Clinical Imaging 42 (2017) 119–125

Contents lists available at ScienceDirect

Clinical Imaging

journal homepage: http://www.clinicalimaging.org

Original Article

Clinical applicability and relevance of fibroglandular tissue segmentation on routine T1 weighted breast MRI

Akshat C. Pujara [a,1], Artem Mikheev [a,b], Henry Rusinek [a,b], Harikrishna Rallapalli [b], Jerzy Walczyk [a,b], Yiming Gao [a,c], Chloe Chhor [a,c], Kristine Pysarenko [a,c], James S. Babb [a,b], Amy N. Melsaether [a,c,*]

[a] Department of Radiology, New York University School of Medicine, 550 First Avenue, New York, NY 10016, USA
[b] Bernard and Irene Schwartz Center for Biomedical Imaging, New York University School of Medicine, 660 First Avenue, 4th Floor, New York, NY 10016, USA
[c] Perlmutter Cancer Center, New York University School of Medicine, 160 East 34th Street, New York, NY 10016, USA

ARTICLE INFO

Article history:
Received 26 September 2016
Received in revised form 7 November 2016
Accepted 2 December 2016

Keywords:
Breast
Magnetic resonance imaging (MRI)
Segmentation
Quantification
Fibroglandular tissue (FGT)

ABSTRACT

Purpose: To evaluate clinical applicability of fibroglandular tissue (FGT) segmentation on routine T1 weighted breast MRI and compare FGT quantification with radiologist assessment.
Methods: FGT was segmented on 232 breasts and quantified, and was assessed qualitatively by four breast imagers.
Results: FGT segmentation was successful in all 232 breasts. Agreement between radiologists and quantified FGT was moderate to substantial (kappa = 0.52–0.67); lower quantified FGT was associated with disagreement between radiologists and quantified FGT ($P \le 0.002$).
Conclusions: FGT segmentation was successful using routine T1 weighted breast MRI. Radiologists were less consistent with quantified results in breasts with lower quantified FGT.

1. Introduction

Fibroglandular tissue (FGT) is the breast parenchyma that can enhance after contrast administration on magnetic resonance imaging (MRI). FGT thus represents the region of clinical interest during breast MRI interpretation. The fraction of breast tissue containing FGT correlates with mammographic breast density [1,2], an independent risk factor for developing breast cancer [3–7], and is recommended to be included in breast MRI reports [8]. Clinical evaluation of breast FGT fraction from high-resolution T1 weighted MRI is performed using a standardized 4-point qualitative scale (a–d) [8] and requires the radiologist to visually integrate information from dozens of slices. As such, qualitative FGT assessment is subject to intra- and inter-reader variability [7]. Quantitative, automated FGT segmentation may be useful for providing an accurate, reproducible measure of FGT. Further, FGT segmentation is an essential first step for automated quantification of background parenchymal enhancement on breast MRI, which has also been proposed as a potential imaging biomarker of breast cancer risk and response to therapy [6,7,9].

FGT segmentation first requires isolation of the breasts from the chest, which can be achieved using contour-based [10–13], region-based [14–16], or atlas-based [17–19] methods. Following whole breast segmentation, the key challenge of accurate separation of FGT and breast adipose tissue on routinely acquired T1 weighted sequences is the variable signal of commercial breast coils [20,21] (Fig. 1). Signal variability often results in misclassification of adipose tissue located distant from the coil as FGT. Separation of FGT from fat thus typically entails sophisticated non-uniformity correction (NUC) followed by two-class tissue separation. Various methods for automatic separation of FGT and fat have been reported, but with limitations. Some authors recommend acquiring phase-shifted data using Dixon techniques [22–24]; however, adding scan time to acquire this additional sequence is not ideal in the clinical routine, and accuracy of Dixon methods can be affected by degree of T1 weighting and MR sequence selection [22]. Atlas-aided fuzzy C-means methods require development of a prior likelihood atlas and are applied two-dimensionally due to the variable amount and complex distribution of FGT in the breast [25]. Lastly, conclusions from studies of hierarchical support vector machines [23] and adaptive thresholding [26] are limited by validation in small numbers of cases (4 and 20, respectively).

We have developed an automated software tool for 3-D separation of FGT and breast adipose tissue that does not require additional sequences or scan time. The main step involves an NUC algorithm that

* Corresponding author at: 160 East 34th Street, 3rd Floor, New York, NY 10016, United States.
E-mail addresses: apujara@gmail.com (A.C. Pujara), artemmikheev@gmail.com (A. Mikheev), hr18@nyu.edu (H. Rusinek), Harikrishna.Rallapalli@nyumc.org (H. Rallapalli), Jerzy.Walczyk@nyumc.org (J. Walczyk), Yiming.Gao@nyumc.org (Y. Gao), Chloe.Chhor@nyumc.org (C. Chhor), Kristine.Pysarenko@nyumc.org (K. Pysarenko), James.Babb@nyumc.org (J.S. Babb), Amy.Melsaether@nyumc.org (A.N. Melsaether).
[1] Department of Radiology, Northwestern University, 250 East Superior Street, 4th Floor, Chicago, IL 60611, USA.

