HHS Public Access

Author manuscript

J Magn Reson Imaging. Author manuscript; available in PMC 2019 September 23.

Repeatability, Reproducibility, and Accuracy of Quantitative MRI of the Breast in the Community Radiology Setting

Anna G. Sorace, PhD[1,2,3], **Chengyue Wu, BS**[3], **Stephanie L. Barnes, PhD**[3,4], **Angela M. Jarrett, PhD**[4], **Sarah Avery, MD**[5], **Debra Patt, MD**[6], **Boone Goodgame, MD**[7,8], **Jeffery J. Luci, PhD**[3,9], **Hakmook Kang, PhD**[10], **Richard G. Abramson, MD**[11], **Thomas E. Yankeelov, PhD**[1,2,3,4,*], and **John Virostko, PhD**[1,2]

[1]Department of Diagnostic Medicine, University of Texas at Austin, Austin, Texas, USA

[2]Livestrong Cancer Institutes, University of Texas at Austin, Austin, Texas, USA

[3]Department of Biomedical Engineering, University of Texas at Austin, Austin, Texas, USA

[4]Institute for Computational Engineering and Sciences, University of Texas at Austin, Austin, Texas, USA

[5]Austin Radiological Association, Austin, Texas, USA

[6]Texas Oncology, Austin, Texas, USA

[7]Seton Hospital, Austin, Texas, USA

[8]Department of Internal Medicine, University of Texas at Austin, Austin, Texas, USA

[9]Department of Neuroscience, University of Texas at Austin, Austin, Texas, USA

[10]Department of Biostatistics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

[11]Department of Radiology and Radiological Sciences, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Abstract

Background—Quantitative diffusion-weighted MRI (DW-MRI) and dynamic contrast-enhanced MRI (DCE-MRI) have the potential to impact patient care by providing noninvasive biological information in breast cancer.

Purpose/Hypothesis—To quantify the repeatability, reproducibility, and accuracy of apparent diffusion coefficient (ADC) and T_1-mapping of the breast in community radiology practices.

Study Type—Prospective.

Subjects/Phantom—Ice-water DW-MRI and T_1 gel phantoms were used to assess accuracy. Normal subjects ($n = 3$) and phantoms across three sites (one academic, two community) were

*Address reprint requests to: T.E.Y., University of Texas at Austin, Austin, Texas 78712. thomas.yankeelov@utexas.edu.

Additional supporting information may be found in the online version of this article.

Level of Evidence: 1

Technical Efficacy: Stage 2

used to assess reproducibility. Test–retest analysis at one site in normal subjects ($n = 12$) was used to assess repeatability.

Field Strength/Sequence—3T Siemens Skyra MRI quantitative DW-MRI and T_1-mapping.

Assessment—Quantitative DW-MRI and T_1-mapping parametric maps of phantoms and fibroglandular and adipose tissue of the breast.

Statistical Tests—Average values of breast tissue were quantified and Bland-Altman analysis was performed to assess the repeatability of the MRI techniques, while the Friedman test assessed reproducibility.

Results—ADC measurements were reproducible across sites, with an average difference of 1.6% in an ice-water phantom and 7.0% in breast fibroglandular tissue. T_1 measurements in gel phantoms had an average difference of 2.8% across three sites, whereas breast fibroglandular and adipose tissue had 8.4% and 7.5% average differences, respectively. In the repeatability study, we found no bias between first and second scanning sessions ($P = 0.1$). The difference between repeated measurements was independent of the mean for each MRI metric ($P = 0.156$, $P = 0.862$, $P = 0.197$ for ADC, T_1 of fibroglandular tissue, and T_1 of adipose tissue, respectively).

Data Conclusion—Community radiology practices can perform repeatable, reproducible, and accurate quantitative T_1- mapping and DW-MRI. This has the potential to dramatically expand the number of sites that can participate in multisite clinical trials and increase clinical translation of quantitative MRI techniques for cancer response assessment.

Quantitative imaging has matured to the point where it is routinely incorporated in clinical trials of solid tumors in the academic setting, with promising results.[1,2] However, the majority of cancer patients receive their care in the community setting,[3] which consists of private physician-owned businesses or regional hospitals and clinics that are not a part of an academic medical center or medical teaching institution. Thus, to maximize the impact of quantitative imaging it must be implemented in the community setting. This would result in two distinct advances: 1) expand the ability of community sites to participate in multisite clinical trials requiring quantitative magnetic resonance imaging (MRI); and 2) extend the ability for quantitative imaging to be implemented into the standard-of-care setting. Furthermore, performing quantitative MRI in the community setting demonstrates the feasibility for future widespread clinical dissemination. The goal of this study was to implement quantitative MRI in the community setting and characterize the repeatability, reproducibility, and accuracy of two MRI measures appropriate for quantitative characterization of breast cancer. The two measures were the 1) the longitudinal relaxation rate (T_1) for quantitative dynamic contrast-enhanced MRI (DCE-MRI); and 2) the apparent diffusion coefficient (ADC) from diffusion-weighted MRI (DW-MRI).

DW-MRI and DCE-MRI have the potential to characterize important characteristics of the tumor and its microenvironment, including vessel perfusion and permeability, cellularity, and cellular proliferation. The ADC, quantified from DW-MRI, has been shown to correlate with cellularity.[4–6] Longitudinally measured changes in ADC values have been utilized to predict the response of breast tumors to therapy[7–9] and have been shown to be a better predictor of eventual response than measurements of tumor size.[10] Accurate T_1 values are

essential for quantitatively assessing DCE-MRI, which consists of the serial acquisition of T_1-weighted images before, during, and after the injection of a gadolinium-based MRI contrast agent. Quantitative pharmacokinetic parameters derived from DCE-MRI, such as K^{trans} (the volume transfer rate related to vascular permeability and perfusion), v_e (extravascular extracellular volume fraction), v_p (plasma volume fraction), and k_{ep} ($=K^{trans}/v_e$) have been shown to be good predictors of pathological response in breast tumors undergoing anticancer therapy prior to any significant changes in tumor size.[8,11–13] Central to the quantitative analysis of DCE-MRI is the accurate measurement of T_1 values of the tissue under investigation prior to contrast administration. Such "relaxometry" measurements are not typically made at community-based radiology facilities. Importantly, these quantitative MRI parameters that characterize underlying tissue physiology may interrogate important changes of breast tumors prior to changes in tumor size.

This study evaluated the ability to implement and assess the accuracy, repeatability, and reproducibility of ADC and T_1-mapping in the community setting through evaluations of both phantom studies and normal human subjects. Accuracy of the quantitative MRI metrics was evaluated through phantom imaging with known T_1 and ADC values. Repeatability of these measurements was assessed in a test–retest population of normal subjects. Reproducibility was assessed in repeat scanning of phantoms and individuals across three sites.

Materials and Methods

Scanning Protocol

Imaging data were acquired at two community imaging facilities and one research facility using 3T Siemens (Erlangen, Germany) Skyra scanners, software VE11a, equipped with 8- or 16-channel receive double-breast coil (Sentinelle; InVivo, Gainesville, FL). In general, a community care imaging center is a nonacademic, nonresearch setting (ie, does not actively train medical residents) and is either an outpatient facility or a hospital that is regionally used. One of the community centers was an outpatient imaging facility, equipped with one 3T and one 1.5T MRI and utilized for only outpatient exams. The second community facility is a regional hospital, ~120 beds, equipped with two MRIs (one 3T and the other 1.5T), and utilized for both inpatient and outpatient services. Both community imaging facilities currently perform breast MRI as part of their routine clinical practice. All three MRIs used in this study are at different sites and on different service contracts, as well as have different quality control guidelines. Additionally, the radiology technologists currently employed at the various sites were directly involved and responsible for positioning the subjects and deploying the research imaging protocols. The human subjects and diffusion phantom were scanned in the sagittal plane, while the T_1 phantom was imaged in the coronal plane. A highresolution anatomical image was acquired using a gradient echo fast low-angle shot (FLASH) 3D sequence for a total scan time of 3 minutes 11 seconds. (Please see Table 1 for complete imaging parameters.)

