Review Article

Current and Future Trends in Magnetic Resonance Imaging Assessments of the Response of Breast Tumors to Neoadjuvant Chemotherapy

Lori R. Arlinghaus,[1] Xia Li,[1] Mia Levy,[2,3] David Smith,[1] E. Brian Welch,[1] John C. Gore,[1,4,5,6] and Thomas E. Yankeelov[1,4,5,7]

[1] *Department of Radiology and Radiological Sciences, Institute of Imaging Science, Nashville, TN 37232-2310, USA*

[2] *Department of Biomedical Informatics, Institute of Imaging Science, Nashville, TN 37232-2310, USA*

[3] *Department of Medicine, Institute of Imaging Science, Nashville, TN 37232-2310, USA*

[4] *Department of Biomedical Engineering, Institute of Imaging Science, Nashville, TN 37232-2310, USA*

[5] *Department of Physics and Astronomy, Institute of Imaging Science, Nashville, TN 37232-2310, USA*

[6] *Department of Molecular Physiology and Biophysics, Institute of Imaging Science, Nashville, TN 37232-2310, USA*

[7] *Department of Cancer Biology, Institute of Imaging Science, Nashville, TN 37232-2310, USA*

Correspondence should be addressed to Thomas E. Yankeelov, thomas.yankeelov@vanderbilt.edu

The current state-of-the-art assessment of treatment response in breast cancer is based on the response evaluation criteria in solid tumors (RECIST). RECIST reports on changes in gross morphology and divides response into one of four categories. In this paper we highlight how dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) and diffusion-weighted MRI (DW-MRI) may be able to offer earlier, and more precise, information on treatment response in the neoadjuvant setting than RECIST. We then describe how longitudinal registration of breast images and the incorporation of intelligent bioinformatics approaches with imaging data have the potential to increase the sensitivity of assessing treatment response. We conclude with a discussion of the potential benefits of breast MRI at the higher field strength of 3T. For each of these areas, we provide a review, illustrative examples from clinical trials, and offer insights into future research directions.

1. Introduction

In recent years, neoadjuvant chemotherapy has become increasingly important for patients with operable breast cancer requiring chemotherapy [1, 2]. There are several important reasons for this trend. These include the potential ability to reduce primary tumor burden and thereby allowing for higher rate of breast conservation surgery, as well as earlier treatment of micrometastatic disease. Neoadjuvant chemotherapy also provides an opportunity to evaluate (*in vivo*) the response of a particular treatment on an individual patient basis which may inform the selection of postoperative chemotherapy. Furthermore, according to National Surgical Adjuvant Breast and Bowel Project B-18 (NSABP B-18)

results, neoadjuvant chemotherapy yields a survival outcome similar to adjuvant treatment in early breast cancer patient [3]. Thus, noninvasive imaging techniques which can report on the status of the response of breast tumors to neoadjuvant chemotherapy is of great import.

While X-ray mammography and ultrasound imaging play a critical role in the detection and diagnosis of breast cancer, they do not provide quantitative information on their response to therapy. Additionally, the 2D nature of planar X-ray mammography makes it difficult to interpret potentially complex 3D tumor structures, thus severely limiting its value in quantitative longitudinal studies. Consequently, there are currently no radiological methods for adequately assessing the state of breast tumors or their response to treatments

to guide clinical decisions. Although serial biopsies may be obtained to measure changes at the cellular and molecular levels, the spatial sampling may be poor and the results misleading [4–7]. Clinical judgments using physical examination, mammograms and ultrasound to measure short-term treatment effects have a high inter-observer variability and are prone to error. The development of appropriate methods of tissue characterization that could be applied early in the course of treatment to assess response would allow clinicians to individualize therapy based on each patient's response to a particular agent. Quantitative imaging techniques which can reliably assess tumor response are urgently needed to improve clinical breast cancer care.

Three generations of solid tumor response criteria have been applied to assess treatment response in breast cancer clinical trials. The original WHO criteria developed in 1981 used the sum of two-dimensional measurements of cancer lesions to compute a composite estimate of tumor burden while the currently accepted standard method Response Evaluation Criteria in Solid Tumors (RECIST) uses single dimensional measurements as measured by ultrasound, X-ray-computed tomography (CT), or magnetic resonance imaging (MRI). While details of the criteria are presented elsewhere [8], we briefly summarize the salient features here.

In the pretreatment scan, "target lesions are identified" and their longest dimension and perpendicular short axis are measured. For lymph node lesions, the short axis is used for the composite tumor burden score while the long axis is used for nonlymph node lesions. The "baseline sum diameters" are then defined by adding the single dimension measurement of all the target lesions. Additional scans are performed after treatment and analyzed similarly. The percent change in the sum of diameters from the baseline assessment to the follow-up assessment after treatment is calculated and this quantity is used to classify response. A complete response (CR) is defined as the disappearance of all target lesions; a partial response (PR) is defined as a 30% decrease in the sum of diameters of the target lesions; progressive disease (PD) is defined as a 20% increase in the sum of diameters of the target lesions or the appearance of any new lesions; and stable disease (SD) is defined as changes that do not meet the previous three criteria. While many clinical trials have used these criteria for assessing cancers at many disease sites (see [9] and references cited therein), it is well recognized that this approach needs to be significantly improved for a number of reasons including practical, technical, and scientific issues. Practical and technical issues included how to select the total number of lesions to include in the sum longest dimension calculation, and how to apply RECIST in trials where progression, and not response, is the primary endpoint. However, more fundamental (scientific) issues have also been raised, including whether or not to include emerging imaging techniques such as, fluorodeoxyglucose PET (FDG-PET) or dynamic contrast-enhanced MRI (DCE-MRI) with RECIST. A second fundamental issue is whether RECIST should be included—or if it is even relevant—in assessing the activities of noncytotoxic anticancer drugs. The metric for positive response is based only on anatomical and morphological changes which are (temporally) downstream

manifestations of underlying pathophysiological changes which may occur earlier. The current (and presumably many future) generations of anticancer therapy use molecularly targeted agents, so changes in morphology may not be the most appropriate method to assess response.

Response criteria continue to evolve and require validation as new technologies for assessing disease emerge. In early 2009, RECIST was updated, as RECIST 1.1, and made substantial progress towards addressing these (and other) issues [9]. A major development was the incorporation of fluorodeoxyglucose positron emission tomography (FDG-PET) in the guidelines, albeit in the limited role of assessing progressive disease. This is at once exciting and sobering; exciting in the sense that this is the first time the working group has incorporated a nonanatomical imaging modality in the assessment of tumor response, a trend that will almost certainly continue in the future. However, it is also sobering because FDG-PET was first employed in a human imaging study in 1976 [10] and required 33 years to advance to the point that it could be recommended as a method for assessing one aspect (i.e., progressive disease) of tumor response. In part, this is due to the difficulty in developing, optimizing, and standardizing advanced imaging techniques to permit results to be obtained and compared across vendors and sites. In light of this, the RECIST working group concluded that "there is not sufficient standardization or evidence to abandon anatomical assessment of tumor burden". Nonetheless, it makes clear that more advanced and specific imaging methods will be incorporated into RECIST in future years and this fact is not lost on the PET community who has begun to formulate the PET Response Criteria in Solid Tumors, or "PERCIST" [11].

Clearly, methods are needed to characterize the underlying pathophysiological changes induced by specific targeting agents. Such methods may be considerably more likely to offer earlier—and more specific—information on response to treatment when compared to changes in longest tumor dimensions. This may be especially true in the case of neoadjuvant treatment of breast cancer. Neoadjuvant treatment offers a unique opportunity for correlation of image-based changes in tumor burden with the pathologic response observed at the time of definitive surgery. Pathologic complete response is highly predictive of survival in breast cancer and offers an earlier endpoint for preliminary validation of image-based response criteria. Several aspects of MRI data analysis, acquisition, and synthesis have matured to the point where they can offer quantitative information on breast tumor status and response to therapy.

In this contribution, we begin by defining what is meant by "quantitative imaging" and then proceed to review both diffusion weighted MRI (DW-MRI) and dynamic contrast-enhanced MRI (DCE-MRI). We then proceed to discuss how these techniques can best be synthesized (through image registration and the incorporation of bioinformatic approaches) to offer optimal insight into treatment effects. In each section we introduce the basis of the technique, offer several illustrative examples, and then highlight areas for further research. We conclude with a discussion of what high magnetic field scanners (3T), which are gradually

permeating clinical practice, can offer for the quantitative imaging of breast cancer.

2. Beyond Signal Intensity: Quantitative Imaging

Any digital image, regardless of whether it was acquired with a common camera or a sophisticated medical imaging device, consists of a matrix of signal intensities. In the case of a photograph captured with a camera, the signal intensities in each pixel are given by a combination of the colors red, green, and blue, whereas in a medical image, the displayed signal intensity can be made proportional to one or more physical or physiological properties of tissue. In standard of care clinical imaging, the signal intensities in each voxel are acquired in such a way so as to optimize the contrasts between different tissue types. This improves the ability of a clinician to detect abnormalities and allows for investigators to apply the RECIST criteria described above. However, the signal intensities are usually a complicated mixture of several different quantities related to both physiology and scanner characteristics that do not have physical units associated with them.

Quantitative imaging connotes techniques in which the signal intensity in each voxel does have physical units and is therefore associated with a particular physical or physiological phenomenon. For example, there are techniques available that report on tissue oxygen status, blood flow, or metabolism; in each of these cases, the signal intensities within the image can be made proportional to mmHg, mL(blood)/mL(tissue)/min, and the rate of glucose metabolism, respectively. Quantitative cancer imaging is motivated by the hypothesis that changes in morphology are only the downstream manifestations of underlying pathophysiological changes (such as those just above), but that if it were possible to reliably measure more relevant properties we would be able to assess treatment response at an early stage and ultimately, perhaps, contribute to individualized patient care. Given the fact that anticancer treatments are increasingly specific, it is necessary to develop diagnostic techniques that can report on the specific effects of those treatments.

3. DCE-MRI

3.1. Fundamentals of DCE-MRI. DCE-MRI involves the rapid acquisition of images before, during, and after the injection of a contrast agent (CA). MRI CAs are pharmaceuticals administered to patients that are designed to increase the contrast between different tissues by changing a tissue's inherent relaxation times, T_1 and/or T_2. These relaxation properties reflect the time it takes for the nuclear magnetization, induced by the external polarizing magnetic field, to recover back towards equilibrium after being disturbed by radiofrequency emissions during the MRI process. The most common MRI CAs are gadolinium-chelates such as, Gd-DTPA. In a typical DCE-MRI imaging session, a region of interest (ROI) is selected for study (e.g., a tumor locus), and a series of heavily T_1-weighted MR images are collected

before, during, and after a CA is injected into the antecubital vein of a patient. Each image in the series corresponds to one time point, and each pixel in each image set gives rise to its own signal intensity time course which can be analyzed mathematically. In general, quantitative analyses are based on compartmental modeling whereby the tissue within the tissue element corresponding to each voxel is divided into a series of partitions, between which contrast agent may move back and forth. In the case of DCE-MRI, each voxel is typically assumed to consist of two compartments, the vascular space and the extravascular space. By considering the rate of change of the concentration of CA in extravascular space, we can write down the following model:

$$\frac{d}{dt}C_t(t) = k_1 C_p(t) - k_2 C_t(t), \tag{1}$$

where $C_t(t)$ and $C_p(t)$ are the time course of concentrations of CA in the tissue and plasma space, respectively, and k_1 and k_2 are the rate constants describing the exchange of the CA from plasma to tissue and tissue to plasma, respectively. By convention, K^{trans} and v_e denote the volume transfer constant and the volume of extravascular extracellular space, respectively. It can readily be shown [12] that $k_1 = K^{\text{trans}}$ and $k_2 = K^{\text{trans}}/v_e$. With these assignments, (1) (details provided in [12]) has the following solution:

$$C_t(T) = K^{\text{trans}} \int_0^T C_p(t) \exp\left(-\frac{K^{\text{trans}}}{v_e(t-T)}\right) dt. \tag{2}$$

Equation (2) is frequently referred to as the standard model and can be extended to include a vascular component explicitly in which case the model is given by (3)

$$C_t(T) = K^{\text{trans}} \int_0^T C_p(t) \exp\left(-\frac{K^{\text{trans}}}{v_e(t-T)}\right) dt + v_p C_p(T), \tag{3}$$

where v_p is the plasma volume fraction. In situations where the blood plasma volume is very small (say, less than 2%), (2) may be adequate to assess the DCE data. However, as v_p gets larger, (2) becomes less applicable and (3) is preferred. Of course, *a priori*, the approximate plasma volume of the tumor may not be known and that is why many investigators prefer using (3) to fit DCE time courses.