http://dx.doi.org/10.1016/j.clinimag.2016.12.002
0899-7071/

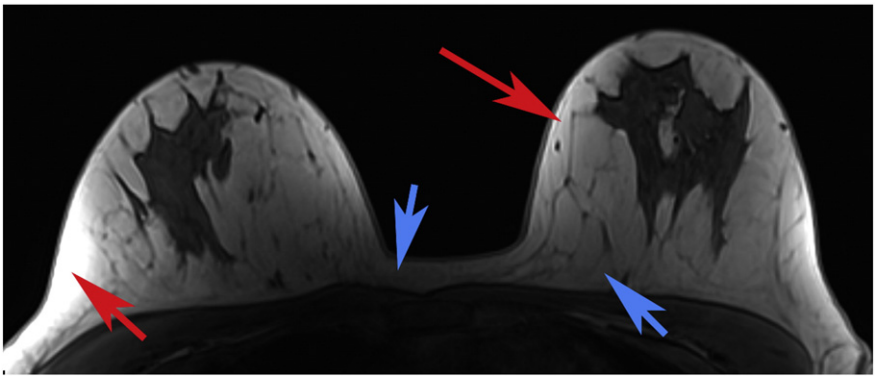

Fig. 1. Typical signal non-uniformities in clinical T1 weighted breast magnetic resonance imaging. Breast tissue located adjacent to coil elements (red arrows) shows brighter signal compared with breast tissue located further from coil elements (blue arrows). Such signal non-uniformities can result in misclassification of adipose tissue located distant from the coil as fibroglandular tissue.

accounts for coil-specific 3-D distributions of non-uniformities present in the acquired images, which pose challenges for existing post-processing algorithms [27]. Specifically, we hypothesized that initial optimization of NUC parameters using images of breast phantoms would allow for successful FGT segmentation using only a single clinically acquired T1 weighted sequence. The objective of this study was to evaluate the clinical applicability of FGT segmentation using this technique on routine T1 weighted breast MRI. Quantified FGT was also compared with radiologist qualitative FGT assessment to evaluate the clinical need for FGT quantification.

2. Material and methods

This study was approved by the Institutional Review Board and compliant with the Health Insurance Portability and Accountability Act. Patients were retrospectively selected for inclusion and written informed consent was waived.

2.1. Phantoms

2.1.1. Assembly and imaging

Four pairs of breast phantoms were assembled using 0.1 mM manganese chloride to represent FGT and canola oil (Crisco, The J.M. Smucker Company, Orrville, OH, USA) to represent adipose tissue. The 0.1 mM dilution was chosen to mimic contrast between FGT and fat on human breast MRI acquired with a clinical T1 weighted sequence (see below for parameters). The manganese chloride and canola oil were either freely poured or contained in 80 cm^3 Mylar pouches (MylarFoil MiniPouch, IMPAK Corporation, Los Angeles, CA, USA) within a one-gallon resealable plastic bag (Ziploc, S.C. Johnson & Son, Racine, WI, USA). The total volume of each breast phantom was either 1106 cm^3 or 1346 cm^3.

Axial T1 weighted, non-contrast, 3-D volumetric scans of the four breast phantom pairs were obtained without fat saturation using a dedicated seven-element surface breast coil (Sentinelle, Invivo, Gainesville, FL, USA) on a 3 Tesla magnet (MAGNETOM Trio, A Tim System, Siemens Medical Solutions, Erlangen, Germany) with clinical acquisition parameters (repetition time/echo time = 4.74 ms/1.79 ms, field-of-view 320 mm^2, matrix 448 × 358, slice thickness 1.10 mm). Each phantom pair was scanned three times. At each acquisition, the bags were rearranged in different positions to assess whether quantification was morphologically dependent. MR images were divided in the midline, resulting in 24 individual phantom breast volumes (Fig. 2).

2.1.2. Non-uniformity correction

NUC for FGT segmentation was predicated on the previously reported BiCal technique [28]. The algorithm represents the multiplicative bias field as a slowly varying function whose logarithm is represented as the linear combination of 3-D Legendre polynomials. The key steps are 1) exclusion of image "edges," or regions of rapid signal transition, and 2) constrained smoothing of remaining areas prior to model fitting. The key parameters are the smoothing radius, R, and the maximum degree of Legendre polynomials, L.

Four phantom breast volumes with variable amounts of FGT and fat were selected as training volumes to optimize BiCal parameters (R, L) for T1 weighted breast examinations. Using locally developed software (FireVoxel, New York University Medical Center, New York, NY, USA), the operator placed a total of 240 pairs of regions of interest (ROIs) in the training data to sample signal uniformity and tissue contrast. Each pair consisted of one circular 6 mm diameter ROI placed on FGT and a similar ROI placed on nearby fat. The two ROIs were placed within 25 mm of one another to estimate signal contrast that is minimally contaminated by non-uniformity artifacts. For each tested set of BiCal parameters, we computed the variability of signal in each tissue type and the distribution of contrast between FGT and fat. We used the fraction of falsely classified voxels at the optimal cutoff as the metric to minimize signal variability while preserving tissue contrast. The optimized phantom-derived parameters ($R = 45$ mm, $L = 10$) were validated on all 24 phantom breast volumes (Fig. 3) and then applied to patient MRI exams.

2.1.3. Signal intensity histogram thresholding

After NUC and exclusion of background air and surface voxels contaminated with partial volume artifact, the signal intensity histogram was well approximated as the linear sum of two Gaussian probability distributions. Each histogram was decomposed as the linear combination of two peaks, with a lower intensity peak representing FGT voxels and a higher intensity peak representing fat voxels. A signal intensity threshold, T, was computed for each histogram as the value where the two signal distributions intersected (Fig. 3) and was used to quantify tissue volumes. Estimated volumes were recorded for comparison with known values.