DW-MRI was acquired using a monopolar single-shot spin echo (SE) echo planar imaging (EPI) sequence in three diagonal diffusion-encoding directions. Six acquisitions were averaged for *b*-values of 0 and 200 s/mm^2, while 18 acquisitions were averaged with a *b*-

value of 800 s/mm^2. This allowed for approximately equal signal-to-noise ratios.[14] DW-MRI was acquired over 10 slices with 5 mm thickness and no slice gap. Spectrally selective adiabatic inversion recovery (SPAIR) fat suppression was included in the human subject protocol for a total scan time of 1 minute 39 seconds.

For the T_1 measurement, a variable flip angle data approach was employed using a 3D spoiled gradient echo (SPGE) sequence at 10 flip angles (2°, 4°, 6°,…,20°) for a total acquisition time of 99 seconds. A Siemens TurboFLASH sequence was used to map the B_1 field to correct for transmit inhomogeneity. Due to the inclusion of a slice gap in the B_1 mapping protocol, two acquisitions were performed to cover the same field of view (FOV) as the above measurements for a total acquisition time of 34 seconds. To confirm the T_1 phantom values calculated using the variable flip angle technique, a partial saturation recovery technique was acquired using a spin echo sequence with 28 different TR values, ranging from 19 msec to 8000 msec.

Figure 1 shows a pipeline of the acquisition, analysis, and visualization of the study metrics.

Phantom Studies

For phantom studies, a coronal image orientation was prescribed in the center of the breast coil containing the phantoms. To test T_1 measurement accuracy, eight gel phantoms (Eurospin II Test System, Diagnostic Sonar), with T_1 values ranging from 300–1700 msec, were submerged in water and scanned at room temperature. To assess the accuracy of ADC measurements, an ice-water diffusion phantom[15] was allowed to reach 0°C and then securely placed in the center of the breast coil for scanning. Phantoms were scanned for intrasite repeatability ($n = 3$ repeated measures with removal of the phantom between sessions) on a 16-channel breast coil at one of the three sites (research facility). Phantoms were also scanned at each of three facilities to assess intersite reproducibility of the measurements at community radiology centers.

Volunteer Studies

All protocols were approved by the Institutional Review Boards (IRBs) and informed consent was collected. Women ($n = 12$) with no previous or currently diagnosed breast disease, who were neither pregnant nor breastfeeding, underwent two quantitative MRI examinations within the same day with the subject removed from the MRI table between scans. Only one breast per subject was imaged (to follow suit with future clinical studies where only the affected breast will be imaged). These repeatability studies were completed on a 16-channel breast coil at a single site. The mean age of the normal subjects was 40 years old (range 22–62). For each breast, fibroglandular tissue and adipose tissue were segmented using the anatomical images (segmentation details provided in the next section). Two of the normal subjects did not have sufficient fibroglandular tissue to be assessed for repeatability and were evaluated solely on their adipose tissue. An additional three subjects were scanned at each of the three facilities to assess reproducibility of the measurements between both academic and community radiology centers.

Quantitative Analysis

SEGMENTATION—The segmentation of fibroglandular and adipose tissues was conducted on downsampled higher spatial resolution anatomical images. Downsampled images were regridded to the same resolution as the acquisitions used for quantification. First, the image was preprocessed to enhance contrast.[16] In brief, a transfer function was determined by assigning a new intensity to each voxel according to an adaptive transfer function that is designed on the basis of the local statistics (based on the min, max, and average for the voxels) in order to take advantage of range stretching. Therefore, this local-statistics-based transfer function was applied to each voxel in the anatomical image, so that the histogram equalization was performed on the local scale, resulting in enhancement and background suppression simultaneously.[16] Then a *k*-means clustering algorithm determined the intensity thresholding to generate the initial masks for fibroglandular, adipose, and voxels that contain both adipose and fibroglandular tissues. The voxels located near the edge of the region that contained both tissue types were systematically eliminated from analysis. Finally, small "islands" of voxels (size <20 voxels) were eliminated from the fibroglandular mask and adipose mask to yield a contiguous, homogeneous section of tissue on which to perform statistics.

For the T_1-mapping of the phantoms, circular regions of interest (ROIs) were manually drawn within each gel phantom using the central slice of the high-resolution image. The average T_1 from each phantom was recorded from each of the T_1 maps obtained from the two techniques (ie, the variable flip angle and the partial saturation recovery). For the ADC mapping of the phantoms, a rectangular ROI was manually drawn to encompass the entire saline phantom inside of the ice bath and the average ADC from each phantom was recorded.

DW-MRI—DW-MRI was collected as described above and used to extract ADC values for every voxel within the phantoms and segmented fibroglandular tissue. ADC maps were calculated from the DW-MRI data by fitting the data to Eq. [1]:

$$S(b) = S_0 \bullet \exp(-ADC \bullet b) \quad (1)$$

where $S(b)$ is the signal intensity in the presence of diffusion gradients at strength b, S_0 is the signal intensity in the absence of diffusion gradients, and the b-value is the strength of the diffusion gradients. Equation [1] was fitted to the signal intensities from the three b-values (0, 200, 800 s/mm^2) on a voxel-by-voxel basis using a nonlinear least squares approach written in MatLab (Math Works, Natick, MA). Accepted reference values of ADC measurements in the ice-water phantom at 0°C are 1.1×10^{-3} mm^2/s.[15] ADC measurements of fibroglandular tissue in normal subjects were evaluated for repeatability at one site and reproducibility across three imaging sites.

T_1-MAPPING—T_1 maps were generated using two distinct methods. The first, which employed variable flip angles, is rapid and appropriate for routine patient use, but is prone to inaccuracies due to the B_1 inhomogeneities commonly seen at higher fields.[17] To address

this issue, we employed a B_1 correction map that quantifies the difference between the prescribed and actual flip angles at each voxel. Quantitative B_1-corrected, variable flip angle, T_1 values were calculated for each voxel by fitting the signal intensity, *S*, data to Eq. [2]:

$$S = S_0 \bullet \left[\frac{\sin(f \bullet \alpha) \bullet \left(1 - \exp\left(-\frac{TR}{T_1}\right)\right)}{1 - \left(\exp\left(-\frac{TR}{T_1}\right) \bullet \cos(f \bullet \alpha)\right)} \right] \quad (2)$$

where S_0 is a constant related to scanner gain and proton density, α is the prescribed flip angle, and f is the flip angle correction factor that accounts for inhomogeneity in B_1. We have also assumed $TE \ll T_2^*$.

The second method employed partial saturation recovery as a complementary measurement of T_1 for comparison with the variable flip angle technique to confirm accurate measurements; however, the acquisition time of this technique makes it unsuitable for routine patient use. Quantitative T_1 values were calculated for each voxel by fitting the signal intensity data from the partial saturation recovery sequence to Eq. [3]:

$$S(TR) = S_0\left(1 - \exp\left(\frac{-TR}{T_1}\right)\right) \quad (3)$$

where $S(TR)$ is the signal intensity measured at each TR and S_0 is the baseline signal intensity. T_1 maps of a phantom that consisted of standardized T_1 vials obtained by a B_1-corrected, variable flip angle approach was compared to those obtained by a partial saturation recovery method to assess accuracy. The results of the B_1-corrected, variable flip angle T_1-mapping was compared for repeatability at one site and reproducibility across imaging sites for fibroglandular tissue and adipose tissue in normal subjects.