The potential clinical utility of K^{trans} and v_p is that they offer the ability to noninvasively and quantitatively report on tumor vascular status; K^{trans} estimates tumor vessel blood flow and permeable, while v_p reports on the volume fraction of a section of tissue that consists of blood. Both of these measures can be used in studies wherein the therapy is designed to affect tumor vascular status. Changes in the extravascular extracellular volume fraction (as estimated by v_e) are more complicated to interpret as there are many treatments that might change the volume and geometry of this space. Most studies that employ DCE-MRI to assess treatment response focus on K^{trans} and v_p.

In order to perform quantitative DCE-MRI analysis, the time courses of $C_t(t)$ and $C_p(t)$ must be estimated, then those time courses can be input into a curve-fitting routine

to extract K^{trans}, v_p, and v_e on a voxel-by-voxel or region-of-interest basis. Figure 1 presents an example of this approach. In DCE-MRI this is not a straightforward process because the CA is not directly measured—only its affect on the tissue's native T_1 relaxation is measured in the study. Thus, the MR signal time course must be converted to a time course representing the concentration of CA. The majority of studies employ (4) for this transformation:

$$R_1 = r_1 C_x + R_{10}, \tag{4}$$

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 $R_{10} = 1/T_{10}$, where T_{10} is the longitudinal relaxation time prior to administration of CA. Equation (4) assumes that the extravascular space is a homogeneous solution and that the system remains in what is called the "fast exchange limit" (FXL) with respect to the water exchange between the extravascular extracellular space and the extravascular intracellular space. In most tissues, most water is intracellular, and because the common Gadolinium chelates cannot access this intracellular water directly, water exchange between the extravascular extracellular space and the extravascular intracellular space must be explicitly incorporated into analytic models under certain circumstances [13, 14]. Similar comments apply to water exchange between the intravascular and extravascular spaces when using an intravascular agent [15]. In these cases, longitudinal relaxation may exhibit a recovery that is not well described by a single exponential or single time constant as discussed by several investigators [16–18]. The effects of slower water exchange may be incorporated into analyses of the MRI signal and then the longitudinal relaxation can often be characterized as exhibiting a bi-exponential decay. More specifically, the signal intensity equation in a spoiled gradient recalled echo acquisition (the main choice for DCE-MRI studies) is given by

$$\frac{S}{S_0} = \frac{1 - \exp(-TR/T_1) \cdot \sin\alpha}{1 - \cos\alpha \cdot \exp(-TR/T_1)}, \tag{5}$$

where α is the flip angle, S_0 is a reference signal that depends on the scanner gain and proton density, TR is the repetition time, and we have assumed that the echo time (TE) is much less than the apparent transverse relaxation rate T_2^*. However, in the case of bi-exponential decay of longitudinal relaxation, (5) must be amended to

$$\frac{S}{S_0} = a_L \frac{1 - \exp(-TR/T_{1L}) \cdot \sin\alpha}{1 - \cos\alpha \cdot \exp(-TR/T_{1L})} + a_S \frac{1 - \exp(-TR/T_{1S}) \cdot \sin\alpha}{1 - \cos\alpha \cdot \exp(-TR/T_{1S})}, \tag{6}$$

where a_L and R_{1L} are the fraction and relaxation rates, respectively, for the component with the longer T_1 (i.e., $R_{1L} \equiv 1/T_{1L}$); a_S and R_{1S} are defined similarly for the

component with the shorter T_1. Furthermore, instead of (4) we have

$$\frac{1}{T_{1S,1L}} = \frac{1}{2}\left(\frac{1}{T_{1A}} + \frac{1}{\tau_A} + \frac{1}{T_{1B}} + \frac{1}{\tau_B}\right) \pm \frac{1}{2}\left[\left(\frac{1}{T_{1A}} + \frac{1}{\tau_A} + \frac{1}{T_{1B}} + \frac{1}{\tau_B}\right)^2 - \left(\left(\frac{1}{T_{1A}} + \frac{1}{\tau_A}\right)\left(\frac{1}{T_{1B}} - \frac{1}{\tau_B}\right) - \frac{1}{\tau_A \tau_B}\right)\right]^{1/2}, \tag{7}$$

where T_{1A} and T_{1B} are the relaxation times for components A and B, respectively, τ_A denotes the average lifetime of a water proton in compartment A, and τ_B denotes the average lifetime in the compartment B. In this formalism the extravascular extracellular space can be taken to be compartment A while the extravascular intracellular space can be taken to be compartment B. The pharmacokinetic parameters obtained using (2), (4), and (5) (or (3), (4), and (5)) and those extracted by (2), (6), and (7) may differ significantly [14, 19–21] and these differences have a significant effect on establishing if a tumor is responding to treatment. However, there is some disagreement as to whether this formalism is truly justified in a standard DCE-MRI study and it is therefore an active area of investigation [22].

The various models have their respective strengths and weaknesses; the FXL models with and without a v_p term are straightforward to apply, but may not be as physically accurate as the exchange models. However, the exchange models require specialized data (high temporal resolution and high signal-to-noise data) that may not always be available in the clinical setting. Determining the optimal DCE analysis method for breast studies is an area of active investigation.

3.2. DCE-MRI in Breast Cancer Clinical Trials. A considerable number of studies of cancer have applied DCE-MRI methods but have employed only qualitative or semi-quantitative analyses of these data, that is, the investigators did not perform complete modeling, rather they studied features such as, the volume of the enhancing region and the area under the signal intensity time course, and so forth. There have been substantially less studies that have included quantitative modeling of DCE-MRI time courses (see, e.g., [23–25]), and when quantitative models have been employed, different authors have not always shown similar results.

To date, the most complete study of contrast-enhanced MRI in assessing the response of breast tumors to therapy was reported by Ah-See et al. [26] who applied both DCE-MRI and dynamic susceptibility MRI (DSC-MRI, for a review see [27]). The authors studied the ability of K^{trans}, k_{ep} ($\equiv K^{\text{trans}}/v_e$), v_e, maximum contrast agent concentration, relative blood flow, relative blood volume, and mean transit time to predict treatment response in 28 patients receiving
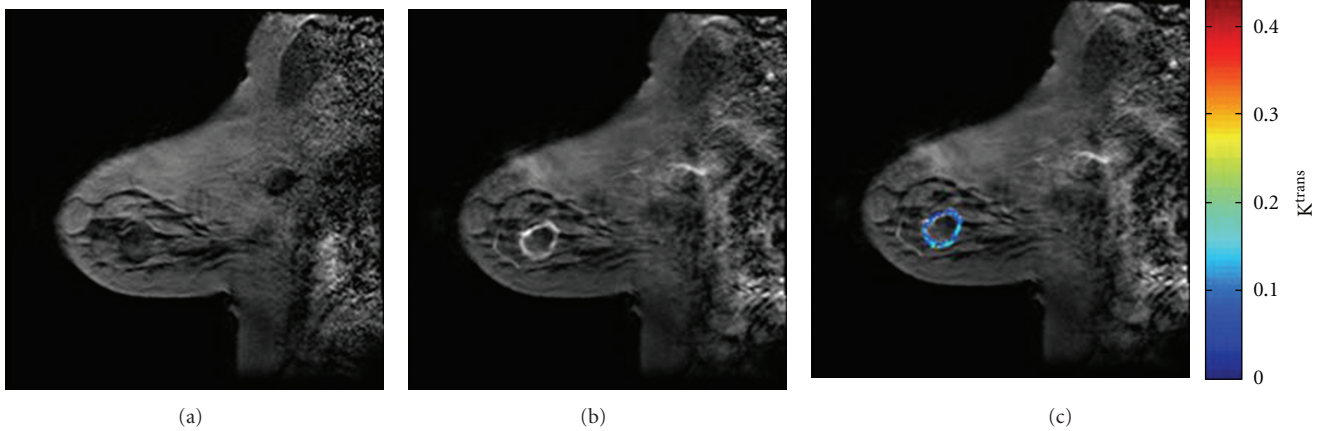
(a)   (b)   (c)

FIGURE 1: (a) and (b) display pre-contrast and postcontrast T_1-weighted sagittal images of a breast tumor. By considering the time course from each voxel, a parametric map can be generated that reports on, for example, the volume transfer constant (K^{trans}) as displayed in (c).

5-fluorouracil, epirubicin and cyclophosphamide neoadjuvant chemotherapy. While all parameters correlated with pathological response, K^{trans} was the best predictor with an area under the receiver operating characteristic curve of 0.93 while changes in tumor size did not predict for pathologic response. This is a notable result because it offers evidence of an emerging functional imaging technique that outperforms conventional morphological imaging.

However, a similar study by Yu et al. reported somewhat different results [28]. These investigators studied the ability of tumor size, K^{trans} and k_{ep} to predict final clinical response based on changes after two and four cycles of anthracycline and cyclophosphamide neoadjuvant chemotherapy in 29 patients with invasive breast cancer. While there was a significant correlation between tumor size and the pharmacokinetic parameters, the area under the receiver operating characteristic curve differentiating between nonresponders and responders was 0.88, 0.77, and 0.63 for tumor size, k_{ep}, and K^{trans}, respectively. The authors concluded that early tumor size changes on MRI after one cycle of therapy provided the most accurate predictor in assessing response to this treatment regimen.

The reasons for such discrepancies are not entirely clear but almost certainly related to the differences in treatment regimen and data acquisition techniques and timing of the second imaging session. (The two groups actually used very similar analysis techniques based on (2) described above.) In particular, Ah-See et al. acquired the second imaging data set after the second cycle of therapy, while Yu et al. acquired the second imaging data set after the first cycle of therapy. It is possible that the changes observed by the two groups are due to the timing of the second imaging session. Also, of particular interest is the issue of temporal resolution and contrast agent injection rate. Ah-See et al. acquired data with a temporal resolution of 12 seconds during their dynamic study after injection of 0.1 mmol/kg of Magnevist (Bayer Pharmaceuticals) at 4 mL/sec, whereas Yu et al. acquired data with a 42-second temporal resolution after the injection of 0.1 mmol/kg of Omniscan (GE Healthcare) at the rate of

approximately 1 mL/sec. (As we will see below, the issue of temporal versus spatial resolution is of critical importance and is a difficult balancing act.) The accuracy of estimating K^{trans} is determined by the initial rapid changes of the signal intensity time course so if the curve is not sampled adequately errors in K^{trans} may result [29]. It is possible that the differences in temporal resolutions and injection rates led to the differences in K^{trans} estimates, and it is generally appreciated that differences in acquisition protocol can lead to substantially different outcomes [30].