2.2. Patients

2.2.1. Patients and MRI exam

All axial breast MRIs performed between 7/2/4/2015 – the first date of axial breast MRI acquisition at our institution – and 11/30/2015 were reviewed. Patients with negative or benign MRI results were included. For the purposes of this study, patients with a possible breast cancer or a history of breast cancer were excluded. As part of clinical protocol, both breasts were imaged using the same dedicated breast coil, 3 Tesla magnet, and T1 weighted parameters used for phantom imaging.

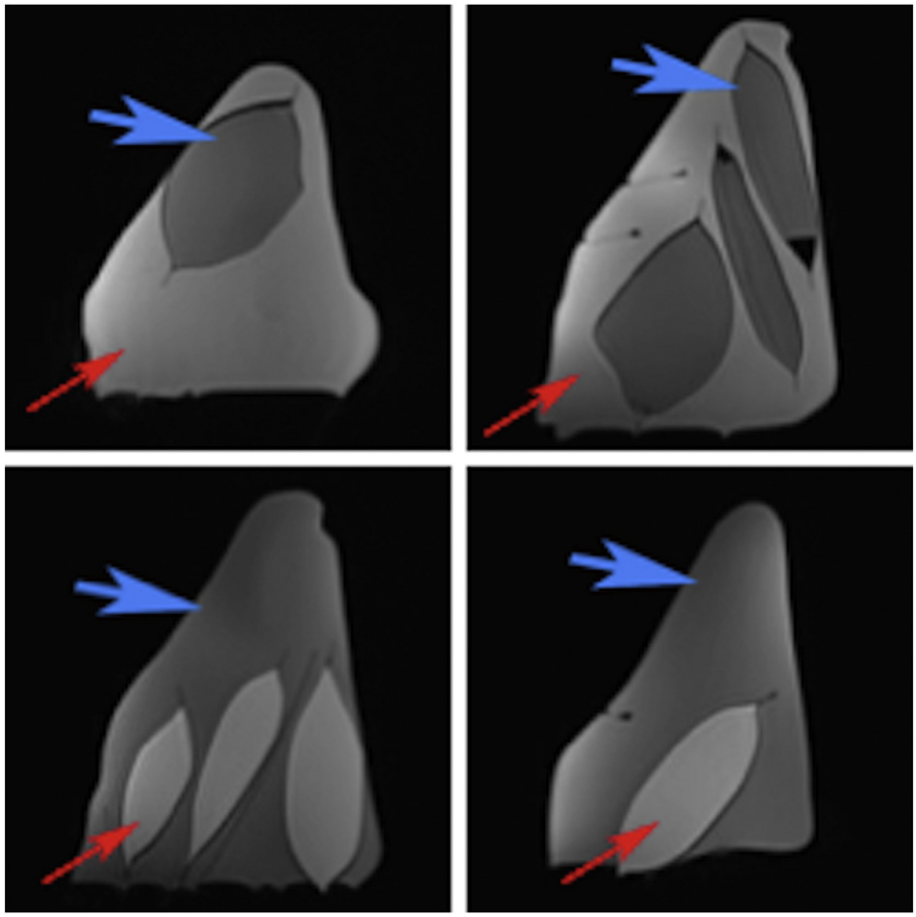

Fig. 2. T1 weighted images of breast phantoms. Breast phantoms with variable amounts of phantom fibroglandular tissue (FGT; blue arrows) and phantom fat (red arrows) were used for derivation of optimal non-uniformity correction parameters. Top left (a) = 14.5% FGT: 160 cm^3 FGT in pouches, 946 cm^3 fat freely poured. Top right (b) = 29.7% FGT: 400 cm^3 FGT in pouches, 946 cm^3 fat freely poured. Bottom left (c) = 70.3% FGT: 946 cm^3 FGT freely poured, 400 cm^3 fat in pouches. Bottom right (d) = 85.5%: 946 cm^3 FGT freely poured, 160 cm^3 fat in pouches.

