarded. An alternative approach is to co-register image sets taken at different points before and after treatment to a common “image space”. In this way, individual voxels can be compared for specific changes. This is the technique that was employed in the functional diffusion maps described above in which the volume of brain tumor showing the greatest increase in diffusion was the strongest predictor of patient survival at one year. It is a reasonable hypothesis that forming such parametric maps may produce similar prognostic markers. While the most progress in this arena has been made in brain cancer (due to the relative ease with which brain images can be co-registered), there have been preliminary efforts at performing this analysis in other disease sites, including the breast.^{29,30}

Repeatability/Reproducibility

Reproducibility refers to the agreement between independent results using the same method on the same subject but under different conditions (e.g., different operators, machines, or time depending on the experiment), whereas repeatability is the agreement of the above under the same condition. Compared to DCE-MRI, there has been substantially less effort on assessing the repeatability and reproducibility of diffusion imaging. Indeed, a recent consensus statement paper indicated the necessity of incorporating baseline patient reproducibility test as part of study design (see reference 38). Here we highlight recent work in this important area.

Gibbs *et al.*³¹ performed a repeatability study on ADC measurements in prostate tumors. They found that ADC values were repeatable to within 35% over both a short time period of a few minutes and a longer time period of a month. Koh *et al.*³² assessed the reproducibility of the ADC as part of a multi-site trial of combretastatin A4 phosphate in solid tumors. The group studied sixteen patients *via* EPI and found that, when care was taken to use high b-value diffusion weighted data, ADC had a good measurement repeatability of approximately

14% as quantified by the coefficient of repeatability. The authors concluded that ADC measurements were highly reproducible in the multi-center trial setting and appeared promising for evaluating the anti-cancer drugs. Paldino *et al.*³³ reported on the repeatability of DWI-derived ADC and fractional anisotropy of patients with glioblastoma multiforme in a very practical fashion. Sixteen patients underwent DWI at two time points prior to intervention. ADC maps were registered to the contrast-enhanced image volumes and the group looked at the repeatability of ADC values within the contrast-enhancing regions; they found that a sample of nine was needed, whereas 30 patients would be required to detect a subtle 5% change.

There also have been a number of reproducibility and repeatability studies in DCE-MRI. Galbraith *et al.*³⁴ studied the size of change needed for significance in a group of patients. Their results indicated that K^{trans} , v_e , and k_{ep} are sufficiently reproducible to detect changes greater than 14 - 17% in a cohort of 16 patients. Ng *et al.*³⁵ studied 23 patients with lung and liver tumors who underwent two DCE-MRI examinations within one week without treatment. The within-patient coefficients of variation for K^{trans} for liver lesions were 8.9% and 17.9% for liver and lung lesions, respectively. Estimates of confidence that changes observed in a given patient were due to intervening therapy rather than variability of the technique were calculated to range from 71% to 87% if a 20% reduction in a parameter was observed. Another scan-rescan investigation in liver and lung lesions compared two common methods of calculating perfusion parameters from DCE-MRI data: the first method employed a single-model AIF across all study data sets, while the second used an automated process to obtain a patient specific AIF.³⁶ The authors found that the use of a data-derived AIF reduced the visit-to-visit coefficient of variation in K^{trans} for liver lesions by approximately 70%, while the improvement was less than 20% for lung lesions. The use of a data-derived AIF in the analysis of DCE-MRI data provided a substantial reduction in the scan-rescan coefficient of variation on the measurement of K^{trans} . These results show a much larger advantage in the liver than in the lungs. As such, these findings provide the foundation for interpretation of the results, and design of clinical trials including estimates of patient sample size in which DCE-MRI studies are used to assess objective responses.

Direct comparison to RECIST criteria

In order for an emerging technique to be incorporated into RECIST, it will need to be shown that the newer technique significantly improves on the current RECIST approach by offering new and/or complimentary information. Unfortunately, there is a paucity of data in the literature regarding a direct comparison among the capabilities of DWI, DCE-MRI, and RECIST to predict a pathologic response. A direct comparison between DCE-MRI and RECIST may be difficult for the reasons described above related to spatial and temporal resolution. A second problem is that imaging acquisition and analysis protocols to perform DWI and DCE-MRI vary greatly from institution to institution and, often, from vendor to vendor. These methods must be standardized so that multi-center trials can be performed and compared to RECIST data.

