Received:
19 July 2019

Revised:
16 October 2019

Accepted:
30 October 2019

<https://doi.org/10.1259/bjr.20190639>

Cite this article as:

Goodburn R, Kousi E, Macdonald A, Morgan V, Scurr E, Reddy M, et al. An automated approach for the optimised estimation of breast density with Dixon methods. *Br J Radiol* 2020; **93**: 20190639.

SHORT COMMUNICATION

An automated approach for the optimised estimation of breast density with Dixon methods

¹ROSIE GOODBURN, MSc, ¹EVANTHIA KOUSI, PhD, ²ALISON MACDONALD, MSc, ²VERONICA MORGAN, MSc, ²ERICA SCURR, MSc, ³MAMATHA REDDY, FRCR, ³LOUISE WILKINSON, FRCR, ²ELIZABETH O'FLYNN, FRCR, ²ROMNEY POPE, FRCR, ²STEVEN ALLEN, FRCR and ¹MARIA ANGÉLICA SCHMIDT, PhD

¹CRUK Cancer Imaging Centre, The Institute of Cancer Research and Royal Marsden Foundation Trust, London, United Kingdom

²The Royal Marsden NHS Foundation Trust, Sutton, United Kingdom

³St George's University Hospitals NHS Foundation Trust, London, United Kingdom

Address correspondence to: Ms Rosie Goodburn
E-mail: rosie.goodburn@icr.ac.uk

Objective: To present and evaluate an automated method to correct scaling between Dixon water/fat images used in breast density (BD) assessments.

Methods: Dixon images were acquired in 14 subjects with different T_1 weightings (flip angles, FA, $4^\circ/16^\circ$). Our method corrects intensity differences between water (W) and fat (F) images via the application of a uniform scaling factor (SF), determined subject-by-subject. Based on the postulation that optimal SFs yield relatively featureless summed fat/scaled-water ($F + W_{SF}$) images, each SF was chosen as that which generated the lowest 95th-percentile in the absolute spatial-gradient image-volume of $F + W_{SF}$. Water-fraction maps were calculated for data acquired with low/high FAs, and BD

(%) was the total percentage water within each breast volume.

Results: Corrected/uncorrected BD ranged from, respectively, 10.9–71.8%/8.9–66.7% for low-FA data to 8.1–74.3%/5.6–54.3% for high-FA data. Corrected metrics had an average absolute increase in BD of 6.4% for low-FA data and 18.4% for high-FA data. BD values estimated from low- and high-FA data were closer following SF-correction.

Conclusion: Our results demonstrate need for scaling in such BD assessments, where our method brought high-FA and low-FA data into closer agreement.

Advances in knowledge: We demonstrated a feasible method to address a main source of inaccuracy in Dixon-based BD measurements.

INTRODUCTION

Mammographic density is a strong independent risk factor for the development of breast cancer¹; high mammographic density is associated not only with a higher risk of developing breast cancer, but also with a lower probability of having breast cancer detected by X-ray Mammography (XRM).² Moreover, for subjects of high breast density (BD), it may be more effective to use MRI rather than of XRM a part of a screening programme.^{3,4} The assessment of BD therefore has a potential role in the stratification of subjects into different screening regimes according to breast cancer risk levels.

The measurement of percentage mammographic-density is based on the relative area occupied by the fibroglandular breast tissue (parenchyma) within a mammogram and depends on compression characteristics. MRI-based measurements, shown to correlate with BD measured by XRM,¹ are derived from three-dimensional data sets and thus potentially offer higher accuracy in BD assessments.

The MRI Dixon method⁵ for fat-water separation is particularly suited to BD assessment, as it produces separate “water”/“fat” images, used as surrogates for fibroglandular/adipose tissues. In theory, Dixon-based approaches can account for partial volume effects in low resolution images by obtaining a water fraction for every voxel.⁶ However, for BD measurements relying on water-fraction estimates, the fact that voxels containing either 100% water or fat do not yield the same signal intensity is a fundamental problem.⁷ Corrections based on signal intensity over specific regions^{7,8} and test-object experiments⁶ have been employed. Because Dixon images can be generated with different contrast characteristics, and because proton density/relaxation times may vary over breast parenchyma, any scaling of water signal should ideally be undertaken on a subject-by-subject basis.