Statistical Analysis

T_1 and ADC measurements were acquired for each individual voxel across the phantom and whole breast imaging, and then summarized as the mean value for the entire ROI of the phantom, fibroglandular, or adipose tissue. Statistical analysis was performed using the statistical toolbox in MATLAB as in previous efforts.[17,18] The repeatability, reproducibility, and accuracy of ADC and T_1 were calculated. The accuracy compares measured values in phantoms at one location compared to known or gold standard measurements. The percent difference (ie, the accuracy) between the measured values and the known or gold standard values was calculated to determine the accuracy of the T_1 and ADC mapping methods in the phantom studies. Repeatability evaluates the repeat scans of breast tissue or phantoms at a single site on the same day. Reproducibility evaluates the phantom and breast tissue scans across multiple sites. Reported values are listed as the mean ± standard error. Specifically, for the T_1 phantom, which includes eight separate vials, the analysis counts each individual gel phantom as one separate component and compares the difference of each individual gel

to its repeat measures. Following the calculated differences, the mean of all the differences was computed. As the sample sizes are small for phantom and subject data, further information is provided in tables and graphs for reproducibility and repeatability.[19]

Repeatability statistics used in this test–retest study follow the methods previously described by Bland and Altman[20] and applied in the breasts of healthy volunteers.[21,22] First, the absolute difference, *d*, was calculated between the ADC and T_1 datasets obtained for each subject. Kendall's tau test was performed to test whether the magnitude of the differences was correlated with the parameter mean of the repeated measurements. An additional Kendall's tau test was performed to test for correlation between each MRI metric and age, as it is known that breast tissue composition changes with age.[23] The Wilcoxon signed-rank test was used to test the null hypothesis of no bias (ie, median difference is zero) between the first and second measurements, with the Spearman correlation testing the effectiveness of the pairing. The repeatability of each MRI metric (ie, ADC and T_1) was assessed. Additionally, the 95% confidence interval (CI), the root-mean-squared deviation, the within-subject standard deviation, and the repeatability value (*r*) were calculated for each MRI metric. The root-mean-square deviation (rMSD) is computed using the difference *d:*

$$rMSD = \sqrt{\frac{\sum d^2}{n}} \quad (4)$$

where *n* is the number of subjects. The 95% CI for a group of *n* subjects:

$$CI = \pm \frac{1.96 \bullet std(d)}{\sqrt{n}} \quad (5)$$

where *std(d)* is the standard deviation of *d*. The confidence interval indicates the range of expected measurement variability in a group of *n* subjects. The within-subject standard deviation *(wSD)* is:

$$wSD = \frac{rMSD}{\sqrt{2}} \quad (6)$$

The repeatability coefficient, *r*, is:

$$r = 2.77 \bullet wSD \equiv 1.96 \bullet rMSD \quad (7)$$

This repeatability coefficient defines the magnitude of the maximum difference (absolute value) between repeated observations expected in 95% of paired observations; for example, *r* defines the expected measurement variability for an individual. Due to our moderate sample size, we used the appropriate *t*-statistic: 2.262 for 10 datasets (fibroglandular tissue), and 2.201 for 12 datasets (adipose tissue).

The reproducibility of each derived MRI metric (ADC of fibroglandular tissue, T_1 of fibroglandular tissue, T_1 of adipose tissue) was quantified for subjects across three MRI scanners using the Friedman test to assess for differences across the three sites. A significance value of $P \leq 0.05$ was used for all statistical tests.

Last, the minimum sample size required to detect a parameter change of 5%, 10%, or 20% was calculated for ADC or T_1 value estimation method with two-tailed statistical significance set at alpha of 0.05 and power of 0.80 and 0.90 and using standard deviations from the reproducibility data. The percentage changes were based on the population mean for each estimation method and the respective scaling metric in this study.

Results

DW-MRI

ADC values of the ice-water phantom were found to be 1.22×10^{-3} $mm^2/s \pm 0.005 \times 10^{-3}$ mm^2/sec, which constitutes a 10% difference compared to literature values for water at ~0°C.[15] Repeatability of ADC revealed a less than 0.01% average difference between scans. Additionally, ADC measurements were compared across the three sites for the standardized DW-MRI phantom (Fig. 2). Phantom reproducibility revealed an average 2.7% difference between the sites, ADC measurements of $1.22 \times 10^{-3} \pm 0.04 \times 10^{-3}$, $1.26 \times 10^{-3} \pm 0.06 \times 10^{-3}$, and $1.24 \times 10^{-3} \pm 0.06 \times 10^{-3}$ mm^2/sec (Supplemental Fig. 1).

ADC measurements of fibroglandular tissue in normal subjects are shown for repeatability at one site (Fig. 3) and reproducibility across three imaging sites (Fig. 2). The average ADC of fibroglandular tissue was $1.34 \times 10^{-3} \pm 0.33 \times 10^{-3}$ mm^2/sec. Repeatability scans of the same subject's fibroglandular tissue showed an average percent difference of 4.5% in ADC measurement between the two scans. There were no significant correlations between subject age and ADC ($P = 0.18$). Repeatability of ADC in fibroglandular tissue was not significantly different ($P = 0.92$), with the pairing significantly correlated ($r_s = 1.0$; [$P < 0.0001$]). Additionally, the difference between repeated measurements was independent of the mean (Fig. 3). The repeatability statistics for ADC measurements are summarized in Table 2. Reproducibility scans of normal breast fibroglandular tissue at the three different sites yielded an average difference of 7.0% in ADC measurement between sites (Fig. 2, Table 3), with average ADC measurements for the three subjects to be $1.41 \times 10^{-3} \pm 0.03 \times 10^{-3}$, $1.44 \times 10^{-3} \pm 0.03 \times 10^{-3}$, and $1.54 \times 10^{-3} \pm 0.04 \times 10^{-3}$ mm^2/sec at each site, respectively. There were no statistical differences detected between the sites ($P = 0.1$).

T_1-Mapping

The average percent difference between the saturation recovery (gold-standard) and B_1-corrected, variable flip angle approach over the eight vials was 3.1% ± 3.0% difference, ranging from 0.2% to 8.1% for any individual vial (Fig. 4, Table 4). Repeatability of T_1 phantom measurements with the B_1-corrected, variable flip angle approach revealed a 1.7% 6 0.77% difference between scans ($P = 0.55$). Reproducibility experiments found an average 2.1% difference between the sites (Fig. 5), as seen in Table 5 and Supplemental Fig. 2 ($P =$

0.36). Statistical comparisons were calculated on the repeats of the individual T_1 vials within the phantom.

The results of the B_1-corrected, variable flip angle T_1-mapping are shown for repeatability at one site and reproducibility across imaging sites for fibroglandular tissue and adipose tissue in normal subjects (Fig. 5). The average T_1 values of fibroglandular tissue and adipose tissue were 948 ± 287 msec and 425 ± 52 msec, respectively. Repeatability scans showed an average percent difference of 6.7% in the T_1 of fibroglandular tissue and 5.9% in the T_1 of adipose tissue (Fig. 6, Supplemental Fig. 3, respectively). There were no significant trends between subject age and T_1 ($P = 0.27$). Repeatability of T_1 in adipose and fibroglandular tissue was not significantly different ($P = 0.09$, $P = 0.08$, respectively), with the pairing significantly correlated ($r_s = 0.78$ [$P = 0.002$], $r_s = 0.94$ [$P = 0.0001$], respectively). Additionally, the difference between repeated measurements was independent of the mean (Fig. 6). The repeatability statistics for T_1 measurements are summarized in Table 2. Reproducibility scans of normal breast fibroglandular tissue at the three different sites yielded an average difference of 8.4% in T_1 measurement across sites in fibroglandular tissue, and 7.5% in T_1 measurement of adipose tissue between sites (Table 3, Fig. 5), with no statistical differences observed between the sites ($P = 0.1$, $P = 0.7$, respectively).