CONCLUSIONS

While there has been a steady increase in both the quality and number of quantitative imaging methods that can report on tumor status, these approaches have not been moved effectively to routine clinical use, or even to routine use in clinical trials. Given the current limitations of DWI and DCE-MRI, we are forced to agree with the conclusion of the RECIST Working Group³⁷ and state that, at present, there is insufficient standardization to encourage the adoption of DWI and DCE-MRI for widespread use in clinical trials for which expert imaging scientists, radiologists, oncologists are not intimately involved. The imaging community recognizes this and is working hard to address these issues by

developing consensus recommendations and moving towards uniform approaches.^{38,39} However, we encourage oncologists to engage imaging specialists and radiologists from the earliest stages of trial design so that these appropriate imaging techniques can be included from the beginning of prospective studies. This can significantly increase the homogeneity with which imaging data are acquired, particularly when multi-site studies are performed.

We have attempted to give a balanced view of the state of the art of these two quantitative MRI techniques, pointing out what they can offer and what barriers remain before they can be readily and regularly incorporated into clinical trials. While there are a number of difficulties that need to be resolved, they are not insurmountable and we believe DWI and DCE-MRI will ultimately find their way into RECIST (or other criteria).

Acknowledgments

We thank the NIH for funding through NCI R01CA129661, NCI P50 CA128323, and NIBIB 1K25EB005936. We thank Dr. Adam Anderson, Ph.D. for reference³².

Supported in part by: NCI R01CA129661, NCI P50CA128323, NIBIB 1K25EB005936

REFERENCES

1. Yankeelov TE, Avison MJ, Damon BM, Manning HC, Peterson TE, Gore JC. Frontiers of Biomedical Imaging Science 2009: workshop report and research opportunities. *Cancer Res.* 2009; 69:7902–4. [PubMed: 19808955]
2. Therasse P, Arbuck SG, Eisenhauer EA, Wanders J, Kaplan RS, Rubinstein L, 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. [PubMed: 10655437]
3. Einstein A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik.* 1905; 322:549–560.
4. Anderson AW, Xie J, Pizzonia J, Bronen RA, Spencer DD, Gore JC. Effects of cell volume fraction changes on apparent diffusion in human cells. *Magn Reson Imaging.* 2000; 18:689–95. [PubMed: 10930778]
5. Hagmann P, Jonasson L, Maeder P, Thiran JP, Wedeen VJ, Meuli R. Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics.* 2006; 26:S205–S223. [PubMed: 17050517]
6. Charles-Edwards EM, deSouza NM. Diffusion-weighted magnetic resonance imaging and its application to cancer. *Cancer Imaging.* 2006; 6:135–43. [PubMed: 17015238]
7. Shellock FG, Kanal E. Safety of magnetic resonance imaging contrast agents. *J Magn Reson Imaging.* 1999; 10:477–84. [PubMed: 10508312]
8. Yankeelov TE, Gore JC. Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples. *Current Medical Imaging Reviews.* 2007; 3:91–107. [PubMed: 19829742]
9. Pope WB, Kim HJ, Huo J, Alger J, Brown MS, Gjertson D, et al. Recurrent Glioblastoma Multiforme: ADC Histogram Analysis Predicts Response to Bevacizumab Treatment. *Radiology.* 2009; 252:182–9. [PubMed: 19561256]
10. Hamstra DA, Chenevert TL, Moffat BA, Johnson TD, Meyer CR, Mukherji SK, et al. Evaluation of the functional diffusion map as an early biomarker of time-to-progression and overall survival in high-grade glioma. *Proc Natl Acad Sci U S A.* 2005; 102:16759–64. [PubMed: 16267128]
11. Hamstra DA, Galbán CJ, Meyer CR, Johnson TD, Sundgren PC, Tsien C, et al. Functional diffusion map as an early imaging biomarker for high-grade glioma: correlation with conventional radiologic response and overall survival. *J Clin Oncol.* 2008; 26:3387–94. [PubMed: 18541899]
12. Sharma U, Danishad KKA, Seenu V, Jagannathan NR. Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy. *NMR Biomedicine.* 2009; 22:104–13.