3.3. Research Opportunities in DCE-MRI of the Breast. As described in Section 3.1, there are a number of methods available for both data acquisition and analysis in a DCE-MRI study and the optimal methods have yet to be identified. For DCE-MRI to have the broadest impact, it will be necessary to establish the most accurate and precise methods for assessing treatment response. It is imperative, particulary when performing multi-institutional studies, to have a standard technique in place for which reproducibility at each site has been established. A particular aspect of the analysis that must be addressed is that of the arterial input function.

Before a curve fitting routine can be employed to extract the pharmacokinetic parameters discussed above, the time rate of change of the concentration of CA in the blood plasma (C_p), which is typically called the arterial input function (AIF), must be established. This time course changes very rapidly, so images have to be acquired with high temporal resolution, but when images are acquired quickly there is a net decrease in signal-to-noise ratio and/or spatial resolution. The need to acquire high temporal resolution data confines quantitative DCE-MRI in two fundamental ways. The first limitation is that it limits the ability to probe tumor heterogeneity and scan large sections of tissue at high resolution. The second limitation is that high temporal resolution acquisitions limit the clinical adoption of the technique because high spatial resolution data covering large sections of tissue are required for clinical reading and applying the RECIST criteria; in particular, high temporal

resolution data is usually limited to scanning one breast which contrasts with the clinical aim of examining both breasts. Some investigators have attempted to overcome this problem by using a population-based AIF [31, 32], reference region models [33, 34], or keyhole imaging techniques [35, 36]. Currently, there are three major kinds of AIFs employed in breast studies, in increasing order of difficulty to obtain, these are: (1) a bi-exponential decay model; (2) a population based AIF, or (3) an individually measured AIF. The first of these does not require any new data acquisition or analysis, while the second method requires measurement of the AIF in a suitable population of patients or volunteers. The third method is the most complex but also (potentially) the most accurate. Using a model or population based AIF reduces the demands on the data acquired, but if there is a substantial difference between the model or population AIF and an individual's AIF then this may lead to significant errors in the pharmacokinetic parameters returned from such an analysis. The accurate characterization of the AIF represents an aspect of breast DCE-MRI that could benefit from further research.

4. DW-MRI

4.1. Fundamentals of DW-MRI. The contrast mechanism behind DW-MRI is based upon the effects of tissue microstructures and organization on the self-diffusion of water. The random motions of water molecules, also known as Brownian motion, are due to thermal energy and were first analyzed mathematically by Einstein [37]. For freely diffusing molecules, the distribution of displacements that the particles will experience over a given amount of time (t) is described by a Gaussian function. The standard deviation is determined by the distribution which is defined by the diffusion coefficient (D) of the molecules, which depends upon the temperature, the size of the particle, and the viscosity of the medium. The mean-squared displacement of the molecules in one dimension over the diffusion time t is given by the equation

$$\langle x^2 \rangle = 2Dt. \tag{8}$$

For example, free diffusion of water molecules at normal body temperature, with a diffusion coefficient of 3×10^{-3} mm^2/s, would result in an average displacement of $17–25\,\mu$m in one dimension over diffusion times of 50–100 ms. However, water molecules within tissue are not freely diffusing on this time scale. They encounter macromolecules, cell membranes, and other tissue microstructures that hinder diffusion, causing a reduction in the average diffusion displacement, compared to that of free water. This effective, or apparent diffusion coefficient (ADC) can be measured with DW-MRI [38].

The signal intensity for each voxel acquired in DW-MRI is dependent on the apparent diffusion coefficient of water (we now use the term D interchangeably with ADC in a single direction), the signal acquired with no diffusion weighting (S_0), and a diffusion-weighting factor, b [39]:

$$S = S_0 \exp(-bD), \tag{9}$$

where

$$b = \gamma^2 \delta^2 G^2 \left(\Delta - \frac{\delta}{3} \right), \tag{10}$$

γ is the gyromagnetic ratio of protons, δ is the duration of the diffusion-weighting gradient G, and Δ is the time (equivalent to t in (8)) over which diffusion effects on the MRI signal are manifest.

The value of b is known and determined by the imaging parameters. Therefore, only two images, one nondiffusion-weighted image ($G = 0$) and one diffusion-weighted image ($G > 0$) are needed to obtain S_0 and S, respectively, and to calculate a map of D in a single direction. In practice, however, it is common to apply diffusion-weighting in three orthogonal directions (i.e., x, y, and z) where $G_x = G_y = G_z > 0$. The ADC maps calculated from the three images are then averaged to obtain a mean ADC map, reducing the effects of anisotropic diffusion.

4.2. DW-MRI in Breast Cancer Clinical Trials. Tumors typically exhibit decreased ADC values compared to the surrounding unaffected tissue because of their increased cellularity [40–42] (e.g., see Figure 2). The sensitivity of DW-MRI to these changes in tissue microstructure and organization has led to increasing interest in the potential of ADC measurements as a biomarker for diagnosis of cancer, tumor differentiation, and treatment assessment. Preclinical studies of the effects of successful treatment on tumor ADC values produced provocative results. For example, Galons et al. [43] demonstrated that successful treatment of human breast cancer xenografts with Paclitaxel resulted in a significant increase in tumor ADC within just two days of the initiation of treatment, and these changes preceded significant reductions in tumor volume.

Although there have been few clinical studies of the effects of treatment on tumor ADC in breast cancer published to date, the results of the existing studies are promising. In a study of 11 women with biopsy-proven infiltrating breast cancer undergoing neoadjuvant chemotherapy (Taxotere or adriamycin/Cytoxan), Yankeelov et al. reported that the mean tumor ADC value increased significantly ($P = .005$) by 24% after a full course of chemotherapy [44]. While this study suggests that ADC may be a useful imaging biomarker for assessing overall response to therapy in breast tumors, it does not provide insight into its usefulness as an early biomarker. Evidence of early changes in ADC due to treatment was reported by Pickles et al. [45]. In this study, both tumor longest diameter and ADC values were measured in eight women with biopsy-proven invasive breast cancer who were undergoing neoadjuvant chemotherapy (four cycles of epirubicin and cyclophosphamide). Measurements were made prior to the start, after the first cycle, and after the second cycle of therapy. A significant increase (14%, $P = .005$) in mean tumor ADC was observed after a single cycle of chemotherapy, while tumor longest diameter remained relatively unchanged (-1%, $P = .852$). After two treatment cycles, the tumor longest diameter began to approach a significant decrease (-20%, $P = .057$) compared
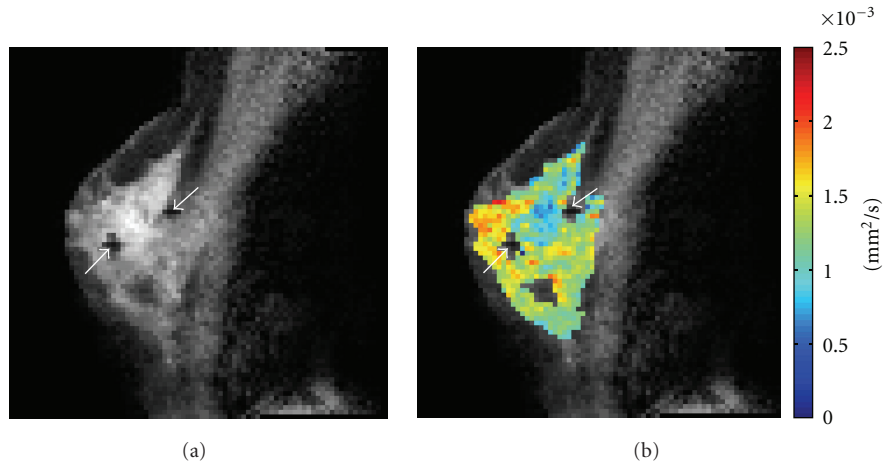
FIGURE 2: ADC values are reduced in breast cancer tumors compared to the unaffected tissue. Shown here is a postcontrast T_1-weighted image (a) for reference and an overlay of the corresponding ADC map (b). The tumor, located above and between the biopsy clips (white arrows), exhibits increased signal in the T_1-weighted image and decreased ADC values, compared to the surrounding tissue.

to the pretreatment value, and the mean ADC value was 27% higher than the pretreatment value ($P = .004$).

More recently, Sharma et al. [46] compared tumor ADC values at multiple time points throughout the course of therapy in a study of 56 women with cytologically proven infiltrating ductal carcinoma undergoing neoadjuvant chemotherapy (cyclophosphamide, adriamycin (doxorubicin) and 5-fluorouracil (5-FU); cyclophosphamide, epirubicin and 5-FU; or paclitaxel and epirubicid). ADC measurements were made before, after the first cycle, after the second cycle, and after the third cycle of treatment, although not all participants were able to be scanned at all four time points. Overall, the mean tumor ADC increased significantly ($P < .01$) from $0.95 \pm 0.11 \times 10^{-3}$ mm^2/s ($N = 53$) before treatment to $1.09 \pm 0.16 \times 10^{-3}$ mm^2/s ($N = 14$) after 1 treatment cycle. In the 11 patients scanned both before treatment and after one treatment cycle, tumor ADC increased significantly ($P = .002$) from $1.01 \pm 0.07 \times 10^{-3}$ mm^2/s before treatment to $1.15 \pm 0.13 \times 10^{-3}$ mm^2/s. Mean tumor ADC continued to increase significantly ($P < .01$) throughout treatment to $1.21 \pm 0.22 \times 10^{-3}$ mm^2/s ($N = 24$) after the second treatment cycle and $1.30 \pm 0.24 \times 10^{-3}$ mm^2/s ($N = 29$) after the third treatment cycle. Tumor diameter and volume reductions were observed only after the second cycle of treatment. The authors suggest that changes in ADC may be more useful for personalized treatment management than the current standard of changes in tumor morphology.

4.3. Research Opportunities in DW-MRI of the Breast. While the results of the DW-MRI studies mentioned in the previous section are encouraging, the application of DW-MRI in breast cancer is still relatively new and there are several technical challenges that need to be addressed. First, the chemotherapy agents used in these studies were all cytotoxic. Successful treatment with these agents leads to cell death, ultimately increasing the extracellular volume and/or reducing the cell density within the tumor, which would be expected to lead to an increase in the tumor ADC. The effects of hormonal treatments and antiangiogenic and cytostatic chemotherapy agents on the tumor ADC are not fully understood yet and need to be explored.

Timing of the ADC measurements is another issue that has not been fully addressed yet. In preclinical studies, a significant increase in ADC in response to treatment has been measured in as few as two days [43]. However, the earliest posttreatment measurements in the clinical studies mentioned above were acquired after the first course of treatment, typically three to four weeks after the start of treatment [45, 46]. The course of tumor ADC changes over time is not fully understood. For example, tumor ADC values may initially increase as tumor cells die and then decrease as healthy tissue replaces it.