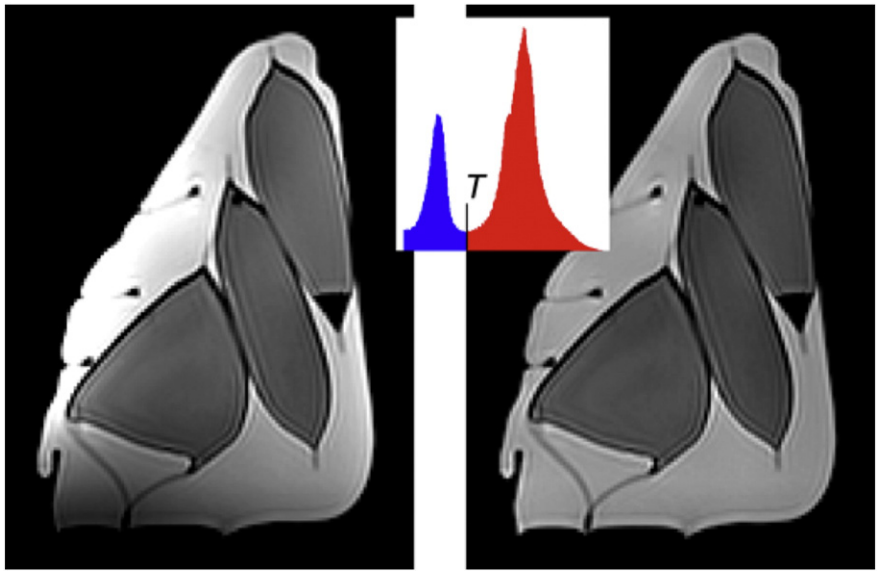

Fig. 3. Phantom breast before and after non-uniformity correction (NUC), with a signal intensity histogram after NUC. The left image (a) shows a representative slice of a phantom breast before NUC. The right image (b) shows the same image after NUC. Inset: the signal intensity histogram of the corrected image is well approximated by a bi-Gaussian probability distribution. The automatically identified threshold between the fibroglandular tissue signal peak (blue) and the fat signal peak (red) is indicated by *T* and was used to quantify tissue volumes.

2.2.2. Whole breast segmentation

A single radiologist drew two-dimensional contours in the axial plane every fifteenth slice (on average ten contours per case) to separate the breasts from the chest wall, following the same pre-defined borders for each case (sternoclavicular joints superiorly, inframammary folds inferiorly, lateral margins of sternum medially, anterior margins of latissimus dorsi muscle anteriorly). Contouring every fifteenth slice allowed for faster segmentation as compared with more frequent contouring, while preserving anatomic accuracy after the contours were automatically filled and interpolated to yield whole breast masks. Interpolation between filled contours *P* and *Q* was achieved by constructing all possible line segments connecting pairs of voxels across *P* and *Q*, and then intersecting each segment with the intermediate slice. Partial volume reduction and skin removal were applied to the whole breast masks using a 1.5 mm morphologic erosion operator within our in-house software (FireVoxel). The resulting 3-D breast ROI was reviewed and divided in the midline. No manual correction of breast segmentation was performed.

2.2.3. FGT segmentation and quantification

Optimized phantom-derived BiCal NUC and signal intensity histogram segmentation were applied to all 232 breasts. The resulting FGT and adipose masks were visually reviewed to assess for successful segmentation; no manual correction of FGT segmentation was performed. Percent FGT was calculated as (absolute FGT volume / total breast volume) × 100. Automated FGT segmentation is illustrated in Fig. 4.

2.2.4. Radiologist FGT assessment

Blinded to quantitative results, four breast imagers (R1–R4) with 2–8 years experience (YG 2, KP 6, AM 7, and CC 8 years) independently and qualitatively graded FGT on the T1 weighted images for 232 breasts on a 4-point scale in accordance with the 2013 Breast Imaging-Reporting and Data System lexicon (up to 25% = a; 26–50% = b; 51–75% = c; >75% = d) [8]. Readers established consensus FGT reads, which were only used for receiver operating characteristic (ROC)-based conversion of percent FGT to reader scoring of a–d.

2.3. Statistical analysis

Quantification of breast phantoms was compared with known composition by calculating absolute errors for calculated percent FGT. Coefficients of variation and intra-class correlation coefficients were computed to assess the repeatability of percent FGT measurement after re-positioning.

For patient MRIs, within-subject correlations were accounted for in each aspect of statistical analysis. Inter-reader agreement was assessed using the linear weighted kappa coefficient (k) with 95% confidence intervals (CI). $k < 0.0$ was interpreted as poor agreement, $0.0 \le k \le 0.20$ as slight agreement, $0.20 < k \le 0.40$ as fair agreement, $0.40 < k \le 0.60$ as moderate agreement, and $k > 0.60$ as substantial agreement [29]. Spearman correlation (r) was performed to assess the relationship between reader FGT and quantified percent FGT. $r > 0.70$ was considered strong, $0.30 \le r \le 0.70$ moderate, and $r < 0.30$ weak.

Three sequential ROC analyses of the 4-point consensus qualitative reads were performed to convert quantitative semi-automated percent FGT results to 4-point ordinal assessments; reader scoring of a–d was recoded to 1–4 for this portion of the analysis. An initial ROC analysis identified the threshold of percent FGT that optimally discriminated between reader scores of 1 and those greater than 1. The data were then pared by removing reader scores of 1, and the ROC analysis was repeated to identify the threshold of percent FGT that best discriminated between reader scores of 2 and those greater than 2, and then again to differentiate between scores of 3 and 4.

Agreement between individual readers and converted percent FGT was then assessed using the linear weighted kappa coefficient (k) as above. Mann-Whitney tests were performed to assess whether the level of quantified percent FGT impacted the frequency of concordance and discordance between readers and semi-automated quantification. All statistical tests were conducted at the two-sided 5% significance level using SAS 9.3 software (SAS Institute, Cary, NC).

3. Results

3.1. Phantom FGT segmentation

For 24 phantom breast volumes overall, mean absolute error for percent FGT was 2.1 ± 0.7%. Tests of intra-phantom percent FGT quantification repeatability demonstrated a coefficient of variation of 0.74% and an intra-class correlation coefficient of 0.99. Quantified percent FGT compared with known percent FGT for all 24 phantom breast volumes is shown in Fig. 5.