13. Ah-See ML, Makris A, Taylor NJ, Harrison M, Richman PI, Burcombe RJ, et al. Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer. *Clin Cancer Res*. 2008; 14:6580–9. [PubMed: 18927299]
14. Johansen R, Jensen LR, Rydland J, Goa PE, Kvistad KA, Bathen TF, et al. Predicting survival and early clinical response to primary chemotherapy for patients with locally advanced breast cancer using DCE-MRI. *J Magn Reson Imaging*. 2009; 29:1300–7. [PubMed: 19472387]
15. Jarnagin WR, Schwartz LH, Gultekin DH, Gönen M, Haviland D, Shia J, et al. Regional chemotherapy for unresectable primary liver cancer: results of a phase II clinical trial and assessment of DCE-MRI as a biomarker of survival. *Ann Oncol*. 2009; 20:1589–95. [PubMed: 19491285]
16. Hutton C, Bork A, Josephs O, Deichmann R, Ashburner J, Turner R. Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage*. 2002; 16:217–40. [PubMed: 11969330]
17. Bito Y, Hirata S, Yamamoto E. Optimum gradient factors for apparent diffusion coefficient measurements. *Proc Intl Soc Magn Reson Med*. 1995; 3:913.
18. Cheng HL. Investigation and optimization of parameter accuracy in dynamic contrast-enhanced MRI. *J Magn Reson Imaging*. 2008; 28:736–43. [PubMed: 18777534]
19. Martel AL, Chan RW, Ramsay E, Plewes DB. Removing undersampling artifacts in DCE-MRI studies using independent components analysis. *Magn Reson Med*. 2008; 59:874–84. [PubMed: 18302238]
20. Huang W, Wang Y, Panicek DM, Schwartz LH, Koutcher JA. Feasibility of using limited-population-based average R10 for pharmacokinetic modeling of osteosarcoma dynamic contrast-enhanced magnetic resonance imaging data. *Magn Reson Imaging*. 2009; 27:852–8. [PubMed: 19282123]
21. Weinmann HJ, Laniado M, Mutzel W. Pharmacokinetics of GdDTPA/Dimeglumine after intravenous injection into healthy volunteers. *Physiol Chem Phys Med NMR*. 1984; 16:167–72. [PubMed: 6505043]
22. Yang C, Karczmar GS, Medved M, Stadler WM. Multiple reference tissue method for contrast agent arterial input function estimation. *Magn Reson Med*. 2007; 58:1266–75. [PubMed: 17969061]
23. Yankeelov TE, Luci JJ, DeBusk LM, Lin PC, Gore JC. Incorporating the effects of transcytolemmal water exchange in a reference region model for DCE-MRI analysis: theory, simulations, and experimental results. *Magn Reson Med*. 2008; 59:326–35. [PubMed: 18228592]
24. Planey CR, Welch EB, Xu L, Chakravarthy AB, Gatenby JC, Freehardt D, et al. Temporal Sampling Requirements for Reference Region Modeling DCE-MRI Data in Human Breast Cancer Analysis. *J Magn Reson Imaging*. 2009; 30:121–34. [PubMed: 19557727]
25. Yang C, Karczmar GS, Medved M, Oto A, Zamora M, Stadler WM. Reproducibility assessment of a multiple reference tissue method for quantitative dynamic contrast enhanced-MRI analysis. *Magn Reson Med*. 2009; 61:851–9. [PubMed: 19185002]
26. Yankeelov TE, DeBusk LM, Billheimer DD, Luci JJ, Lin PC, Price RR, et al. Repeatability of a reference region model for analysis of murine DCE-MRI data at 7T. *J Magn Reson Imaging*. 2006; 24:1140–7. [PubMed: 17024660]
27. Huang W, Li X, Morris EA, Tudorica LA, Seshan VE, Rooney WD, et al. The magnetic resonance shutter speed discriminates vascular properties of malignant and benign breast tumors in vivo. *Proc Natl Acad Sci U S A*. 2008; 105:17943–8. [PubMed: 19004780]
28. Buckley DL, Kershaw LE, Stanisiz GJ. Cellular-interstitial water exchange and its effect on the determination of contrast agent concentration in vivo: dynamic contrast-enhanced MRI of human internal obturator muscle. *Magn Reson Med*. 2008; 60:1011–9. [PubMed: 18956419]
29. Ma B, Meyer CR, Pickles MD, Chenevert TL, Bland PH, Galbán CJ, et al. Voxel-by-voxel functional diffusion mapping for early evaluation of breast cancer treatment. *Inf Process Med Imaging*. 2009; 21:276–87. [PubMed: 19694270]
30. Li X, Dawant BM, Welch EB, Chakravarthy AB, Freehardt D, Mayer I, et al. A Nonrigid Registration Algorithm for Longitudinal Breast MR Images and the Analysis of Breast Tumor Response. *Magn Reson Imaging*. 2009; 27:1258–70. [PubMed: 19525078]