Cardiac, respiratory, and bulk subject motion during diffusion-encoding can introduce artifacts into the images and lead to inaccurate ADC measurements. To avoid this, a fast imaging technique called echo planar imaging (EPI) [47] is typically employed. Unfortunately, EPI is prone to severe image distortions at the interface of materials with different magnetic susceptibilities, such as, the air/tissue and adipose/glandular tissue interfaces in the breast and chest wall. Alternative fast image acquisition schemes, such as fast spin echo (FSE) [48], which are not as susceptible to these distortions can be used to prevent them. However, these alternative acquisition schemes are more prone to motion-induced artifacts and often require the collection of additional images to correct the motion-induced artifacts. Susceptibility-induced distortions can be corrected using image post-processing techniques; however, this requires the acquisition of an accurate map of the main magnetic field [49].

Subject motion between the acquisition of slices or different diffusion-weightings can cause misregistration between the nondiffusion-weighted and diffusion-weighted images,

leading to errors in the ADC calculation. Multiple repetitions are often acquired and averaged to increase signal to noise, and misregistration between the individual images will also lead to inaccurate ADC calculations. Slice-based affine registration of the individual diffusion-weighted images to the nondiffusion-weighted image is typically used in DW-MRI studies of the brain. Until recently, however, these techniques have not been used in DW-MRI studies of the breast [50].

Although guidelines have been published [51], little consensus exists yet within the field regarding the optimum set of b-values that should be used in DW-MRI of breast cancer. Blood flow in the capillaries causes an artificial decrease in signal at very low b-values (e.g., 0–100 s/mm^2). This effect is related to the perfusion fraction, which is the fraction of total water located within the capillaries in the tissue [52]. Tumors typically exhibit a broad range of perfusion factors (20%–80% [53, 54]), increasing the effects of perfusion on the signal measured at low b-values. Therefore, values of 50–100 s/mm^2 are typically used instead of $b = 0$ s/mm^2 in cancer applications. At higher b-values (e.g., ≥ 3000 s/mm^2), there is greater signal attenuation due to diffusion and the measured signal may appear to be artificially increased by rectified noise in magnitude images as the noise floor is reached. The rate at which the noise floor is approached is related to the ADC value; the higher the ADC value, the faster the signal is attenuated at a lower b-value. Therefore, the optimal high b-value for estimation of ADC values is related to the ADC value of the tissue being probed. A wide range of high b-values were used in the studies described in Section 4.2 (300–1000 s/mm^2). The use of multiple b-values can mitigate this confusion but requires additional images and therefore takes more time.

In addition to the technical challenges of DW-MRI acquisition, there exist challenges related to the analysis methods used in longitudinal studies of ADC. The most common method is region-of-interest (ROI) analysis, where ADC values are averaged over a manually defined set of voxels within the tumor. These mean ADC values are then compared between measurements. This method is straightforward and relatively easy to implement. However, it is prone to partial volume effects and inter- and intrauser errors in ROI selection. Voxelwise statistical comparisons may provide a better means of analyzing tumors, particularly if the tumors exhibit heterogeneous ADC values. Ma et al. [55] recently applied a Voxelwise analysis method called "functional diffusion mapping" (fDM) in a small study evaluating early treatment response in breast cancer. Voxel-based analyses require coregistration of images over time, and there are several challenges to doing this, which will be described in more detail in the next section.

5. Longitudinal Registration of Breast Images

5.1. Fundamentals of Image Registration. Image registration is the process of calculating a transformation to align one image to a target image. It has been widely used in the field of medical image analysis to, for example, integrate complementary information from multiple modalities, estimate the variations in anatomical structures within a population, follow diseases over time, or detect tumor response to therapy.

Two key elements are involved in image registration: calculation of a function which measures the similarity between two images such as their mutual information or a cost function that quantifies differences, and a transformation which maps corresponding points of two images. The goal of registration is to find an optimum transformation which matches two images maximally through minimizing/maximizing the cost/similarity function. The transformation is usually categorized as rigid, affine, or nonrigid. Only rotation and translation are involved in rigid body transformation, while there are additional scaling and shearing parameters in affine transformation. Nonrigid transformation is used to incorporate additional local elastic deformations. In general, nonrigid deformation fields can be modeled by a linear combination of basis functions which can be expressed in various ways [56–62], cubic B-splines [59, 60], or Wu's compactly supported positive definite radial basis functions [61, 62].

Parametric registration methods can be classified as point-based, surface-based, or intensity-based. Point-based registration algorithms require identifying and localizing fiducial points in two images. Fiducial points can be markers placed in images intentionally for easy identification or salient points in anatomical structures. The cost function could be the distance between two sets of points (e.g., for rigid body registration) or with additional constraints (e.g., a penalty term [63] to regularize the transformation in nonrigid registration). The rigid or affine transformation parameters or the coefficients of basis functions need to be calculated through optimizing the cost function during registration.

Surface representation needs to be determined first and then registered in surface-based registration methods. The representation includes salient points, boundaries or curves of anatomical structures, or regions with homogeneous intensities. Various surface-based registration algorithms have been developed. The iterative closest point (ICP) [64] is one of the most robust methods applied widely for the rigid body registration of medical images. This algorithm minimizes the differences between two sets of points and calculates the transformation iteratively. The robust point matching (RPM) algorithm [65] is another approach commonly used in nonrigid registration tasks. Instead of generating a one-to-one correspondence between two sets of points as in ICP, RPM builds a "fuzzy" correspondence matrix whose elements are weights assigned to each point in one image, corresponding to each point in the other image. A virtual correspondence is then calculated based on the matrix and the TPS based transformation is obtained iteratively.

Intensity-based registration methods employ the image intensity of each pixel (or voxel) to conduct the alignment, thus containing more comprehensive information than point- and surface-based registration. Among numerous cost or similarity functions, mutual information (MI) and normalized mutual information (NMI) [66–68] are the most common approaches in medical image registration [69–72].

MI measures the statistical dependence between two images, instead of using the image intensity directly. NMI is less sensitive to the overlap size between two images than MI [68]. MI or NMI has been employed widely as a successful similarity function and can be extended to solve more registration problems.

The MI-based rigid registration algorithm has been proposed to calculate rigid body transformation through maximizing MI using Powell's optimization method [71]. One of the representative MI-based nonrigid registration algorithms [73] combines a global transformation (rigid or affine) and a local transformation (the B-spline-based free form deformation) to model the deformation field. The cost function is composed of a similarity term (NMI) and a smoothness term (the bending energy [63]) and optimized using the gradient descent technique. The method was applied to breast MR images and then evaluated by a finite element method- (FEM-) based validation algorithm [74]. Another popular MI-based nonrigid registration algorithm is the adaptive bases algorithm (ABA) [61]. Wu's [62] radial basis functions are compactly supported and employed to compute the transformation. The coefficients of basis functions are searched using a gradient descent algorithm combined with a line minimization algorithm.

In addition to parametric registration algorithms, medical images can also be aligned using nonparametric algorithms, such as the hierarchical attribute matching mechanism for elastic registration technique [75], Maxwell's demons-based registration [76], or fluid registration [77]. Those techniques compute the entire deformation field without using basis functions.

5.2. Review of the Literature. Although RECIST is the most widely used approach to measure tumor response, it has well-recognized deficiencies including discarding important tumor spatial information. This is also true of methods which report the mean, skewness, or kurtosis of, for example, a K^{trans} distribution. The registration of longitudinal breast cancer images is an alternative approach which may allow for quantitative assessment of changes in tumor characteristics on a voxel-by-voxel basis. However, there are only a few investigations regarding the registration of longitudinal breast images, although there are a number of methods and techniques which have been developed to register a variety of imaging data. Registering breast images presents a number of difficulties that are not necessarily encountered in the registration of other organs. The challenges include bulk breast motion and deformation caused by patient repositioning, in addition to any changes in the tumor induced by therapy. The most challenging problem is that common registration techniques will compress the breast tumor imaged before treatment to match the tumor shape observed after treatment because the tumor shape typically changes and the tumor size decreases posttreatment. This kind of compression is misleading in regards to tumor response. Hence, any longitudinal breast image registration technique must be able to accommodate changes within the breast without forcing posttreatment tumor shape and volume to match pretreatment features.

Chittineni et al. [78] presented a work which used a B-spline-based method to register longitudinal breast MR images. To maintain tumor volumes, a rigidity constraint was imposed on the nonrigid registration. Specifically, the B-spline control points that cover the tumor were forced to remain equidistant during the registration process.

Li et al. [79] proposed another registration method to align breast MR images obtained at different time points throughout the course of chemotherapy. The adaptive basis algorithm (ABA), which can accommodate nonrigid coregistrations, was extended through incorporating a volume-preserving constraint [80] into the MI-based cost function. The algorithm was used to register high-resolution T_1-weighted MR images obtained pretreatment, post one cycle of treatment, and after all cycles of treatment. The generated transformations were then applied to longitudinal parametric maps obtained from breast DCE-MRI data. The alignment of the parameters which were related to tumor characteristics allowed a voxel-by-voxel-based analysis of tumor response. Li et al. [81] also developed a realistic phantom based on biological and clinical observations to quantitatively and comprehensively evaluate the proposed registration algorithm. The validation experiments demonstrated that the proposed method was highly accurate. Figure 3 presents an example of the performance of the algorithm. The T_1-weighted MR image before neoadjuvant chemotherapy (panel a) is registered to the image posttreatment (panel d) using the constrained algorithm and the original ABA algorithm, respectively. The registered image obtained with the constrained algorithm (panel b) shows the tumor is preserved successfully, while the original ABA algorithm compresses the tumor substantially (panel c).

As mentioned above, Ma et al. [55] conducted voxel-by-voxel analyses on ADC maps to evaluate breast cancer treatment through registering diffusion weighted MRI scans obtained pretreatment, early posttreatment, and late posttreatment. The registration algorithm [82] used by them required the user selection of control points. The TPS transformation mapping two sets of control points was calculated and applied to the images and the MI between two images was then evaluated. The MI-based cost function was minimized through automatically adjusting the locations of control points in the optimization process and recalculating the TPS transformation and the MI iteratively. The voxel-based analysis showed that ADC changes were able to predict early breast tumor response to treatment.

Though these are exciting and promising early efforts, we stress that longitudinal registration of breast images obtained at different time points is still very much experimental and much work is left to be done.