3.2. Patient FGT segmentation

Between 7/24/2015 and 11/27/2015, 369 consecutive, unique patients (ages 23–83 years, mean 53 years) underwent axial breast MRI. Of these 369 patients, 116 (31.4%) patients (ages 23–81 years, mean 49 years) met the inclusion criteria of no history of breast cancer and

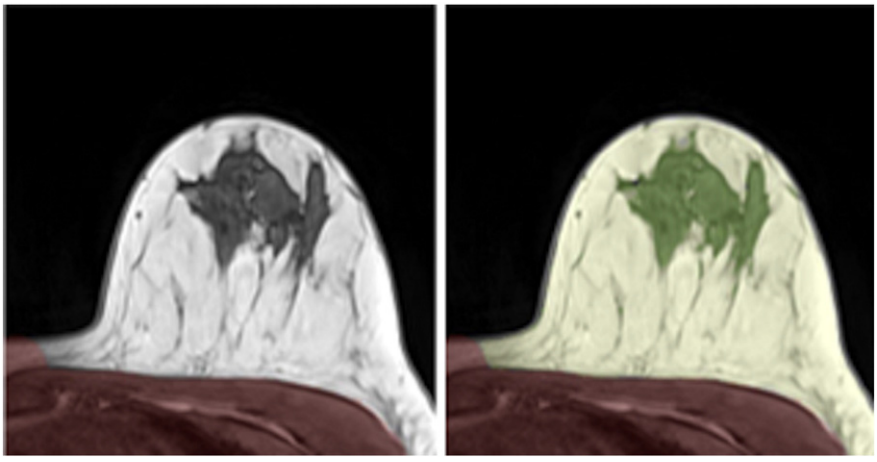

Fig. 4. Automated fibroglandular tissue (FGT) segmentation. The left image (a) shows a representative T1 weighted image of a human breast following separation from the chest, division in the midline, and non-uniformity correction. The right image (b) shows the same image after application of the signal intensity threshold for FGT segmentation and generation of a FGT mask shown in green. The accompanying adipose tissue mask is shown in yellow.

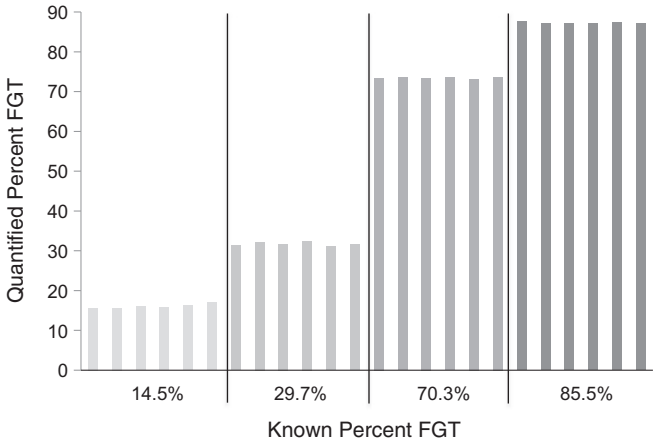

Fig. 5. Quantified percent fibroglandular tissue (FGT) of breast phantoms as compared with known values. Each of the four pairs of breast phantoms was scanned in three different positions, for a total of 12 scans. Division of images in the midline resulted in 24 phantom breast volumes (six for each FGT composition). For all 24 phantom breast volumes, the mean absolute error for percent FGT compared with known values was 2.1 + 0.7%. The coefficient of variation was 0.74% and the intra-class correlation coefficient was 0.99, indicating good reproducibility.

negative or benign MRI results. Indications for MRI included screening in 89/116 (76.7%; BRCA1 positive in 11, BRCA2 positive in 14, family history of breast cancer in 56, history of atypia in 5, and history of lobular carcinoma in situ in 3 patients), follow-up of prior findings assessed as probably benign on screening MRI in 15/116 (12.9%), and problem solving in 12/116 (10.3%) patients.

Segmentation of each pair of breasts from the chest wall was the only non-automatic step in FGT segmentation and averaged approximately 7 min per case. After saving breast masks as image files, the remaining steps were fully automatic and 232 volumes were analyzed as a single batch processing job at a rate of <1 min per breast on a desktop computer. Visual inspection of the 232 processed volumes demonstrated anatomically successful FGT segmentation in all cases.

3.3. Agreement for patient FGT

Inter-reader agreement for FGT across all reader pairs was substantial (k = 0.69). The distribution of quantified percent FGT as a function of individual reader assessment is shown in Fig. 6. Spearman correlations between reader FGT and quantified percent FGT were strong (R1 $r = 0.85$, R2 $r = 0.84$, R3 $r = 0.84$, R4 $r = 0.81$; $P < 0.001$ for all 4 correlations). Sequential ROC analyses of consensus reads indicated that reader FGT was best predicted by semi-automated percent FGT ≤9.1 = a, 9.1–13.4 = b, 13.5–26.1 = c, >26.1 = d. Agreement between reader FGT and converted semi-automated FGT among the 232 breasts was moderate to substantial [R1 k = 0.67 (95% CI 0.61–0.73); R2 k = 0.60 (95% CI = 0.54–0.67); R3 k = 0.62 (95% CI 0.55–0.69); R4 k = 0.52 (95% CI 0.46–0.58)]. Lower levels of quantified percent FGT were associated with decreased frequency of concordance between readers and semi-automated quantification for all 4 readers ($P \leq 0.002$) (Table 1).