Here, we propose and implement an automated method to compensate for the difference between water and fat signals in MRI-based BD measurements employing Dixon-based

Figure 1. Our correction algorithm demonstrated with a simulated water and fat Dixon images for a numerical phantom based on signal and noise values in low- and high-FA breast images. In the top row, the fat and SF-corrected water images for high-FA data are displayed with the corresponding water-fraction map. The plots below show the intensity values of the profiles through fat (orange), water (blue), and water fraction (grey), where uncorrected data are represented with solid lines and corrected data with dotted lines. SFs of 4.5 and 1.8 were generated using our automated method for high- and low-FA data, respectively. FA, flip angle; SF, scaling factor.

methods. We evaluate this method in a group of subjects, and consider the impact of this correction on MRI-based BD assessment.

METHODS

MRI-based breast density measurements

For three-dimensional MRI, two approaches are used to measure BD. The first involves allocating all voxels within breast to either fat or parenchyma.^{9–12} In this approach, higher spatial resolution implies higher precision. The second approach uses Dixon methods⁵ to calculate the percentage water (or fat) in each voxel. In the latter approach, partial volume effects are accounted for making possible measurements on low resolution data sets.⁶ However, Dixon sequences can be employed with different contrast mechanisms and it is incorrect to assume that the same volume of pure adipose and pure fibroglandular tissue will produce the same signal intensity. In fact, while adipose tissue appears quite uniform over the breast, fibroglandular tissue presents anatomical structures in T_1 weighted images. Proton-density weighted images are thus preferred for BD measurements.⁷

Figure 1 shows a numerical test object with simulated signal and noise, containing a gradient between two regions of 100% water and 100% fat. If water and fat yielded equivalent signals, the sum of water and fat images would be uniform, and the water-fraction correctly estimated. However, as water and fat intensities are

different in this model, the estimated water-fraction is distorted. Although we expect the error to be relatively small for voxels containing mostly fat or water, low resolution data sets are more severely affected, since these are more affected by partial volume effects.

We propose to correct this difference by multiplying water images (W) by a uniform scaling factor (SF) that is determined automatically for each data set. By considering that breast is occupied by either fat or water (parenchyma), the sum of fat (F) and SF-corrected water ($W_{SF} = SF \cdot W$) images should ideally produce an image ($F + W_{SF}$) with few intensity variations within the breast (Figure 2). Our calculation of SF thus aims to identify a value that generates a relatively featureless $F + W_{SF}$ image, achieved by calculating the gradient of intensity over the $F + W_{SF}$ image, and minimizing it (see *Data Processing*). Figure 1 demonstrates how this correction acts to rectify the water fraction along the water-fat gradient so that it varies linearly, as expected.

Clinical study

As part of research investigating the characteristics of breast parenchyma in populations at different levels of breast-cancer risk, subjects were invited for MRI examinations following the reporting of negative XRM at The Royal Marsden. Inclusion criteria were: radiologically healthy breast tissue as assessed by XRM; age range 39.5–50.5; pre-menopausal status; able to attend

Figure 2. Top row (same windowing): sum of the Dixon water and fat images (high FA, right breast) for an example subject slice location where no SF is applied (a) and an SF of 4.9 (generated using our automated method) is applied (b). Bottom row: %W maps (c, d) corresponding to the above images. SF, scaling factor.

MRI screening within 6 weeks of XRM. Exclusion criteria were: previous breast cancer diagnosis and/or treatment; hormonal treatment; bilateral Salpingo-Oophrectomy. All subjects provided written consent for this institutional review board-approved (14/LO/1908) single-centre study, and Dixon data for the first 14 subjects were used for the work described in this article.

All subjects underwent MRI at 3 T (mDixon, Achieva, Philips Healthcare) using a 16-channel breast coil in a prone position. Dixon water and fat images were acquired for two levels of T_1 weighting [flip angles (FA) 16° and 4° ; repetition time 3.60 ms; two averages; slice thickness 2 mm; acquisition matrix 172×172 ; reconstruction matrix 432×432 ; parallel imaging (SENSE) acceleration factor 2].

Data processing

All data processing was performed using MATLAB (MathWorks, Natick, MA).