31. Gibbs P, Pickles MD, Turnbull LW. Repeatability of echo-planar-based diffusion measurements of the human prostate at 3 T. *Magn Reson Imaging*. 2007; 25:1423–9. [PubMed: 17499468]
32. Koh DM, Blackledge M, Collins DJ, Padhani AR, Wallace T, Wilton B, et al. Reproducibility and changes in the apparent diffusion coefficients of solid tumours treated with combretastatin A4 phosphate and bevacizumab in a two-centre phase I clinical trial. *Eur Radiol*. 2009; 19:2728–38. [PubMed: 19547986]
33. Paldino MJ, Barboriak D, Desjardins A, Friedman HS, Vredenburgh JJ. Repeatability of quantitative parameters derived from diffusion tensor imaging in patients with glioblastoma multiforme. *J Magn Reson Imaging*. 2009; 1199–205. [PubMed: 19388113]
34. Galbraith SM, Lodge MA, Taylor NJ, Rustin GJ, Bentzen S, Stirling JJ, et al. Reproducibility of dynamic contrast-enhanced MRI in human muscle and tumours: comparison of quantitative and semi-quantitative analysis. *NMR Biomed*. 2002; 15:132–42. [PubMed: 11870909]
35. Ng CS, Raunig DL, Jackson EF, Ashton EA, Kelcz F, Kim KB, et al. Reproducibility of perfusion parameters in dynamic contrast-enhanced MRI of lung and liver tumors: effect on estimates of patient sample size in clinical trials and on individual patient responses. *AM J Roentgenol*. 2010; 194:W134–40. [PubMed: 20093564]
36. Ashton E, Raunig D, Ng C, Kelcz F, McShane T, Evelhoch J. Scan-rescan variability in perfusion assessment of tumors in MRI using both model and data-derived arterial input functions. *J Magn Reson Imaging*. 2008; 28:791–6. [PubMed: 18777526]
37. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009; 45:228–47. [PubMed: 19097774]
38. Padhani AR, Liu G, Koh DM, Chenevert TL, Thoeny HC, Takahara T, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia*. 2009; 11:102–25. [PubMed: 19186405]
39. Evelhoch J, Garwood M, Vigneron D, Knopp M, Sullivan D, Menkens A, et al. Expanding the use of magnetic resonance in the assessment of tumor response to therapy: workshop report. *Cancer Res*. 2005; 65:7041–4. [PubMed: 16103049]

Fig 1.

Panels A through C depict sagittal views of a breast acquired with different flip angle values. In each of these three images the contrast reflects variations in signal intensity that do not have physical units associated with them and are difficult to interpret in terms of individual physiological or physical parameters. By analyzing the data from multiple images, and by assuming particular mathematical relationships between images, the intrinsic properties can be computed. These three data sets can be used to calculate a quantitative T_1 map (panel D) wherein every voxel in the image corresponds to the tissue's T_1 in units of seconds.

Fig 2.

Panels A and B depict sagittal non-diffusion weighted and diffusion-weighted images of an invasive ductal carcinoma. Using a model, these two data sets can be used to compute a quantitative apparent diffusion coefficient (ADC) map (panel C) wherein every voxel in the image corresponds to the tissue's ADC in units of mm²/s.

Figure 3.

Panels A and B depict two T_1 -weighted sagittal images taken from a time series of a dynamic contrast enhanced MRI study of a patient with an invasive ductal carcinoma; panel A is the pre-contrast image, while panel B corresponds to the 5 minute post-contrast image. Using a model (e.g., Eq. [2] or [3]) and the time series data, a quantitative parametric map wherein every voxel in the image corresponds to the tissue's contrast agent transfer coefficient (K^{trans}) in units of mL(blood)/mL(tissue)/(min).