4. Discussion

Breast MRI poses unique challenges for FGT segmentation due to variability in breast size, morphology, distribution of FGT, and signal non-uniformities. We have developed a phantom-validated, semi-automated FGT segmentation technique that can be performed on routine clinical T1 weighted breast MRI without the added scan time required by Dixon techniques [22–24]. Breast phantoms helped to optimize the key NUC step and showed accurate and reproducible results for FGT segmentation. The clinical robustness of this technique was demonstrated by successful FGT segmentation of 232 breasts imaged over a 4 month time period without manual correction. These attributes of the method for FGT segmentation reported here are favorable for integration into clinical workflows. Comparison between quantified FGT and radiologist qualitative FGT assessment suggests a need for incorporation of FGT quantification into clinical practice.

The FGT segmentation method described in this study includes only one non-automatic step, namely the use of semi-automated chest wall contouring to first isolate the breasts from the chest, as in several other studies [10–13]. Fully automated chest wall detection algorithms are being evaluated to automate this step, but with noteworthy

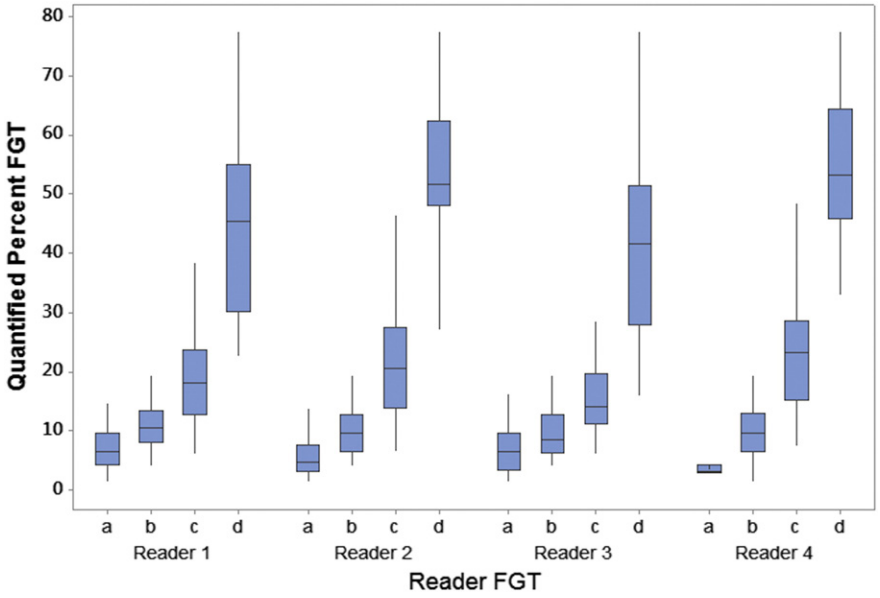

Fig. 6. Distribution of quantified percent fibroglandular tissue (FGT) as a function of individual reader assessment of FGT. Each box indicates the median percent FGT and interquartile range (IQR) corresponding to each level of reader qualitative assessment. Whiskers extend 1.5 × IQR above and below. Readers did not discriminate levels of quantified percent FGT into discrete groups, indicating intra-reader variability. Variation between the readers' sets of four boxplots (a–d) indicates inter-reader variability.

Table 1
Lower percent fibroglandular tissue (FGT) in cases of discordance between readers and converted percent FGT.

Reader	Concordant cases		Discordant cases		*P*
	n[a] (%)	Mean % FGT ± SD[b]	*n*[a] (%)	Mean % FGT ± SD[b]	
Reader 1	147/232 (63.4)	25.2 ± 19.5	85/232 (36.6)	14.1 ± 8.1	<0.001
Reader 2	133/232 (57.3)	24.3 ± 19.3	99/232 (42.7)	16.9 ± 12.5	0.002
Reader 3	136/232 (58.6)	27.4 ± 19.4	96/232 (41.4)	12.4 ± 6.5	<0.001
Reader 4	114/232 (49.1)	25.2 ± 18.7	118/232 (50.9)	17.3 ± 14.5	<0.001

[a] *n* = number of cases out of 232 total in which reader qualitative assessment of FGT was concordant or discordant with percent FGT after receiver operating characteristic-based conversion of percent FGT to reader scoring.
[b] SD = standard deviation.

limitations. Wu et al. reported an automated processing time of about 4.5 min for 56-slice sequences [12]; in our study, semi-automated chest wall contouring of T1 weighted sequences that routinely consisted of nearly three times as many slices (>150 slices) took approximately 7 min per case. In a study of 14 patients, Rosado-Toro et al. described a technique for automated whole breast segmentation that used Radial Gradient and Spin Echo sequences not routinely acquired in clinical breast MRI [30]. In addition, inaccuracies in chest wall detection using automated techniques have been reported due to low image contrast or extreme FGT [10,31,32]. van der Welden et al. reported fully automated whole breast segmentation in a study of patients with unilateral breast cancer; however, manual correction of erroneous breast masks had to be performed [33]. Thus, the combination of fast, clinically translatable, and accurate isolation of the breasts from the chest wall without user intervention remains a challenge.