Power Analysis

For ADC, a minimum number of 3-16 subjects is necessary to statistically power a study in which a 5-20% change is expected. For changes in T_1 (as required for DCE-MRI analysis), a minimum number of 3-29 subjects is necessary to statistically power a study in which a 5-20% change is expected. These results are summarized in Table 6.

Discussion

We found that a quantitative breast MRI protocol can be deployed at community imaging centers with acceptable accuracy, repeatability, and reproducibility, thereby demonstrating the feasibility of future widespread clinical dissemination. This study has two separate, interrelated, goals: 1) increase the ability for community centers to participate in multisite clinical trials, and 2) improve patient care by dissemination of quantitative MRI for early prediction of response to NAT. Importantly, repeatability in normal subjects was tested across a range of ages comparable to the age ranges of women diagnosed with breast cancer. Furthermore, reproducibility of both phantoms and normal subjects was tested across multiple scanners at three sites, allowing for additional verification of dissemination of quantitative MRI.

The ADC values of the ice-water phantom were found to be consistent with the range reported in the literature values of $1.05–1.20 \times 10^{-3}$ mm^2/s[15,24,25] for water at ~0°C. Malyarenko et al have shown that repeatability of ADC for day-to-day measurements in a phantom was 4.5%, while the errors between academic sites was 3% at the isocenter and ~10% off center.[25] In normal subjects, Clendenen et al measured the ADC of fibroglandular tissue as 1.15×10^{-3} mm^2/s, ranging from 0.84×10^{-3} mm^2/s to 1.55×10^{-3} mm^2/s, consistent with our measurements.[26] Furthermore, in normal subjects at 3T at another academic site, Giannotti et al observed a 6% error between scanners of ADC phantoms, and

an 8% error on test-retest repeatability of 10 subjects.[27] Aliu et al found an 11% within-subject error at an academic site in fibroglandular tissue in nine healthy subjects (accounting for menstrual cycle changes).[28] Our results acquired at community centers reveal similar differences for repeatability in normal subjects and phantoms. Additionally, repeatability of ADC in fibroglandular tissue was not significantly different, suggesting a lack of bias between the first and second scanning sessions. Furthermore, we found a low average difference for breast tissue across multiple community locations in the same subject across multiple days.

T_1 measurements were found to have similar accuracy and repeatability in the community setting when compared to the academic setting. The T_1 values for fibroglandular tissue and adipose tissue calculated were consistent with previously reported values of 1324–1444 msec for fibroglandular tissue and 366–449 msec for adipose tissue in the breast at 3T.[29] Bane et al observed a median error of 0.7–25.8% in test-retest repeatability of a T_1-phantom when using a variable flip angle (without B_1 correction) approach at eight different centers.[30] Whisenant et al led a study evaluating repeatability of T_1-mapping in normal breast tissue and revealed that precision and accuracy were improved when implementing a B_1 correction in T_1-mapping of breast tissue, reporting a 3.2% and 2.6% difference in repeat analysis of adipose tissue and fibroglandular tissue, respectively, of the same subjects at the same site.[17] Using a similar B_1 correction, our study found a similar difference in adipose and fibroglandular tissue in test–retest repeat analysis at the same site. Repeatability of T_1 in adipose and fibroglandular tissue was not significantly different, suggesting a lack of bias between the first and second scanning sessions. Furthermore, for reproducibility across various sites, our study showed a low average difference in adipose tissue and fibroglandular tissue.

While there has been much effort in making quantitative imaging techniques available for clinical trials in academic research centers, there have been comparatively few studies in community-based radiology centers—where most patients receive their imaging care.[3] The results from the phantom and normal subject scans indicate that quantitative MRI of the breast, which has previously proven useful for assessing breast tumors in both the diagnostic[31,32] and prognostic settings[7,33] at academic medical centers, can also be successfully and reliably implemented at community imaging centers. Evaluating each of these quantitative MRI techniques across multiple community sites provides evidence that these metrics can be employed in (for example) a multisite clinical trial using community centers to assess treatment response of standard-of-care or novel therapeutics. This is of great importance because it presents the ability for medical providers—other than academic researchoriented medical centers—the ability to be intimately involved in clinical trials that require advanced imaging. Incorporating networks of community practice sites into multisite trials might allow us to collect more data in less time, thus enhancing the overall research effort and compressing the timeframe of clinical translation.

There are, however, acute differences in MRI scanner logistics between the academic and community setting. MRI scanners in the community setting do not typically have a research contract, thereby limiting the available scans (including some sequences that academic settings runs as standard-of-care) and preventing the scanners from running any “work-in-

progress" imaging sequences. Additionally, although hardware and software between scanners should essentially be the same, there are known differences across any MRI scanners, including frequency and phase drift, magnet inhomogeneity, and temperature instability. Although the magnets and coils tested in this study were the same model from the same manufacturer, with the exception of an 8-channel coil at one site, they were sited at different institutions and subject to individual variations (as mentioned above). Importantly, our results show that the reproducibility of these metrics is within a reasonable range to allow for future widespread dissemination of these techniques. Furthermore, although most sites follow American College of Radiology (ACR) accreditation, academic medical centers must typically follow more rigorous standards in order to participate in clinical trials, sometimes utilizing phantom studies similar to those included in this work.[24,34,35]

Lack of standardization in both the acquisition and analysis of quantitative imaging, due to the tremendous number of potentially adjustable parameters at different sites, is also of notable concern. Major challenges that exist for both academic and community imaging are harmonization in acquisition protocols (both across centers and vendors) and scanner upgrades. To be adopted as routine imaging biomarkers, the accuracy, repeatability, and reproducibility of MRI metrics must be established and validated with multisite studies. The Quantitative Imaging Network (QIN) has led a number of efforts to address this limitation by standardizing methods for image acquisition and analysis.[17,36] Furthermore, the QIN has led numerous efforts in assessing the repeatability and reproducibility of ADC and T_1-mapping at multiple academic sites. Again, there has been limited effort to replicate such data in the community setting. Additionally, there are currently no multisite, multivendor studies validating the quantitative MRI techniques.[37] Future work is needed across multiple vendors and magnet strengths (ie, 1.5T and 3T) to characterize the robustness of quantitative MRI techniques and assess their use as imaging biomarkers in community imaging centers. This analysis is important for dissemination of quantitative metrics to community imaging centers.

There are a number of limitations in the present study. For example, one of the sites had an 8-channel breast coil, while the other two sites used 16-channel breast coils. Importantly, this single site still was able to reproduce statistically similar quantitative metrics for all parameters assessed, suggesting that these measurements can be robust across variation in coil hardware. A second limitation of our study is that we did not assess T_1 and ADC across multiple vendors or magnet strengths, and as previously mentioned, this is a requirement for widespread dissemination of quantitative MRI. This study, however, does outline the necessary steps needed to perform a power analysis for future clinical studies and multisite multivendor studies. Additionally, we did not control for the timing of the menstrual cycle when the scans were performed on the subjects who were imaged at multiple institutions. This can be a potential source of error, as the relative timing within a woman's menstrual cycle can cause fluctuations in breast density, which may[38] or may not[39] affect ADC measurements. Importantly, the techniques evaluated were robust despite these differences. As performing reproducibility studies in individuals across multiple locations with strict scheduling logistics is exceedingly difficult, we were limited to three such subjects in the present study who were scanned at each location. Furthermore, to implement these quantitative metrics of the breast into clinical studies that evaluate breast lesions and tumors,

additional efforts on tumor segmentation will need to be completed. For our current normal subject studies, we have only segmented fibroglandular and adipose tissue and voxels that encompassed both adipose and fibroglandular tissue were eliminated from this analysis; however, for future applications to breast cancer, it will be important to evaluate all voxels that include any voxel within the tumor. Additionally, T_1 mapping provides the basis for DCE-MRI; however, future steps in tumor evaluation will require parameter evaluation of quantitative DCE-MRI parameters K^{trans}, v_e, v_p, and k_{ep}.