5.3. Research Opportunities in Longitudinal Registration of Breast Images. The paucity of studies relating to longitudinal registration of breast images leaves many research opportunities. All current investigations rely on intensity-based registration algorithms, neglecting the inherent physical properties of tumor and healthy tissues. One potential approach is to incorporate breast tissue characteristics into the registration process. Schnabel et al. [74] constructed
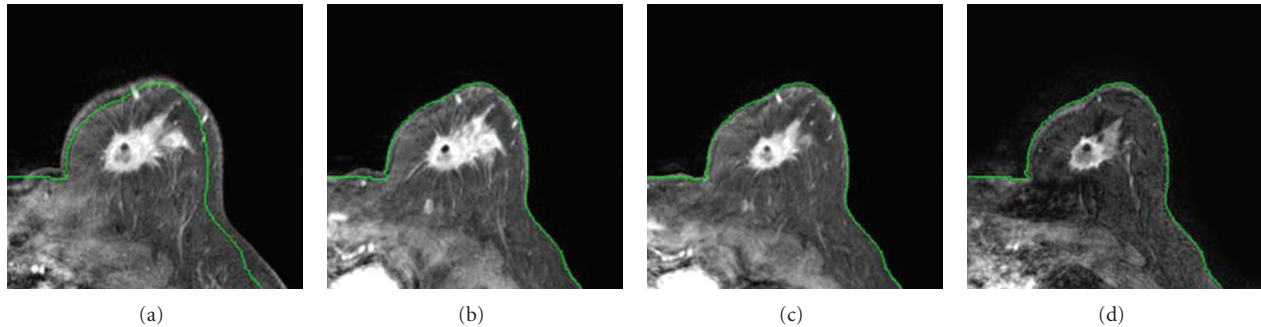
(a) (b) (c) (d)

Figure 3: The T_1-weighted MR image before neoadjuvant chemotherapy (a) is registered to the image posttreatment (d) using the constrained ABA algorithm and the original ABA algorithm, respectively. The registered image using the constrained algorithm (b) shows that the tumor is preserved successfully, while the original ABA algorithm compresses the tumor substantially (c).

an FEM model which simulated biomechanical tissue deformations to validate a nonrigid registration algorithm, but the model was applied to pre- and postcontrast breast MR images, instead of pre- and posttreatment data. In addition to the challenging problems described above in longitudinal breast image registration, it is also difficult to measure tumor and tissue mechanical features on an individual basis and build patient-specific models. Furthermore, the breast tumor generally has different material properties before and after treatment. Solving those difficulties may improve the performance of registration substantially. Another potential opportunity is the use of physiological DCE-MRI parameters during registration. Integrating the parameters related to tumor features into the registration will possibly improve the performance of registration.

6. Biomedical Informatics Systems to Support Image-Based Treatment Response Assessment

Current approaches for applying and developing response criteria have significant limitations. In this section, we describe the current approaches and systems used to apply and develop response criteria, describe the limitations of these approaches, existing systems that may offer solutions, and additional research opportunities.

6.1. Limitations of Current Approaches to Applying Response Criteria. There are several challenges with the current approach to manage clinical trial response data and apply response criteria. First, the raw imaging data are not routinely stored as part of clinical trial data management systems. The DICOM (digital imaging and communications in medicine) image files are often stored on CDs that can be shipped to central sites for review or maintained on the local institution's PACS (picture archiving and communications system). As a result, many large multicenter clinical trials often have incomplete image data sets when it is time to perform a secondary central review. Second, the current radiologist-oncologist communication paradigm is not optimally coordinated with respect to tracking target lesions of interest, resulting in incomplete information for evaluating

tumor response status being recorded in the medical record [83]. Current reporting of radiology results in the medical record focuses on summarizing findings, which is insufficient for consistent application of quantitative methods to evaluate response to treatment.

Furthermore, the use of manually transcribed records of response has several disadvantages including transcription errors, ambiguity of lesion identifiers when multiple lesions are present in the same organ or image, lack of a direct link to the source image data, and limited functionality for calculating response. Response criteria contain multiple calculation and classification rules that are often inconsistently applied within and across studies. This makes it difficult to compare trial response outcomes between cohort arms within the same study and between different trials within the same clinical domain. Automated interpretation methods are needed to improve the consistency of response classification. However, manual data collection in unstructured data formats does not facilitate automated data interpretation. The lack of a direct link between the recorded lesion values and the raw image data source also makes it difficult for outside reviewers to audit the response assessment interpretation results. Finally, there is often a delay in formal application of response criteria to assist treatment decision making and the response data are not integrated into the electronic medical record in such a way that it can readily be used at the point of care by the treating oncologists. Clinical and clinical research systems are needed to (1) support the transfer, storage, and retrieval of clinical and clinical research DICOM files, (2) manage the workflow for interpretation of image data, (3) acquire, store, and retrieve image interpretation meta-data, (4) interpret image findings for response assessment, and (5) enable review of image meta-data and their interpretations with other clinical data as needed for treatment decision making and the aggregation and auditing of response outcomes in clinical trials.

6.2. Limitations of Current Approaches to Development of Response Criteria. Several European and American research centers are involved in the development and testing of new oncology response assessment criteria. For the recent update of the RECIST criteria to version 1.1, a large retrospective

database of target lesion measurements was developed to test the impact of modifications to the criteria [84]. Meta-data on 18,000 potential target lesions were obtained from 6512 patients in 16 metastatic cancer clinical trials. The database was used to evaluate the impact of changes to RECIST on the classification of patient response to treatment [85]. The RECIST 1.1 criterion was thus validated only by comparing it to the previous standard approach and not by evaluating its correlation with survival endpoints.

The ongoing multicenter I-SPY II (investigation of serial studies to predict your therapeutic response with imaging and molecular analysis 2) neoadjuvant breast cancer clinical trial [86] is attempting to address some of these issues by evaluating if breast MRI response predicts treatment outcomes. The NCI's cancer Biomedical Informatics Grid (caBIG) (https://cabig.nci.nih.gov/) program is supporting the I-SPY trial through integration and analysis of diverse data types including clinical, MRI imaging, and tumor biomarkers throughout the breast cancer treatment cycle. The integrated platform is designed to enable correlation of molecular data with MRI patterns to identify surrogate markers for early treatment response.

Despite these recent advances in the development of large databases to conduct research on response criteria, several limitations remain. First, these data repositories currently do not use data standards for sharing imaging metadata across institutions so as to enable the collection of large imaging data sets for research. Similarly, there is a lack of publicly available data sets containing baseline and follow-up imaging studies and image annotations, along with the corresponding diagnoses, therapies, and clinical outcomes. A publicly available database containing data from multiple institutions could be used to evaluate new image processing algorithms and response criteria.

6.3. Biomedical Informatics Systems to Support Response Assessment. Several biomedical informatics systems have been developed that could support the application and development of response criteria. These include systems to support (1) the creation of image repositories, (2) the acquisition, transfer and storage of image meta-data, (3) the automated interpretation of response data, and (4) the visualization of image meta-data for treatment decision making.

Several systems have been developed to support the creation of image repositories for clinical research. The National Biomedical Imaging Archive (NBIA) (https://cabig.nci.nih.gov/tools/NCIA) provides web-based access to deidentified DICOM images, markups, and annotations using role-based security. Publically available image collections available though the NBIA include the I-SPY trial and the Reference Image Database to Evaluate Therapy Response (RIDER) (https://wiki.nci.nih.gov/display/CIP/RIDER) which includes serial DCE-MRI studies of the breast. These publically available image data sets are available *via* the NCI hosted NBIA, but institutions can adapt the open source NBIA software for data storage by setting up a local instance of NBIA at the institution.

The American College of Radiology Imaging Network (ACRIN) (http://www.acrin.org/) has developed the TRIAD (https://triad.acr.org) system to support the management of large image repositories for clinical research. The TRIAD web client allows users to search, download, and view available DICOM image series, while the graphical user interface client provides advanced functions for DICOM series routing, image processing, and annotation layer management. TRIAD is used to manage the image data for all of ACRIN's clinical trials but remains proprietary software.

In addition to advances in the management of clinical trial image data, several more recent advances have emerged in the management of image meta-data. Image meta-data refers to data that describe the content of images. Another initiative of the caBIG program was the development of the image Physician Annotation Device (iPad) [87]. iPad is a plugin to the popular open source OsiriX [88] image viewing application and implements the Annotation and Image Markup (AIM) standard [89]. AIM provides the data structure for storing the key semantic lesion information. "Semantics" refers to meaning in images including image metadata pertinent to quantitative criteria of disease, such as, lesion identification, location, size measurements, method of measurement, and other quantitative features. AIM utilizes the RadLex [90] controlled terminology for describing the contents of medical images, and provides a standard information model for semantic annotations. Information about image annotations is recorded in AIM as XML compliant with the AIM schema, enabling the consistent representation, storage, and transfer of the semantic meaning of imaging features.

The use of iPad has been recently evaluated for tumor lesion tracking and semantic annotation of image meta-data to automatically populate lesion flow sheets for several cancer clinical trials [91]. The use of a semantic image annotation tool to directly populate lesion flow sheets provides several advantages. First, the use of an information model and structured terminology to encode cancer lesion image findings provides a consistent representation for storage and sharing of image meta-data needed for cancer clinical trials. The information model also provides a foundation to enable reasoning with and querying over image meta-data for response assessment. Image annotation disambiguates lesion identifiers by linking them directly to the source image and image mark-up. Image annotation tools also directly generate meta-data about image mark-up such as, length calculations eliminating possible transcription errors for length measurements.

Automated response assessment methods have also been developed to calculate RECIST variables such as, sum of diameters and the percent change from baseline. For example, a semantic reasoning method for response assessment utilizing AIM data was recently developed [92]. The method evaluates both the qualitative and quantitative features of tumor lesion annotations to calculate and classify treatment response according to the RECIST criteria. A vendor solution developed by MEDIAN technologies (http://www.mediantechnologies.com/) includes systems to measure and track cancer lesions and calculate quantitative

RECIST metrics. Use of such automated approaches could improve the consistency and ease with which response criteria are applied in clinical research and routine care.

The abstraction and visualization of cancer lesion meta-data have also been demonstrated and integrated into the clinical work flow of treating medical oncologists [93]. Such visualizations enable rapid and accurate assessments of the image-based response data for use in treatment decision-making. Additional work however, is needed to incorporate image and response meta-data into the electronic health record and to integrate them with other clinical data such as treatments administered and laboratory response biomarkers.

6.4. Research Opportunities Bioinformatics Systems. We have described above several existing and emerging informatics systems that manage various aspects of image-based cancer treatment response assessment. However, several biomedical informatics research challenges remain to be solved to support application and development of response criteria for breast cancer. One of the greatest challenges remains integration of these various systems to support end-to-end data acquisition, interpretation and visualization for clinical trial data management and treatment decision support. Figure 4 shows an informatics architecture integrating multiple data acquisition and visualization systems with their respective data storage systems. Automated response calculators also need to be integrated into systems to support more consistent application of response criteria.

Research systems are also needed to better support the development of response criteria for novel imaging techniques such as DW-MRI and DCE-MRI of the breast. Large de-identified data sets are needed of image and image meta-data linked to patient treatments and clinical outcomes. While the iSPY trial presents an interesting pilot of such a system, larger data sets will likely be required to validate new response assessment modalities and criteria.

7. Breast MRI at 3T

For most of the past 20 years, human body imaging has been performed mainly at 1.5T. In recent years, clinical scanners operating at the higher field of 3T have become routinely available for clinical studies. To date, 1.5T is still the main field for studies of the breast.