Following breast segmentation, the phantom-based technique for FGT segmentation in this study was fully automated. Direct comparison of FGT segmentation results with other methods is difficult due to differences in MRI data sets, yet technical considerations warrant mention. In our study, BiCal NUC and signal intensity histogram segmentation of 232 breasts from clinical T1 weighted breast MRI required on average 1 min per breast. Wu et al. used an atlas-aided fuzzy C-means method to automatically segment FGT on 120 breasts. Processing time was approximately 30 s for each 56-slice case [12,25], but this approach requires the development of a learned FGT likelihood atlas. In a study by Gubern-Mérida et al. with 100 cases, the authors used an expectation-maximization algorithm to automatically segment FGT and noted oversegmentation of FGT in large breasts due to incomplete bias field correction [17]. Wang et al. described the use of hierarchical support vector machines for automated FGT segmentation, however the technique incorporated non-clinical Dixon sequences and was validated in only 4 cases [23].

In our study, ROC analyses of consensus reads were used to convert quantified percent FGT to reader scoring. The resulting thresholds (percent FGT ≤9.1 = a, 9.1–13.4 = b, 13.5–26.1 = c, >26.1 = d) were similar to the mean percent FGT corresponding to reader assessment in a recent study of breast MRI FGT segmentation (a = 4.6%, b = 8.7%, c = 18.1%, d = 37.4%) [34], as well as thresholds used for automated mammographic breast density assessment [35–37]. Subsequent comparison between converted percent FGT and individual reader assessment showed strong correlations and moderate to substantial agreement, suggesting reader consistency with quantified results. However, the overlapping ranges of percent FGT for each level of reader assessment in Fig. 6, and the variations between the readers' sets of overlapping ranges indicate intra- and inter-reader variability, respectively. Further, reader concordance with percent FGT was less frequent at lower levels of percent FGT (Table 1). A reference for comparison of these results could not be found in the literature, yet these individual reader inconsistencies in grading FGT suggest a need for standardized, quantitative FGT assessment.

The current study has some limitations. First, results were based on the T1 weighted breast MRI protocol and coil design of a single institution. Adjustments to phantom-derived NUC parameters may initially be necessary for use with different acquisition parameters. Second, phantom studies showed that quantified percent FGT was overestimated by 2.1 ± 0.7%, likely due to partial volume effects at the interfaces of phantom FGT and phantom adipose tissue [38]. Third, whole breast segmentation was semi-automated and was the rate limiting step. Fourth, determination of segmentation success on clinical images was subjective. Further development of the current breast and FGT segmentation techniques is ongoing to achieve full automation.

5. Conclusions

In this study, we report a phantom-validated workflow for FGT segmentation on routine T1 weighted breast MRI that, after initial training, demonstrated high accuracy and precision across a wide spectrum of FGT composition. Clinical application of this method of FGT segmentation was successful in 232 processed breasts. Qualitative radiologist FGT assessment showed intra- and inter-reader variability, and consistency with quantified FGT varied with the level of quantified FGT. These findings suggest a need for incorporation of FGT quantification into the clinical routine, which can be achieved without added scan time and may be useful for breast cancer risk estimation and personalized screening protocols.

References

[1] Pike MC, Pearce CL. Mammographic density, MRI background parenchymal enhancement and breast cancer risk. Ann Oncol 2013;24(Suppl. 8):viii37–41.
[2] Wei J, Chan HP, Helvie MA, et al. Correlation between mammographic density and volumetric fibroglandular tissue estimated on breast MR images. Med Phys 2004; 31(4):933–42.
[3] Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92(13):1081–7.
[4] McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006;15(6):1159–69.
[5] Albert M, Schnabel F, Chun J, et al. The relationship of breast density in mammography and magnetic resonance imaging in high-risk women and women with breast cancer. Clin Imaging 2015;39(6):987–92.
[6] Dontchos BN, Rahbar H, Partridge SC, et al. Are qualitative assessments of background parenchymal enhancement, amount of fibroglandular tissue on MR images, and mammographic density associated with breast cancer risk? Radiology 2015; 276(2):371–80.
[7] King V, Brooks JD, Bernstein JL, Reiner AS, Pike MC, Morris EA. Background parenchymal enhancement at breast MR imaging and breast cancer risk. Radiology 2011; 260(1):50–60.
[8] Morris EA, Comstock CE, Lee CH, et al. ACR BI-RADS® magnetic resonance imaging. ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. Reston, VA: American College of Radiology; 2013.
[9] Schrading S, Schild H, Kühr M, Kuhl C. Effects of tamoxifen and aromatase inhibitors on breast tissue enhancement in dynamic contrast-enhanced breast MR imaging: a longitudinal intraindividual cohort study. Radiology 2014;271(1).
[10] Giannini V, Vignati A, Morra L, et al. A fully automatic algorithm for segmentation of the breasts in DCE-MR images. Conf Proc IEEE Eng Med Biol Soc 2010;2010:3146–9.
[11] Milenkovic J, Chambers O, Marolt Music M, Tasic JF. Automated breast-region segmentation in the axial breast MR images. Comput Biol Med 2015;62:55–64.
[12] Wu S, Weinstein SP, Conant EF, Schnall MD, Kontos D. Automated chest wall line detection for whole-breast segmentation in sagittal breast MR images. Med Phys 2013; 40(4):042301.