In conclusion, quantitative T_1 and ADC mapping of the breast can be implemented in the community-based imaging setting, thereby dramatically increasing the access of these techniques for patients who can benefit from these technologies. We recently implemented these techniques in a community-focused clinical trial to predict the response of locally advanced breast cancer to neoadjuvant therapy.[40]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Contract grant sponsor: National Cancer Institute; contract grant number: U01CA174706 and U01CA142565; Contract grant sponsor: Cancer Prevention and Research Institute of Texas; contract grant number: RR160005; T.E.Y. is a CPRIT scholar of cancer research

We thank the volunteers for participating in this study. We also thank the MRI technicians at Austin Radiological Associates and Seton Hospital for assistance with running scanning protocols.

References

1. Abramson RG, Arlinghaus LR, Dula AN, et al. MR imaging biomarkers in oncology clinical trials. Magn Reson Imaging Clin N Am. 2016; 24:11–29. [PubMed: 26613873]
2. Rosenkrantz AB, Mendiratta-Lala M, Bartholmai BJ, et al. Clinical utility of quantitative imaging. Acad Radiol. 2015; 22:33–49. [PubMed: 25442800]
3. Forte GJ, Hanley A, Hagerty K, Kurup A, Neuss MN, Mulvey TM. American Society of Clinical Oncology National Census of Oncology Practices: Preliminary report. J Oncol Pract. 2013; 9:9–19. [PubMed: 23633966]
4. Padhani AR, Liu G, Koh DM, et al. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: Consensus and recommendations. Neoplasia. 2009; 11:102–125. [PubMed: 19186405]
5. Chenevert TL, Brunberg JA, Pipe JG. Anisotropic diffusion in human white matter: Demonstration with MR techniques in vivo. Radiology. 1990; 177:401–405. [PubMed: 2217776]
6. Patterson DM, Padhani AR, Collins DJ. Technology insight: Water diffusion MRI—a potential new biomarker of response to cancer therapy. Nat Clin Pract Oncol. 2008; 5:220–233. [PubMed: 18301415]
7. Galban CJ, Ma B, Malyarenko D, et al. Multi-site clinical evaluation of DW-MRI as a treatment response metric for breast cancer patients undergoing neoadjuvant chemotherapy. PLoS One. 2015; 10:e0122151. [PubMed: 25816249]
8. Li X, Abramson RG, Arlinghaus LR, et al. Multiparametric magnetic resonance imaging for predicting pathological response after the first cycle of neoadjuvant chemotherapy in breast cancer. Invest Radiol. 2015; 50:195–204. [PubMed: 25360603]
9. Wu LM, Hu JN, Gu HY, Hua J, Chen J, Xu JR. Can diffusion-weighted MR imaging and contrast-enhanced MR imaging precisely evaluate and predict pathological response to neoadjuvant

chemotherapy in patients with breast cancer? Breast Cancer Res Treat. 2012; 135:17–28. [PubMed: 22476850]

10. Sharma U, Danishad KK, Seenu V, Jagannathan NR. Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy. NMR Biomed. 2009; 22:104–113. [PubMed: 18384182]
11. Abramson RG, Arlinghaus LR, Weis JA, et al. Current and emerging quantitative magnetic resonance imaging methods for assessing and predicting the response of breast cancer to neoadjuvant therapy. Breast Cancer (Dove Med Press). 2012; 2012:139–154. [PubMed: 23154619]
12. Tudorica A, Oh KY, Chui SY, et al. Early prediction and evaluation of breast cancer response to neoadjuvant chemotherapy using quantitative DCE-MRI. Transl Oncol. 2016; 9:8–17. [PubMed: 26947876]
13. Tateishi U, Miyake M, Nagaoka T, et al. Neoadjuvant chemotherapy in breast cancer: Prediction of pathologic response with PET/CT and dynamic contrast-enhanced MR imaging—prospective assessment. Radiology. 2012; 263:53–63. [PubMed: 22438441]
14. Burdette JH, Durden DD, Elster AD, Yen YF. High b-value diffusion-weighted MRI of normal brain. J Comput Assist Tomogr. 2001; 25:515–519. [PubMed: 11473179]
15. Chenevert TL, Galban CJ, Ivancevic MK, et al. Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies. J Magn Reson Imaging. 2011; 34:983–987. [PubMed: 21928310]
16. Yu ZeyunBajaj C. International Conference on Image Processing. Vol. 2. IEEE; 2004. A fast and adaptive method for image contrast enhancement.
17. Whisenant JG, Dortch RD, Grissom W, Kang H, Arlinghaus LR, Yankeelov TE. Bloch-Siegert B1-mapping improves accuracy and precision of longitudinal relaxation measurements in the breast at 3T. Tomography. 2016; 2:250–259. [PubMed: 28044146]
18. Whisenant JG, Ayers GD, Loveless ME, Barnes SL, Colvin DC, Yankeelov TE. Assessing reproducibility of diffusion-weighted magnetic resonance imaging studies in a murine model of HER2 + breast cancer. Magn Reson Imaging. 2014; 32:245–249. [PubMed: 24433723]
19. Weissgerber TL, Milic NM, Winham SJ, Garovic VD. Beyond bar and line graphs: Time for a new data presentation paradigm. PLoS Biol. 2015; 13:e1002128. [PubMed: 25901488]
20. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999; 8:135–160. [PubMed: 10501650]
21. Dula AN, Arlinghaus LR, Dortch RD, et al. Amide proton transfer imaging of the breast at 3T: Establishing reproducibility and possible feasibility assessing chemotherapy response. Magn Reson Med. 2013; 70:216–224. [PubMed: 22907893]
22. Zhu H, Arlinghaus LR, Whisenant JG, Li M, Gore JC, Yankeelov TE. Sequence design and evaluation of the reproducibility of water-selective diffusion-weighted imaging of the breast at 3T. NMR Biomed. 2014; 27:1030–1036. [PubMed: 24986756]
23. Rutter CM, Mandelson MT, Laya MB, Seger DJ, Taplin S. Changes in breast density associated with initiation, discontinuation, and continuing use of hormone replacement therapy. JAMA. 2001; 285:171–176. [PubMed: 11176809]
24. Mulkern RV, Ricci KI, Vajapeyam S, et al. Pediatric brain tumor consortium multisite assessment of apparent diffusion coefficient z-axis variation assessed with an ice-water phantom. Acad Radiol. 2015; 22:363–369. [PubMed: 25435183]
25. Malyarenko D, Galban CJ, Londy FJ, et al. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. J Magn Reson Imaging. 2013; 37:1238–1246. [PubMed: 23023785]
26. Clendenen TV, Kim S, Moy L, et al. Magnetic resonance imaging (MRI) of hormone-induced breast changes in young premenopausal women. Magn Reson Imaging. 2013; 31:1–9. [PubMed: 22898693]
27. Giannotti E, Waugh S, Priba L, Davis Z, Crowe E, Vinnicombe S. Assessment and quantification of sources of variability in breast apparent diffusion coefficient (ADC) measurements at diffusion weighted imaging. Eur J Radiol. 2015; 84:1729–1736. [PubMed: 26078100]
28. Aliu SO, Jones EF, Azziz A, et al. Repeatability of quantitative MRI measurements in normal breast tissue. Transl Oncol. 2014; 7:130–137. [PubMed: 24772216]