7.1. Motivation for Higher Field Imaging. All magnetic resonance imaging techniques are constrained by a compromise between temporal and spatial resolution and signal-to-noise ratio (SNR). The fundamental motivations for high-field breast imaging are (1) a higher SNR, which provides gains that can then be kept or traded for higher spatial and/or temporal resolution, and (2) spectral spreading, which improves the resolution of magnetic resonance spectroscopy (discussed below in Section 7.3) and can improve fat suppression. These improvements can, potentially, lead to shortened acquisition times or increasing sensitivity and specificity for DW-MRI and DCE-MRI and, ultimately,
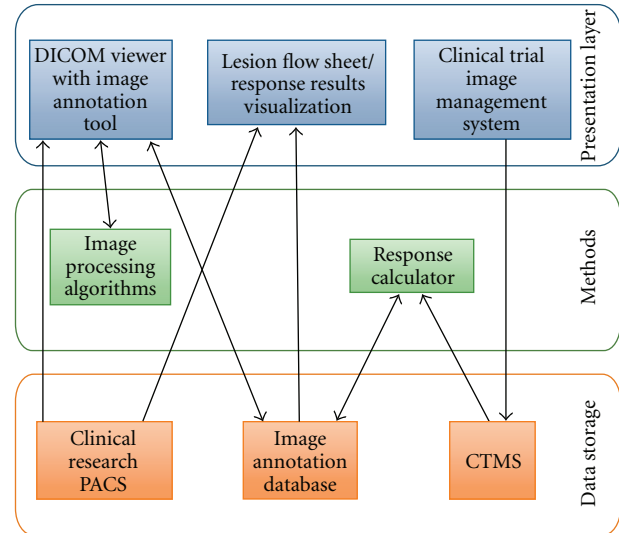

FIGURE 4: Biomedical informatics system architecture to support cancer treatment response assessment. Integrated systems are needed to support end-to-end management of response assessment data including data acquisition, analysis, and visualization. Data acquisition includes systems to acquire and store clinical images for research as well as systems to create image annotation data whether manual image annotation or automated image processing algorithms. Response calculation methods also need to be integrated with structured image annotation databases to perform data analysis to generate response interpretations. The images, image annotations, and response interpretations must be integrated with Clinical Trial Management Systems (CTMS) to enable visualization of response interpretations for treatment decision-making and auditing of clinical trial image response data.

better patient care. Unfortunately, direct comparison studies between 1.5 and 3T are difficult to execute due to the typically short pre-operative window in breast cancer and the requirement for a patient to make multiple trips to an imaging site because separate studies must be performed on separate days to eliminate image contamination by previous contrast material [94]. Thus, only a few studies have performed such a direct comparison, but they suggest that 3T is at least equivalent, and often better, than 1.5T.

Elsamaloty et al. [95] retrospectively examined sensitivity, specificity, and predictive values of breast MRI in 434 women and found MRI at 3T to be more sensitive than mammography and sonography (100% versus 81.8% and 86.4%, resp.) in detection of breast cancer and characterization of lesions smaller than 4 mm. They also found that 3T MRI had a higher sensitivity than lower-field MRI (100% versus 71%–94.4% reported for low-field studies) with no significant difference in specificity. Schmidt et al. [96] compared FDG-PET-CT to whole-body MRI at 1.5T and 3T and found that whole-body MRI using either field strength gave excellent sensitivity (93%) in detecting distant metastatic disease in breast cancer patients, but was less accurate at identifying involved lymph nodes (only 16 of 21 identified). Overall, they found "comparably good performance on both scanners with only slightly increased

artifacts at 3T, without significantly restraining influence on image quality" and a shorter scan time at 3T (43 versus 51 min). An interesting example of a 3T breast protocol was offered by Pinker et al. [97]. Noting the conflicting demands of high temporal resolution imaging (required for quantitative DCE analysis) and high spatial resolution (required for clinical reading), the authors designed an imaging approach that alternates between a T_1-weighted volume-interpolated-breathhold-examination (VIBE) sequence, for high temporal resolution data, and a T_1-weighted fast-low-angle-shot (FLASH) sequence with fat suppression for morphologic analysis.

A final 3T versus 1.5T study we will discuss compared contrast-enhanced imaging of the breast in the same patient set at the two field strengths [98]. The goal of the study was to prospectively compare the the image quality and diagnostic performance of the data obtained at 3T and 1.5T. Image quality was (visually) scored from one to five, with one being nondiagnostic and five being excellent. Differential diagnosis was assessed by the receiver operator characteristic (ROC) analysis. Overall image quality was judged to be higher at 3T; 3T received an average score of 4.6 ± 0.5, while 1.5T received an average score of 4.2 ± 0.5 ($P < .01$). This difference in image quality translated into an improvement in diagnostic performance; an ROC score of 0.98 at 3T compared to. 0.84 at 1.5T ($P < .05$). The authors were careful to state that this result should not be read as an indication that contrast-enhanced breast MRI at 3T is more sensitive or specific than at 1.5T, but they did conclude that contrast-enhanced MRI at 3T is beginning to be ready for routine clinical use.

7.2. Challenges for Breast Imaging at 3T. Kuhl reviewed several issues involved with moving clinical breast imaging to 3T and stressed the need for maximum SNR [94]. Soher et al. provide an excellent review of relevant physics involved in going from 1.5 to 3T [99]. Major barriers to 3T imaging of the breast include increased B_0 and B_1 inhomogeneity (due to shorter RF wavelengths), higher specific absorption rate (SAR), changes in relaxation rate (lengthening T_1 and shortening T_2, which can adversely affect tissue contrast) changes, chemical shift artifacts, susceptibility gradients, and medical device safety. Some of the implications of these changes are described in more detail below.

As B_0 increases, the precession frequencies of nuclear magnets increase, in turn reducing the efficiency of spin-lattice relaxation processes and a longer persistence of the longitudinal magnetization. This translates to an increased T_1 relaxation time at 3T. In order to preserve T_1 contrast, a longer T_1 would require longer repetition time (TR) and/or longer inversion delay values which may lead to lower temporal resolution and/or increased scan times. T_1 values for both adipose and fibroglandular tissue have been shown to increase by approximately 20% in going from 1.5T to 3T [100]. This has a proportional effect on the inversion times used for fat suppression and provides a good example of sequence parameters that must change in going to 3T. Additionally, since the relaxivity of clinically accepted contrast agents decreases with field strength, dosing for DCE-MRI may need to be modified from the 1.5T values to achieve comparable contrast at 3T [101].

SAR scales with the square of the field strength, so all other parameters being equal, the SAR level at 3T is expected to be four times that at 1.5T. Increased SAR may require modification of pulse sequences, such as smaller flip angles, longer echo times, or more data (i.e., k-space) lines acquired per excitation, to reduce RF power deposition to 1.5T levels, thus compromising some of the SNR and temporal resolution gains available at 3T.

Increased B_1 inhomogeneity due to dielectric effects can cause spatial variations in SNR. The sensitivity of the overall scan is limited by the lowest sensitivity of any voxel in the volume so the overall scan sensitivity suffers when there are significant areas of low B_1. Azlan et al. have quantified B_1 inhomogeneity across the breast in normal volunteers at 3T [102]. They find a median 40% reduction of the expected B_1 in the right breast on T_1-weighted spoiled gradient echo images.

7.3. Research Opportunities for Breast Imaging at 3T. Maintaining or improving image characteristics such as, contrast, resolution, and SNR at 3T will require not only modifications to conventional sequences but also improved hardware designs. Larger multichannel coil arrays will help maintain existing scan times with increased parallel acceleration factors while maintaining comparable SNR [103].

The stronger background magnetic field produces a wider separation of fat and water frequencies, though this improvement in spectral resolution is somewhat balanced by increased susceptibility effects that broaden the lines. Assuming that the line broadening does not completely nullify the wider separation of the spectral lines, fat suppression should be more complete at 3T. This gain, plus the possibility of increased time resolution, could improve fat-suppressed DW-MRI (important for limiting artifacts) as well as DCE-MRI studies [104]. However, B_0 and B_1 inhomogeneity make successful fat suppression difficult in practice. MR spectroscopy (MRS) is also hampered by the limited SNR available at 1.5T and could theoretically benefit from the increased SNR at 3T [105]. Many groups have employed MRS for measuring the response of tumors to treatment and the interested reader is referred to the excellent recent reviews in references [106, 107]. In the context of breast cancer, the MRS-visible molecule that has generated the most interest is choline. Since choline containing compounds are implicated in synthesis and metabolism of cell membranes, the standard interpretation of increases in total choline signal is increased cell turnover. However, at lower fields measuring this signal can be quite difficult and time consuming which can be a burden on the patient. It has been established that the sensitivity of proton MRS scales roughly linearly with field strength [108, 109]. Also, the increased spectral spreading with higher field strength leads to the separation of previously overlapping spectral peaks and ultimately the ability to resolve and quantify more compounds. Thus, the use of higher magnetic fields offers the opportunity for more precise and shorter MRS studies.

Finally, medical devices or implants deemed MR compatible at 1.5T may not be at 3T, or they may more markedly degrade the image quality as field strength increases [105]. However, Peters et al. concluded that MRI-guided large-core needle breast biopsy is safe and effective at 3T, despite fears over increased susceptibility artifacts from the biopsy needle [110]. More studies of device safety and effects on image quality must be done for the range of devices currently approved for 1.5T use.

8. Summary

In this paper we have highlighted several areas of quantitative magnetic resonance imaging that will have to be advanced in order to report with high sensitivity and specificity on the response of breast tumors to treatment. While the current state of the art in quantitative breast MRI is not quite ready for clinical acceptance, we believe that future developments in each of these areas will catalyze the use of more advanced MR techniques for breast cancer imaging. In fact, there are a number of organizations designed to assist in these efforts by standardizing various acquisition and analysis protocol; these include QIBA (Quantitative Imaging Biomarkers Alliance) and UPICT (Uniform Protocols for Imaging in Clinical Trials). There is also a recent NCI initiative to encourage the incorporation of emerging quantitative imaging in clinical trials (and, therefore, larger controlled patient populations) through the Quantitative Imaging Network (QIN).

Acknowledgments

References

[1] M. Kaufmann, G. von Minckwitz, R. Smith et al., "International expert panel on the use of primary (preoperative) systemic treatment of operable breast cancer: review and recommendations," *Journal of Clinical Oncology*, vol. 21, no. 13, pp. 2600–2608, 2003.

[2] B. Fisher, J. Bryant, N. Wolmark et al., "Effect of preoperative chemotherapy on the outcome of women with operable breast cancer," *Journal of Clinical Oncology*, vol. 16, no. 8, pp. 2672–2685, 1998.

[3] B. Fisher, A. Brown, E. Mamounas et al., "Effect of preoperative chemotherapy on local-regional disease in women with operable breast cancer: findings from national surgical adjuvant breast and bowel project B-18," *Journal of Clinical Oncology*, vol. 15, no. 7, pp. 2483–2493, 1997.

[4] T. A. Buchholz, D. W. Davis, D. J. McConkey et al., "Chemotherapy-induced apoptosis and Bcl-2 levels correlate with breast cancer response to chemotherapy," *Cancer Journal*, vol. 9, no. 1, pp. 33–41, 2003.

[5] S. K. Mohsin, H. L. Weiss, M. C. Gutierrez et al., "Neoadjuvant trastuzumab induces apoptosis in primary breast cancers," *Journal of Clinical Oncology*, vol. 23, no. 11, pp. 2460–2468, 2005.

[6] V. Stearns, B. Singh, T. Tsangaris et al., "A prospective randomized pilot study to evaluate predictors of response in serial core biopsies to single agent neoadjuvant doxorubicin or paclitaxel for patients with locally advanced breast cancer," *Clinical Cancer Research*, vol. 9, no. 1, pp. 124–133, 2003.

[7] A. B. Chakravarthy, M. C. Kelley, B. McLaren et al., "Neoadjuvant concurrent paclitaxel and radiation in stage II/III breast cancer," *Clinical Cancer Research*, vol. 12, no. 5, pp. 1570–1576, 2006.