[13] Klifa C, Suzuki S, Aliu S, et al. Quantification of background enhancement in breast magnetic resonance imaging. J Magn Reson Imaging 2011;33(5):1229–34.
[14] Nie K, Chen JH, Chan S, et al. Development of a quantitative method for analysis of breast density based on three-dimensional breast MRI. Med Phys 2008;35(12):5253–62.
[15] Lin M, Chen JH, Wang X, Chan S, Chen S, Su MY. Template-based automatic breast segmentation on MRI by excluding the chest region. Med Phys 2013;40(12):122301.
[16] Ivanovska T, Laqua R, Wang L, Liebscher V, Volzke H, Hegenscheid K. A level set based framework for quantitative evaluation of breast tissue density from MRI data. PLoS One 2014;9(11), e112709.
[17] Gubern-Mérida A, Kallenberg M, Mann RM, Martí R, Karssemeijer N. Breast segmentation and density estimation in breast MRI: a fully automatic framework. IEEE J Biomed Health Inform 2015;19(1):349–57.
[18] Khalvati F, Gallego-Ortiz C, Balasingham S, Martel AL. Automated segmentation of breast in 3-D MR images using a robust atlas. IEEE Trans Med Imaging 2015;34(1):116–25.
[19] Ortiz CG, Martel AL. Automatic atlas-based segmentation of the breast in MRI for 3D breast volume computation. Med Phys 2012;39(10):5835–48.
[20] Konyer NB, Ramsay EA, Bronskill MJ, Plewes DB. Comparison of MR imaging breast coils. Radiology 2002;222(3):830–4.
[21] Mann RM, Kuhl CK, Kinkel K, Boetes C. Breast MRI: guidelines from the European Society of Breast Imaging. Eur Radiol 2008;18(7):1307–18.
[22] Ledger AE, Scurr ED, Hughes J, et al. Comparison of Dixon sequences for estimation of percent breast fibroglandular tissue. PLoS One 2016;11(3), e0152152.
[23] Wang Y, Morrell G, Heibrun ME, Payne A, Parker DL. 3D multi-parametric breast MRI segmentation using hierarchical support vector machine with coil sensitivity correction. Acad Radiol 2013;20(2):137–47.
[24] Wengert GJ, Pinker-Domenig K, Helbich TH, et al. Influence of fat-water separation and spatial resolution on automated volumetric MRI measurements of fibroglandular breast tissue. NMR Biomed 2016;29(6):702–8.
[25] Wu S, Weinstein SP, Conant EF, Kontos D. Automated fibroglandular tissue segmentation and volumetric density estimation in breast MRI using an atlas-aided fuzzy C-means method. Med Phys 2013;40(12):122302.
[26] Wei CH, Li Y, Huang PJ, Gwo CY, Harms SE. Estimation of breast density: an adaptive moment preserving method for segmentation of fibroglandular tissue in breast magnetic resonance images. Eur J Radiol 2012;81(4):e618–24.
[27] Belaroussi B, Milles J, Carme S, Zhu YM, Benoit-Cattin H. Intensity non-uniformity correction in MRI: existing methods and their validation. Med Image Anal 2006;10(2):234–46.
[28] Mikheev A, Rusinek H, Wiggins G. Non-uniformity normalization using 3D Canny edges and Legendre polynomial approximation of the bias field: validation on 7 T T1W brain images. Proceedings of the 21st Annual Meeting of ISMRM, Salt Lake City; 2013 (abstract 2695).
[29] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33(1):159–74.
[30] Rosado-Toro JA, Barr T, Galons JP, et al. Automated breast segmentation of fat and water MR images using dynamic programming. Acad Radiol 2015;22(2):139–48.
[31] Ertas G, Gulcur HO, Osman O, Ucan ON, Tunaci M, Dursun M. Breast MR segmentation and lesion detection with cellular neural networks and 3D template matching. Comput Biol Med 2008;38(1):116–26.
[32] Twellmann T, Lichte O, Nattkemper TW. An adaptive tissue characterization network for model-free visualization of dynamic contrast-enhanced magnetic resonance image data. IEEE Trans Med Imaging 2005;24(10):1256–66.
[33] van der Velden BH, Dmitriev I, Loo CE, Pijnappel RM, Gilhuijs KG. Association between parenchymal enhancement of the contralateral breast in dynamic contrast-enhanced MR imaging and outcome of patients with unilateral invasive breast cancer. Radiology 2015;276(3):675–85.
[34] Ha R, Mema E, Guo X, et al. Quantitative 3D breast magnetic resonance imaging fibroglandular tissue analysis and correlation with qualitative assessments: a feasibility study. Quant Imaging Med Surg 2016;6(2):144–50.
[35] Singh T, Sharma M, Singla V, Khandelwal N. Breast density estimation with fully automated volumetric method: comparison to radiologists' assessment by BI-RADS categories. Acad Radiol 2016;23(1):78–83.
[36] van der Waal D, den Heeten GJ, Pijnappel RM, et al. Comparing visually assessed BI-RADS breast density and automated volumetric breast density software: a cross-sectional study in a breast cancer screening setting. PLoS One 2015;10(9), e0136667.
[37] Youk JH, Gweon HM, Son EJ, Kim JA. Automated volumetric breast density measurements in the era of the BI-RADS fifth edition: a comparison with visual assessment. AJR Am J Roentgenol 2016;206(5):1056–62.
[38] Lu LJ, Nishino TK, Johnson RF, et al. Comparison of breast tissue measurements using magnetic resonance imaging, digital mammography and a mathematical algorithm. Phys Med Biol 2012;57(21):6903–27.