29. Rakow-Penner R, Daniel B, Yu H, Sawyer-Glover A, Glover GH. Relaxation times of breast tissue at 1.5T and 3T measured using IDEAL. J Magn Reson Imaging. 2006; 23:87–91. [PubMed: 16315211]
30. Bane O, Hectors SJ, Wagner M, et al. Accuracy, repeatability, and interplatform reproducibility of T1 quantification methods used for DCE-MRI: Results from a multicenter phantom study. Magn Reson Med. 2017 Epub ahead of print.
31. Abe H, Mori N, Tsuchiya K, et al. Kinetic analysis of benign and malignant breast lesions with ultrafast dynamic contrast-enhanced MRI: Comparison with standard kinetic assessment. AJR Am J Roentgenol. 2016:1–8.
32. Huang W, Tudorica LA, Li X, et al. Discrimination of benign and malignant breast lesions by using shutter-speed dynamic contrast-enhanced MR imaging. Radiology. 2011; 261:394–403. [PubMed: 21828189]
33. Yankeelov TE, Lepage M, Chakravarthy A, et al. Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: Initial results. Magn Reson Imaging. 2007; 25:1–13. [PubMed: 17222711]
34. Chen CC, Wan YL, Wai YY, Liu HL. Quality assurance of clinical MRI scanners using ACR MRI phantom: Preliminary results. J Digit Imaging. 2004; 17:279–284. [PubMed: 15692871]
35. Jackson EF. Standards for imaging endpoints in clinical trials: standardization and optimization of image acquisitions: MR. FDA Workshop. 2010
36. Arlinghaus LR, Dortch RD, Whisenant JG, Kang H, Abramson RG, Yankeelov TE. Quantitative magnetization transfer imaging of the breast at 3.0T: Reproducibility in healthy volunteers. Tomography. 2016; 2:260–266. [PubMed: 28090588]
37. Yankeelov TE, Mankoff DA, Schwartz LH, et al. Quantitative imaging in cancer clinical trials. Clin Cancer Res. 2016; 22:284–290. [PubMed: 26773162]
38. Hovhannisyan G, Chow L, Schlosser A, Yaffe MJ, Boyd NF, Martin LJ. Differences in measured mammographic density in the menstrual cycle. Cancer Epidemiol Biomarkers Prev. 2009; 18:1993–1999. [PubMed: 19567508]
39. Kim JY, Suh HB, Kang HJ, et al. Apparent diffusion coefficient of breast cancer and normal fibroglandular tissue in diffusion-weighted imaging: The effects of menstrual cycle and menopausal status. Breast Cancer Res Treat. 2016; 157:31–40. [PubMed: 27091644]
40. Sorace AG, Virostko J, Wu C. , et al. San Antonio Breast Cancer Symposium. Vol. P4-02-08. San Antonio, TX: 2017. Repeatability and reproducibility of quantitative breast MRI in community imaging centers: Preliminary results.

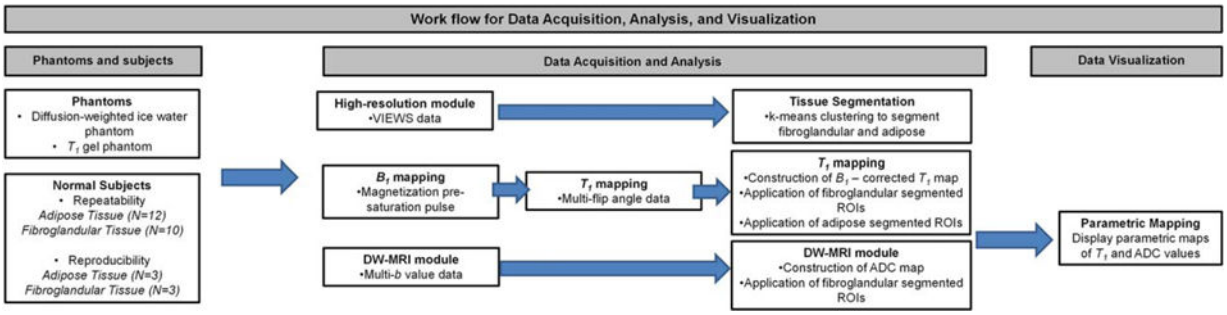

FIGURE 1.

Pipeline of data acquisition, analysis, visualization, and processing for quantitative ADC and T_1 measurements in phantoms and normal breast tissue.

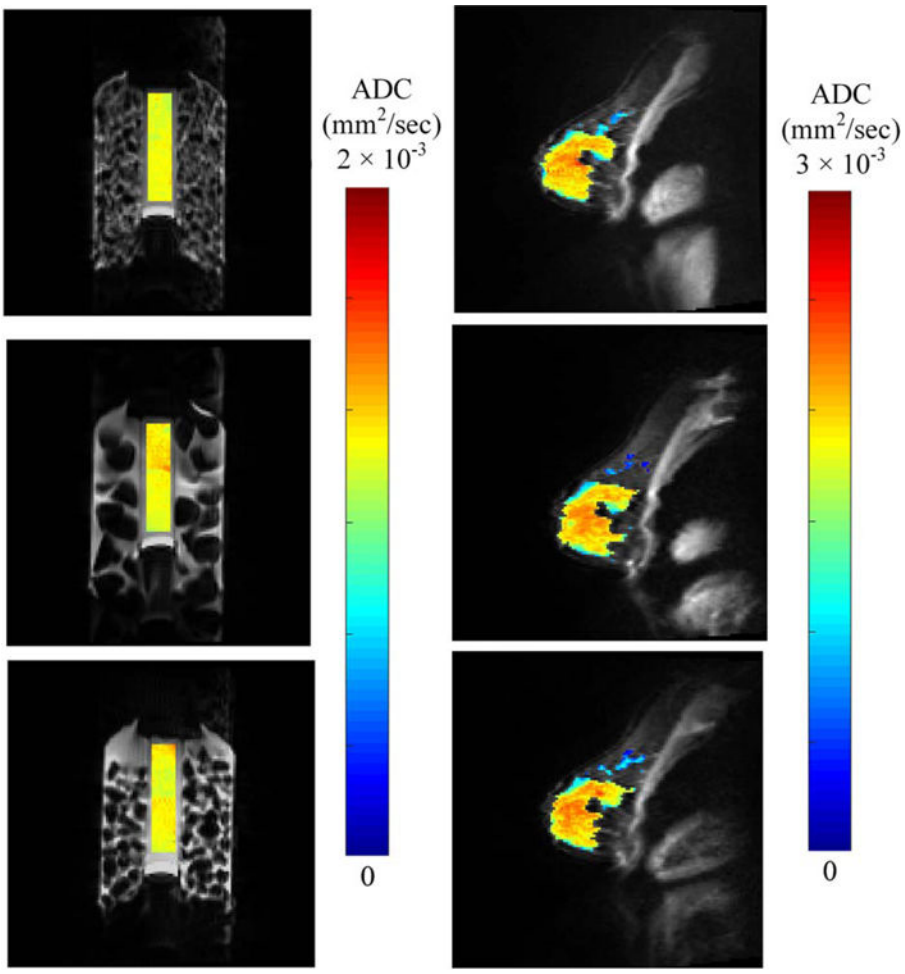

FIGURE 2.

Reproducibility of ADC reveals less than a 10% change across three community sites. Shown are the cross-sectional images of ADC values overlaid on the high-resolution anatomical image of (a) phantom and (b) one example normal subject ADCs from the three sites.