[8] P. Therasse, S. G. Arbuck, E. A. Eisenhauer et al., "New guidelines to evaluate the response to treatment in solid tumors," *Journal of the National Cancer Institute*, vol. 92, no. 3, pp. 205–216, 2000.

[9] E. A. Eisenhauer, P. Therasse, J. Bogaerts et al., "New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1)," *European Journal of Cancer*, vol. 45, no. 2, pp. 228–247, 2009.

[10] A. Alavi and M. Reivich, "Guest editorial: the conception of FDG-PET imaging," *Seminars in Nuclear Medicine*, vol. 32, no. 1, pp. 2–5, 2002.

[11] R. L. Wahl, H. Jacene, Y. Kasamon, and M. A. Lodge, "From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors," *Journal of Nuclear Medicine*, vol. 50, supplement 1, pp. 122S–150S, 2009.

[12] T. E. Yankeelov and J. C. Gore, "Dynamic contrast enhanced magnetic resonance imaging in oncology: theory, data acquisition, analysis, and examples," *Current Medical Imaging Reviews*, vol. 3, no. 2, pp. 91–107, 2007.

[13] T. E. Yankeelov, W. D. Rooney, W. Huang et al., "Evidence for shutter-speed variation in CR bolus-tracking studies of human pathology," *NMR in Biomedicine*, vol. 18, no. 3, pp. 173–185, 2005.

[14] X. Li, W. Huang, T. E. Yankeelov, A. Tudorica, W. D. Rooney, and C. S. Springer Jr., "Shutter-speed analysis of contrast reagent bolus-tracking data: preliminary observations in benign and malignant breast disease," *Magnetic Resonance in Medicine*, vol. 53, no. 3, pp. 724–729, 2005.

[15] K. M. Donahue, R. M. Weisskoff, and D. Burstein, "Water diffusion and exchange as they influence contrast enhancement," *Journal of Magnetic Resonance Imaging*, vol. 7, no. 1, pp. 102–110, 1997.

[16] H. M. McConnell, "Reaction rates by nuclear magnetic resonance," *The Journal of Chemical Physics*, vol. 28, no. 3, pp. 430–431, 1958.

[17] D. E. Woessner, "Nuclear transfer effects in nuclear magnetic resonance pulse experiments," *The Journal of Chemical Physics*, vol. 35, no. 1, pp. 41–48, 1961.

[18] C. S. Landis, X. Li, F. W. Telang et al., "Determination of the MRI contrast agent concentration time course in vivo following bolus injection: effect of equilibrium transcytolemmal water exchange," *Magnetic Resonance in Medicine*, vol. 44, no. 4, pp. 563–574, 2000.

[19] T. E. Yankeelov, J. J. Luci, L. M. DeBusk, P. C. Lin, and J. C. Gore, "Incorporating the effects of transcytolemmal water exchange in a reference region model for DCE-MRI analysis: theory, simulations, and experimental results," *Magnetic Resonance in Medicine*, vol. 59, no. 2, pp. 326–335, 2008.

[20] M. Lowry, B. Zelhof, G. P. Liney, P. Gibbs, M. D. Pickles, and L. W. Turnbull, "Analysis of prostate DCE-MRI: comparison

of fast exchange limit and fast exchange regimen pharmacokinetic models in the discrimination of malignant from normal tissue," *Investigative radiology*, vol. 44, no. 9, pp. 577–584, 2009.

[21] S. Kim, H. Quon, L. A. Loevner et al., "Transcytolemmal water exchange in pharmacokinetic analysis of dynamic contrast-enhanced MRI data in squamous cell carcinoma of the head and neck," *Journal of Magnetic Resonance Imaging*, vol. 26, no. 6, pp. 1607–1617, 2007.

[22] D. L. Buckley, L. E. Kershaw, and G. J. Stanisz, "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," *Magnetic Resonance in Medicine*, vol. 60, no. 5, pp. 1011–1019, 2008.

[23] R. Johansen, L. R. Jensen, J. Rydland et al., "Predicting survival and early clinical response to primary chemotherapy for patients with locally advanced breast cancer using DCE-MRI," *Journal of Magnetic Resonance Imaging*, vol. 29, no. 6, pp. 1300–1307, 2009.

[24] L. Martincich, F. Montemurro, G. De Rosa et al., "Monitoring response to primary chemotherapy in breast cancer using dynamic contrast-enhanced magnetic resonance imaging," *Breast Cancer Research and Treatment*, vol. 83, no. 1, pp. 67–76, 2004.

[25] Y.-C. Cheung, S.-C. Chen, M.-Y. Su et al., "Monitoring the size and response of locally advanced breast cancers to neoadjuvant chemotherapy (weekly paclitaxel and epirubicin) with serial enhanced MRI," *Breast Cancer Research and Treatment*, vol. 78, no. 1, pp. 51–58, 2003.

[26] M.-L. W. Ah-See, A. Makris, N. J. Taylor et al., "Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer," *Clinical Cancer Research*, vol. 14, no. 20, pp. 6580–6589, 2008.

[27] E. S. Paulson and K. M. Schmainda, "Comparison of dynamic susceptibility-weighted contrast-enhanced MR methods: recommendations for measuring relative cerebral blood volume in brain tumors," *Radiology*, vol. 249, no. 2, pp. 601–613, 2008.

[28] H. J. Yu, J.-H. Chen, R. S. Mehta, O. Nalcioglu, and M.-Y. Su, "MRI measurements of tumor size and pharmacokinetic parameters as early predictors of response in breast cancer patients undergoing neoadjuvant anthracycline chemotherapy," *Journal of Magnetic Resonance Imaging*, vol. 26, no. 3, pp. 615–623, 2007.

[29] H.-L. M. Cheng, "Investigation and optimization of parameter accuracy in dynamic contrast-enhanced MRI," *Journal of Magnetic Resonance Imaging*, vol. 28, no. 3, pp. 736–743, 2008.

[30] J. Evelhoch, M. Garwood, D. Vigneron et al., "Expanding the use of magnetic resonance in the assessment of tumor response to therapy: workshop report," *Cancer Research*, vol. 65, no. 16, pp. 7041–7044, 2005.

[31] R. Meng, S. D. Chang, E. C. Jones, S. L. Goldenberg, and P. Kozlowski, "Comparison between population average and experimentally measured arterial input function in predicting biopsy results in prostate cancer," *Academic Radiology*, vol. 17, no. 4, pp. 520–525, 2010.

[32] Y. Wang, W. Huang, D. M. Panicek, L. H. Schwartz, and J. A. Koutcher, "Feasibility of using limited-population-based arterial input function for pharmacokinetic modeling of osteosarcoma dynamic contrast-enhanced MRI data," *Magnetic Resonance in Medicine*, vol. 59, no. 5, pp. 1183–1189, 2008.

[33] C. R. Planey, E. B. Welch, L. Xu et al., "Temporal sampling requirements for reference region modeling of DCE-MRI data in human breast cancer," *Journal of Magnetic Resonance Imaging*, vol. 30, no. 1, pp. 121–134, 2009.

[34] C. Yang, G. S. Karczmar, M. Medved, and W. M. Stadler, "Multiple reference tissue method for contrast agent arterial input function estimation," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1266–1275, 2007.

[35] L. Dougherty, G. Isaac, M. A. Rosen et al., "High frame-rate simultaneous bilateral breast DCE-MRI," *Magnetic Resonance in Medicine*, vol. 57, no. 1, pp. 220–225, 2007.

[36] H. K. Song and L. Dougherty, "Dynamic MRI with projection reconstruction and KWIC processing for simultaneous high spatial and temporal resolution," *Magnetic Resonance in Medicine*, vol. 52, no. 4, pp. 815–824, 2004.

[37] A. Einstein, "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen," *Annalen der Physik*, vol. 322, pp. 549–560, 1905.

[38] D. Le Bihan, E. Breton, and D. Lallemand, "MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders," *Radiology*, vol. 161, no. 2, pp. 401–407, 1986.

[39] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient," *The Journal of Chemical Physics*, vol. 42, no. 1, pp. 288–292, 1965.

[40] T. Sugahara, Y. Korogi, M. Kochi et al., "Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas," *Journal of Magnetic Resonance Imaging*, vol. 9, no. 1, pp. 53–60, 1999.

[41] K. M. Gauvain, R. C. McKinstry, P. Mukherjee et al., "Evaluating pediatric brain tumor cellularity with diffusion-tensor imaging," *American Journal of Roentgenology*, vol. 177, no. 2, pp. 449–454, 2001.

[42] K. Kono, Y. Inoue, K. Nakayama et al., "The role of diffusion-weighted imaging in patients with brain tumors," *American Journal of Neuroradiology*, vol. 22, no. 6, pp. 1081–1088, 2001.

[43] J.-P. Galons, M. I. Altbach, G. D. Paine-Murrieta, C. W. Taylor, and R. J. Gillies, "Early increases in breast tumor xenograft water mobility in response to paclitaxel therapy detected by non-invasive diffusion magnetic resonance imaging," *Neoplasia*, vol. 1, no. 2, pp. 113–117, 1999.

[44] T. E. Yankeelov, M. Lepage, A. Chakravarthy et al., "Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: initial results," *Magnetic Resonance Imaging*, vol. 25, no. 1, pp. 1–13, 2007.

[45] M. D. Pickles, P. Gibbs, M. Lowry, and L. W. Turnbull, "Diffusion changes precede size reduction in neoadjuvant treatment of breast cancer," *Magnetic Resonance Imaging*, vol. 24, no. 7, pp. 843–847, 2006.

[46] U. Sharma, K. K. A. Danishad, V. Seenu, and N. R. Jagannathan, "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 in Biomedicine*, vol. 22, no. 1, pp. 104–113, 2009.

[47] P. Mansfield, "Multi-planar image formation using NMR spin echoes," *Journal of Physics C*, vol. 10, no. 3, pp. L55–L58, 1977.

[48] J. Hennig, A. Nauerth, and H. Friedburg, "RARE imaging: a fast imaging method for clinical MR," *Magnetic Resonance in Medicine*, vol. 3, no. 6, pp. 823–833, 1986.

[49] P. Jezzard, A. S. Barnett, and C. Pierpaoli, "Characterization of and correction for eddy current artifacts in echo planar diffusion imaging," *Magnetic Resonance in Medicine*, vol. 39, no. 5, pp. 801–812, 1998.

[50] S. C. Partridge, R. S. Murthy, A. Ziadloo, S. W. White, K. H. Allison, and C. D. Lehman, "Diffusion tensor magnetic resonance imaging of the normal breast," *Magnetic Resonance Imaging*, vol. 28, no. 3, pp. 320–328, 2010.

[51] A. R. Padhani, G. Liu, D. Mu-Koh et al., "Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations," *Neoplasia*, vol. 11, no. 2, pp. 102–125, 2009.

[52] D. Le Bihan, E. Breton, D. Lallemand, M.-L. Aubin, J. Vignaud, and M. Laval-Jeantet, "Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging," *Radiology*, vol. 168, no. 2, pp. 497–505, 1988.

[53] H. J. J. A. Bernsen, P. F. J. W. Rijken, T. Oostendorp, and A. J. Van der Kogel, "Vascularity and perfusion of human gliomas xenografted in the athymic nude mouse," *British Journal of Cancer*, vol. 71, no. 4, pp. 721–726, 1995.