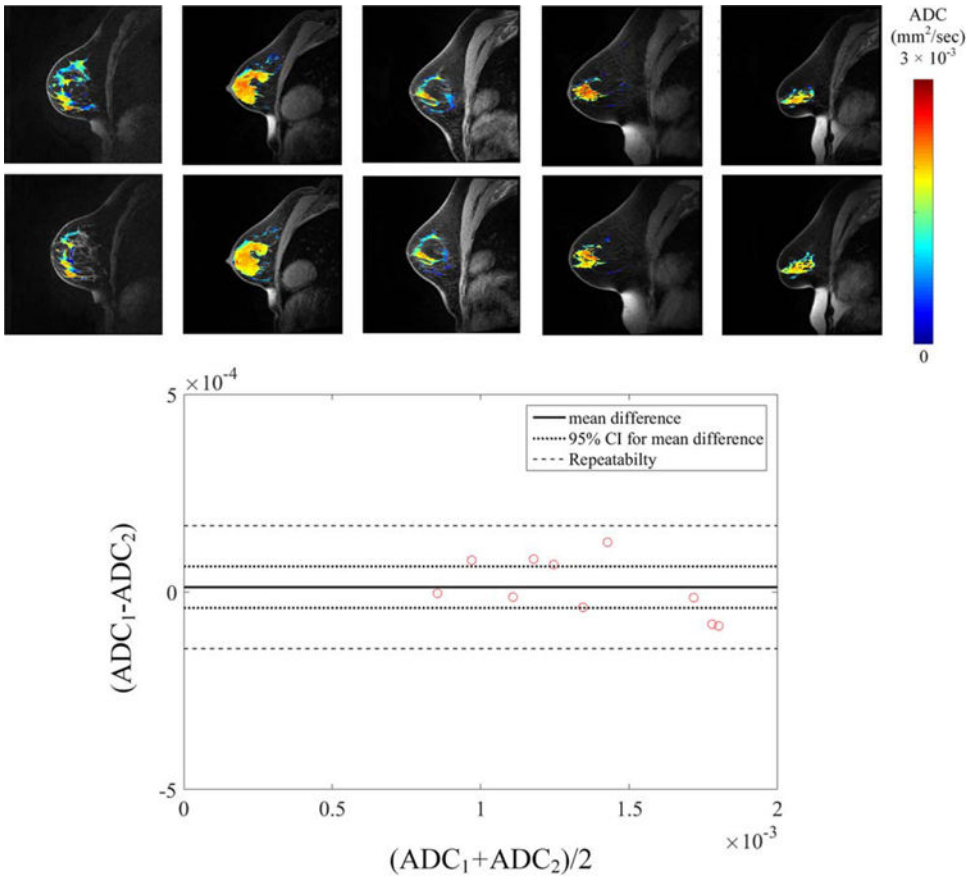

FIGURE 3.

Repeatability of ADC reveals high accuracy of obtaining accurate repeat measurements in breast tissue. Shown are (a) the cross-sectional images of ADC values of fibroglandular tissue overlaid on the high-resolution anatomical image of five test-retest subjects and the associated (b) Bland-Altman plots from all subjects.

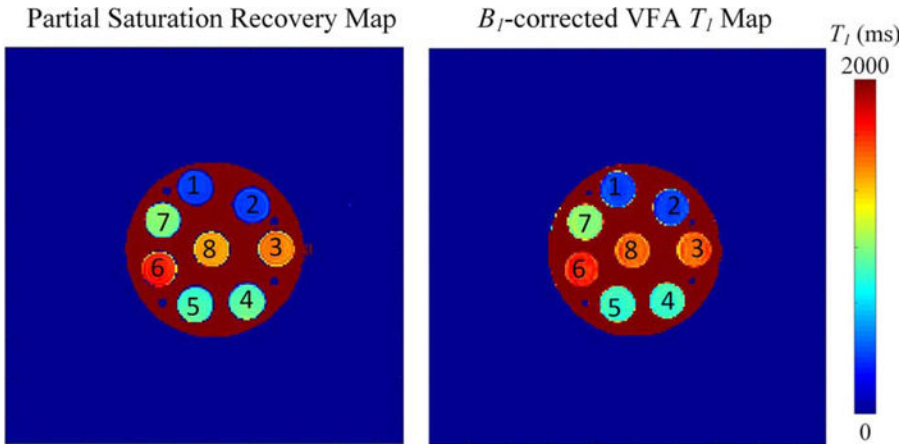

FIGURE 4.

Accuracy of T_1-mapping with partial saturation recovery T_1 map in a T_1 phantom reveals that the B_1-corrected T_1 values obtained are both accurate and clinically able to be obtained.

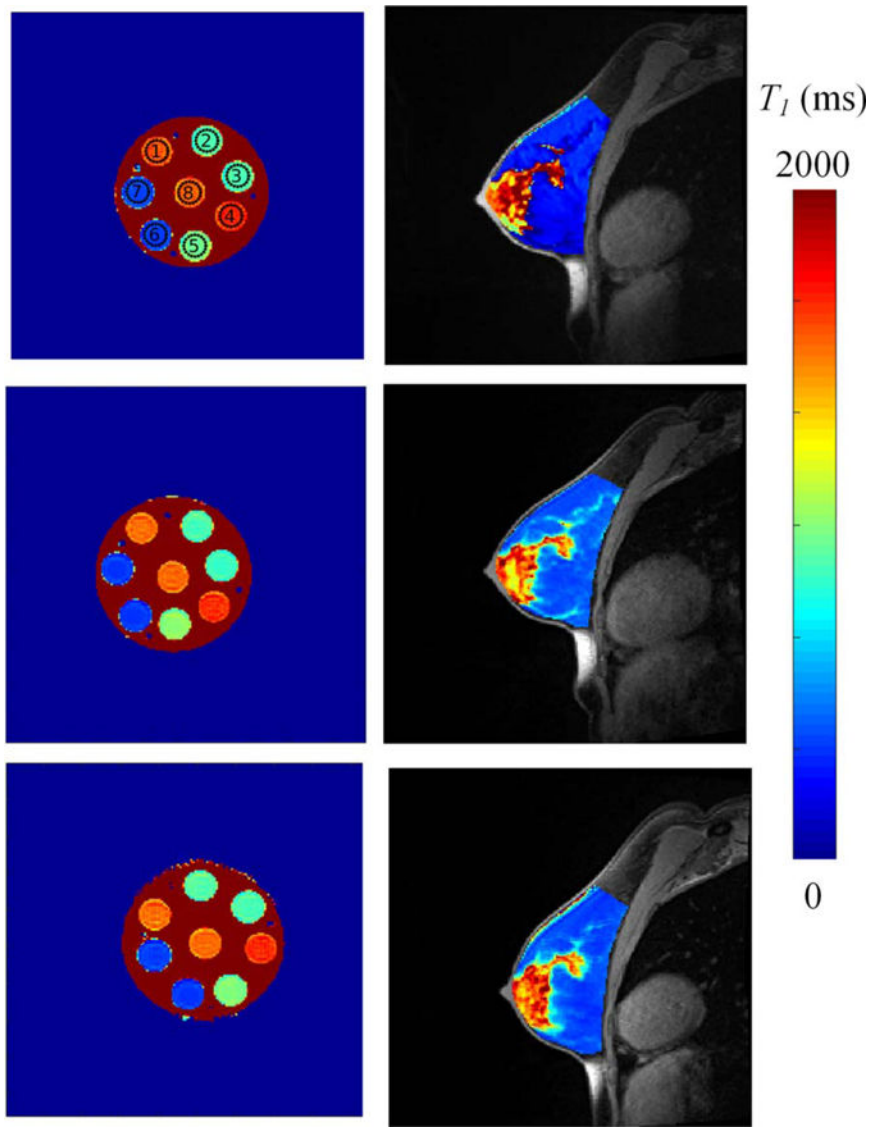

FIGURE 5.

Reproducibility of T_1-mapping reveals less than 10% change across three community sites. Shown are the crosssectional images of T values overlaid on the high-resolution anatomical image of (a) phantom and (b) example normal subject B_1-corrected T_1 map from the three sites.