[54] S. F. Riches, K. Hawtin, E. M. Charles-Edwards, and N. M. de Souza, "Diffusion-weighted imaging of the prostate and rectal wall: comparison of biexponential and monoexponential modelled diffusion and associated perfusion coefficients," *NMR in Biomedicine*, vol. 22, no. 3, pp. 318–325, 2009.

[55] B. Ma, C. R. Meyer, M. D. Pickles et al., "Voxel-by-voxel functional diffusion mapping for early evaluation of breast cancer treatment," in *Proceedings of the 21st International Conference on Information Processing in Medical Imaging (IPMI '09)*, vol. 5636, pp. 276–287, Williamsburg, Va, USA, July 2009.

[56] F. L. Bookstein, "Principal warps: thin-plate splines and the decomposition of deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. v, no. n, pp. 567–585, 1992.

[57] A. Goshtasby, "Registration of images with geometric distortions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 60–64, 1988.

[58] R. L. Harder and R. N. Desmarais, "Interpolation using surface splines," *Journal of Aircraft*, vol. 9, no. 2, pp. 189–191, 1972.

[59] S. Lee, G. Wolberg, K.-Y. Chwa, and S. Y. Shin, "Image metamorphosis with scattered feature constraints," *IEEE Transactions on Visualization and Computer Graphics*, vol. 2, no. 4, pp. 337–354, 1996.

[60] S. Lee, G. Wolberg, and S. Y. Shin, "Scattered data interpolation with multilevel b-splines," *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 3, pp. 228–244, 1997.

[61] G. K. Rohde, A. Aldroubi, and B. M. Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration," *IEEE Transactions on Medical Imaging*, vol. 22, no. 11, pp. 1470–1479, 2003.

[62] Z. Wu, "Compactly supported positive definite radial functions," *Advances in Computational Mathematics*, vol. 4, no. 1, pp. 283–292, 1995.

[63] G. Wahba, *Spline Models for Observational Data*, Society for Industrial and Applied Mathematics, 1990.

[64] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[65] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pp. 114–141, 2003.

[66] A. Collignon, F. Maes, and D. Delaere, "Automated multimodality medical image registration using information theory," in *Proceedings of the 14th International Conference on Information Processing in Medical Imaging, Computational Imaging and Vision*, vol. 3, pp. 263–274, 1995.

[67] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[68] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, no. 1, pp. 71–86, 1999.

[69] J. West, J. M. Fitzpatrick, M. Y. Wang et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques," *Journal of Computer Assisted Tomography*, vol. 21, no. 4, pp. 554–566, 1997.

[70] W. M. Wells III, P. Viola, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," in *Medical Robotics and Computer Assisted Surgery*, pp. 55–62, Wiley-Liss, New York, NY, USA, 1995.

[71] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997.

[72] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 986–1004, 2003.

[73] D. Rueckert, "Nonrigid registration using free-form deformations: application to breast mr images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[74] J. A. Schnabel, C. Tanner, A. D. Castellano-Smith et al., "Validation of nonrigid image registration using finite-element methods: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 238–247, 2003.

[75] D. Shen and C. Davatzikos, "HAMMER: hierarchical attribute matching mechanism for elastic registration," *IEEE Transactions on Medical Imaging*, vol. 21, no. 11, pp. 1421–1439, 2002.

[76] J.-P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[77] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 1996.

[78] R. Chittineni, M. Y. Su, and O. Nalcioglu, "Breast MR registration for evaluation of Neoadjuvant chemotherapy response," in *Proceedings of the International Society of Magnetic Resonance in Medicine*, vol. 16, p. 3095, 2008.

[79] X. Li, B. M. Dawant, E. B. Welch et al., "A nonrigid registration algorithm for longitudinal breast MR images and the analysis of breast tumor response," *Magnetic Resonance Imaging*, vol. 27, no. 9, pp. 1258–1270, 2009.

[80] T. Rohlfing, C. R. Maurer Jr., D. A. Bluemke, and M. A. Jacobs, "Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint," *IEEE Transactions on Medical Imaging*, vol. 22, no. 6, pp. 730–741, 2003.

Journal of Oncology

17

[81] X. Li, B. M. Dawant, E. B. Welch et al., "Validation of an algorithm for the nonrigid registration of longitudinal breast MR images using realistic phantoms," *Medical Physics*, vol. 37, no. 6, pp. 2541–2552, 2010.

[82] C. R. Meyer, J. L. Boes, B. Kim et al., "Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations," *Medical image analysis*, vol. 1, no. 3, pp. 195–206, 1997.

[83] M. A. Levy and D. L. Rubin, "Tool support to enable evaluation of the clinical response to treatment," in *Proceedings of the American Medical Informatics Aassociation Annual Symposium*, pp. 399–403, 2008.

[84] J. Bogaerts, R. Ford, D. Sargent et al., "Individual patient data analysis to assess modifications to the RECIST criteria," *European Journal of Cancer*, vol. 45, no. 2, pp. 248–260, 2009.

[85] C. S. Moskowitz, X. Jia, L. H. Schwartz, and M. Gönen, "A simulation study to evaluate the impact of the number of lesions measured on response assessment," *European Journal of Cancer*, vol. 45, no. 2, pp. 300–310, 2009.

[86] A. D. Barker, C. C. Sigman, G. J. Kelloff, N. M. Hylton, D. A. Berry, and L. J. Esserman, "I-SPY 2: an adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy," *Clinical Pharmacology and Therapeutics*, vol. 86, no. 1, pp. 97–100, 2009.

[87] D. L. Rubin, C. Rodriguez, P. Shah, and C. Beaulieu, "iPad: Semantic annotation and markup of radiological images," in *Proceedings of the American Medical Informatics Association Annual Symposium*, pp. 626–630, 2008.

[88] A. Rosset, L. Spadola, and O. Ratib, "OsiriX: an open-source software for navigating in multidimensional DICOM images," *Journal of Digital Imaging*, vol. 17, no. 3, pp. 205–216, 2004.

[89] D. S. Channin, P. Mongkolwat, V. Kleper, K. Sepukar, and D. L. Rubin, "The caBIG^{TM} Annotation and Image Markup Project," *Journal of Digital Imaging*, vol. 23, no. 2, pp. 217–225, 2009.

[90] C. P. Langlotz, "RadLex: a new method for indexing online educational materials," *Radiographics*, vol. 26, no. 6, pp. 1595–1597, 2006.

[91] M. A. Levy and D. L. Rubin, "Tool support for cancer lesion tracking and quantitative assessment of disease response," in *Proceedings of the Society for Imaging in Informatics Annual Symposium*, 2010.

[92] M. A. Levy, M. J. O'Connor, and D. L. Rubin, "Semantic reasoning with image annotations for tumor assessment," in *Proceedings of the American Medical Informatics Association Annual Symposium*, vol. 11, pp. 359–363, 2009.

[93] M. A. Levy, A. Garg, A. Tam, Y. Garten, and D. L. Rubin, "LesionViewer: a tool for tracking cancer lesions over time," in *Proceedings of the American Medical Informatics Association Annual Symposium*, pp. 443–447, 2007.

[94] C. K. Kuhl, "Breast MR imaging at 3T," *Magnetic Resonance Imaging Clinics of North America*, vol. 15, no. 3, pp. 315–320, 2007.

[95] H. Elsamaloty, M. S. Elzawawi, S. Mohammad, and N. Herial, "Increasing accuracy of detection of breast cancer with 3-T MRI," *American Journal of Roentgenology*, vol. 192, no. 4, pp. 1142–1148, 2009.

[96] G. P. Schmidt, A. Baur-Melnyk, A. Haug et al., "Comprehensive imaging of tumor recurrence in breast cancer patients using whole-body MRI at 1.5 and 3 T compared to FDG-PET-CT," *European Journal of Radiology*, vol. 65, no. 1, pp. 47–58, 2008.

[97] K. Pinker, G. Grabner, W. Bogner et al., "A combined high temporal and high spatial resolution 3 Tesla MR imaging protocol for the assessment of breast lesions: initial results," *Investigative radiology*, vol. 44, no. 9, pp. 553–558, 2009.

[98] C. K. Kuhl, P. Jost, N. Morakkabati, O. Zivanovic, H. H. Schild, and J. Gieseke, "Contrast-enhanced MR imaging of the breast at 3.0 and 1.5 T in the same patients: initial experience," *Radiology*, vol. 239, no. 3, pp. 666–676, 2006.

[99] B. J. Soher, B. M. Dale, and E. M. Merkle, "A review of MR physics: 3T versus 1.5T," *Magnetic Resonance Imaging Clinics of North America*, vol. 15, no. 3, pp. 277–290, 2007.

[100] R. Rakow-Penner, B. Daniel, H. Yu, A. Sawyer-Glover, and G. H. Glover, "Relaxation times of breast tissue at 1.5T and 3T measured using IDEAL," *Journal of Magnetic Resonance Imaging*, vol. 23, no. 1, pp. 87–91, 2006.

[101] S. Trattnig, K. Pinker, A. Ba-Ssalamah, and I.-M. Nöbauer-Huhmann, "The optimal use of contrast agents at high field MRI," *European Radiology*, vol. 16, no. 6, pp. 1280–1287, 2006.

[102] C. A. Azlan, P. Di Giovanni, T. S. Ahearn, S. I.K. Semple, F. J. Gilbert, and T. W. Redpath, "B1 transmission-field inhomogeneity and enhancement ratio errors in dynamic contrast-enhanced MRI (DCE-MRI) of the breast at 3T," *Journal of Magnetic Resonance Imaging*, vol. 31, no. 1, pp. 234–239, 2010.

[103] E. M. Hecht, R. F. Lee, B. Taouli, and D. K. Sodickson, "Perspectives on body MR imaging at ultrahigh field," *Magnetic Resonance Imaging Clinics of North America*, vol. 15, no. 3, pp. 449–465, 2007.

[104] L. W. Turnbull, "Dynamic contrast-enhanced MRI in the diagnosis and management of breast cancer," *NMR in Biomedicine*, vol. 22, no. 1, pp. 28–39, 2009.

[105] C. K. Kuhl, F. Träber, J. Gieseke et al., "Whole-body high-field-strength (3.0-T) MR imaging in clinical practice: part II. Technical considerations and clinical applications," *Radiology*, vol. 247, no. 1, pp. 16–35, 2008.

[106] I. S. Haddadin, A. Mclntosh, S. Meisamy et al., "Metabolite quantification and high-field MRS in breast cancer," *NMR in Biomedicine*, vol. 22, no. 1, pp. 65–76, 2009.

[107] P. Stanwell and C. Mountford, "In vivo proton MR spectroscopy of the breast," *Radiographics*, vol. 27, pp. S253–S266, 2007.

[108] D. I. Hoult and D. Phil, "Sensitivity and power deposition in a high-field imaging experiment," *Journal of Magnetic Resonance Imaging*, vol. 12, no. 1, pp. 46–67, 2000.

[109] J. T. Vaughan, M. Garwood, C. M. Collins et al., "7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images," *Magnetic Resonance in Medicine*, vol. 46, no. 1, pp. 24–30, 2001.

[110] N. H. G. M. Peters, C. Meeuwis, C. J. G. Bakker et al., "Feasibility of MRI-guided large-core-needle biopsy of suspicious breast lesions at 3T," *European Radiology*, vol. 19, no. 7, pp. 1639–1644, 2009.