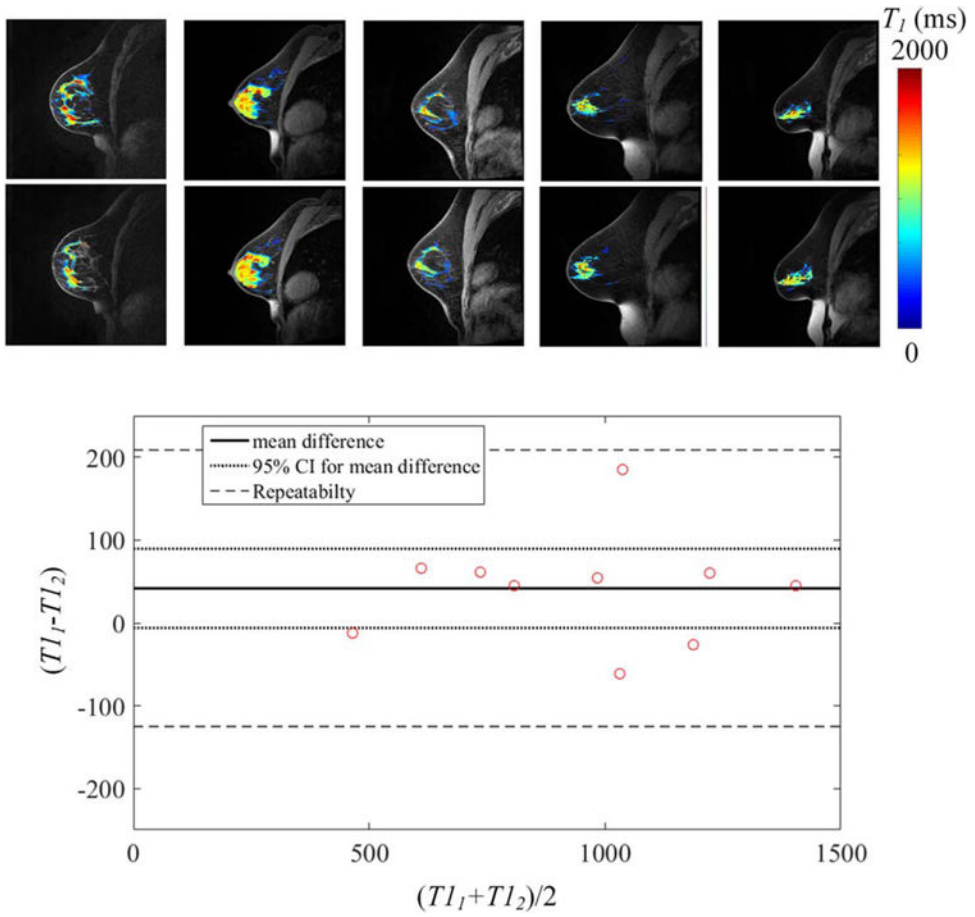

FIGURE 6.

Repeatability of T_1 reveals high accuracy of obtaining accurate repeat measurements in breast tissue. Shown are (a) the cross-sectional images of B_1-corrected T_1 values of fibroglandular tissue overlaid on the high-resolution anatomical image of five test–retest subjects and the associated (b) Bland–Altman plots from all subjects.

TABLE 1

MRI Scan Parameters

MRI parameters	Anatomical scan	DW-MRI	B_1 mapping	T_1-mapping: variable flip angle	T_1-mapping: partial saturation recovery
Scan sequence	T_1 –weighted 3D gradient-echo FLASH	Single-shot spin echo (SE) echo planar (EPI)	3D spoiled gradient-echo	T_1–weighted 3D spoiled gradient-echo	T_1–weighted 3D spoiled gradient-echo
TR (msec)	5.3	3000	8680	7.9	19–8000 (28 different values)
TE (msec)	2.3	52	2	2.71	12
Flip angle (degrees)	20	90	8	2, 4, 6, 8, 10, 12, 14, 16, 18, 20	90
Acquisition matrix	256 × 256	128 × 128	96 × 96	192 × 192	192 × 192
FOV (mm)	256 × 256	256 × 256	256 × 256	256 × 256	256 × 256
Slice thickness (mm)	1	5	5	5	5
GRAPPA acceleration factor	2	2	N/A	3	N/A
Fat suppression	SPAIR	SPAIR	N/A	N/A	N/A

TR = repetition time; TE = echo time; FOV = field of view; GRAPPA = generalized autocalibrating partial parallel acquisition; SPAIR = spectral attenuated inversion recovery.

TABLE 2

Repeatability Statistics for Normal Subjects (N = 10–12)

	ADC (N = 10)	**T_1 (B_1 corrected variable flip angle) fibroglandular tissue (N = 10)**	**T_1 (B_1 corrected variable flip angle) adipose tissue (N = 12)**
Kendall's tau (age vs. mean)	−0.360	−0.494	−0.260
Kendall's tau (difference vs. mean)	0.378	−0.067	0.303
Kendall's tau, P-value (age vs. mean)	0.178	0.0593	0.271
Kendall's tau, P-value (difference vs. mean)	0.156	0.862	0.197
95% CI (percentage of mean)	5.25×10^{-5} (3.91%)	47.8 (5.04%)	21.3 (5.00%)
Root mean square deviation, (percentage of mean)	7.07×10^{-5} (5.27%)	75.9 (8.00%)	34.2 (8.03%)
Within-subject standard deviation (percentage of mean)	5.00×10^{-5} (3.72%)	53.7 (5.66%)	24.2 (5.68%)
Repeatability value (r), (percentage of mean)	1.56×10^{-4} (11.59%)	167.0 (17.62%)	75.2 (17.68%)

TABLE 3

Mean Reproducibility ADC and B_1-Corrected, T_1 Values for Each of the Normal Subjects Across the Three Sites

	Site 1	Site 2	Site 3
ADC (mm^2/sec)			
Fibroglandular tissue			
Subject 1	0.00113	0.00115	0.00117
Subject 2	0.00136	0.00154	0.00157
Subject 3	0.00173	0.00161	0.00187
B_1-corrected, T_1 (msec)			
Fibroglandular tissue			
Subject 1	944.9	918.3	946.7
Subject 2	1062.1	1217.2	1218.2
Subject 3	1191.6	1242.0	1473.3
Adipose tissue			
Subject 1	517.8	426.7	408.7
Subject 2	396.3	407.1	403.4
Subject 3	454.6	470.7	489.5

TABLE 4

Accuracy of the B_1-Corrected, Variable Flip Angle, T_1-Mapping Sequence for Each of the Phantom Vials

Phantom mean T_1 (msec)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
B_1-corrected variable flip angle	351.3	351.3	1527	844.6	835.4	1661.4	1004	1526
Partial saturation recovery	350.6	346.8	1503	915.8	877.9	1691.9	999.8	1426
% Difference	0.2%	1.3%	1.6%	8.1%	5.0%	1.8%	0.4%	6.8%

TABLE 5

Mean B_1-Corrected, T_1 Values for Each of the Phantom Vials Across the Three Sites

	Mean T_1 (msec)							
Site location	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Site 1	1578.8	893.2	891.7	1671.3	964.3	359.9	380.8	1554.8
Site 2	1538.7	897.5	854.4	1653.2	969.7	350.8	351.6	1531.5
Site 3	1559.8	896.6	883.8	1592.1	1031.0	361.1	362.6	1486.3

TABLE 6

Power Analysis for Future Clinical Trials

	Minimum sample size estimations					
	To detect a 5% change		To detect a 10% change		To detect a 20% change	
Quantitative imaging metric	ADC	T_1	ADC	T_1	ADC	T_1
Power = 80%	12	22	5	7	3	3
Power = 90%	16	29	6	9	3	4