Breast-Mask construction

A semi-automated recursive three-dimensional region-growing algorithm¹³ was applied to summed water and fat images to

generate breast masks. User interaction was required to draw separate boxes (projected axially along the superior-inferior direction) to enclose each breast (Figure 3). Following smoothing to remove fine detail, the region-growing algorithm was applied in each breast volume separately, seeding from medial, anterior high-signal pixels. Since the resulting binary mask may include rib/liver or miss dense tissue near the base of the breast, segmentations were edited at this point with a three-dimensional paint/erase tool in ITK-SNAP (v. 3.60).¹⁴ Finally, morphological image processing was applied to smooth and remove islands/holes (opening/closing) from the mask and exclude skin/chest-wall (erosion). For any given subject, the same mask was used throughout this investigation.

Data correction

Our software performs the following functions for left and right masked breast images separately, with SFs iteratively tested in steps of 0.1 from 1.0 to 6.0:

- (1) Multiply water images by SF and sum with fat images.
- (2) Calculate (voxelwise) maps of the absolute value of the image-intensity spatial gradient, as:

Figure 3. Right and left breast masks constructed for an example subject. Every 17th slice location is shown from the first (superior, a) to last (inferior, h) slices of the masks, overlaid on Dixon water images.

$$|G(r)| = \sqrt{G_x(r)^2 + G_y(r)^2 + G_z(r)^2} \quad (1)$$

where $G_x(r)/G_y(r)/G_z(r)$ is spatial gradient of the $F + W_{SF}$ image in the $x/y/z$ dimension.

- (3) Output SF which produces the lowest 95th-percentile voxel-value of the gradient magnitude histogram.

Our approach does not presume every spatial gradient present in the combined $F + W_{SF}$ image arises due to intensity differences between water and fat—some gradients may result from proton-density/T1 variations or from random noise. However, as the SF is increased, only gradients associated with the mismatch between water and fat signals are expected to initially decrease, before increasing as scaled water intensity is brought above that of fat. Therefore, we hypothesise that minimising the 95th-percentile

value of the gradient histogram will be sufficient to minimise differences in intensity between water and fat.

Breast-density calculation

Voxelwise water-fraction maps for both uncorrected (% W) and corrected (% W_{SF}) water images were calculated via:

$$\%W(r) = 100 \times W(r) / (W(r) + F(r))$$

$$\%W_{SF}(r) = 100 \times W_{SF}(r) / (W_{SF}(r) + F(r)) \quad (2)$$

where $W(r)/W_{SF}(r)$ and $F(r)$ represent the intensity of an uncorrected/corrected water and fat voxels, respectively. BD was estimated as the total water-percentage within each masked volume.

Statistical analysis

Differences and correlations between the left and right breasts in volume, SF, and BD were evaluated using Mann–Whitney U tests (taking p -values < 0.05 as significant) and Pearson coefficients.

RESULTS

Corrected/uncorrected BD for the population studied ranged from, respectively, 10.9–71.8%/8.9–66.7% for low-FA data and 8.1–74.3%/5.6–54.3% for high-FA data. Correction consistently increased BD and the effect of correction was larger for data with higher T_1 weighting, with an average absolute increase of 6.4% for low-FA data and 18.4% for high-FA data.

As indicated in Table 1, volume and BD showed no statistical differences between left and right breasts ($p \gg 0.05$) and indicated high levels of left-right correlation (Pearson > 0.97). Left-right differences between SFs were more significant ($p < 0.05$ for low-FA data), where this may be explained by systematic asymmetry in patient positioning within the breast coil, leading to different sensitivity distributions for each side.

Table 1. Comparison of results between right and left breasts

	Left breast	Right breast	Mann–Whitney U p -value	Pearson coefficient
Total breast volume (cc)	1424 ± 521	1331 ± 553	0.87	0.99
Low flip-angle data				
Scaling factor	1.90 ± 0.10	1.60 ± 0.15	0.02*	0.89
Breast density (%)	25.4 ± 9.4	21.4 ± 8.7	0.96	0.99
Corrected breast density (%)	33.2 ± 10.1	28.3 ± 8.6	0.66	0.99
High flip-angle data				
Scaling factor	4.65 ± 0.20	4.15 ± 0.30	0.09	0.95
Breast density (%)	19.6 ± 7.8	15.8 ± 6.6	0.86	0.98
Corrected breast density (%)	41.4 ± 8.8	37.6 ± 6.8	0.68	0.99
(Median ± median absolute deviation)				

All Pearson coefficients were significant with $p < 0.001$.

Figure 4. (A-B) Bar graph of SFs calculated across 14 subjects for low- and high-FA Dixon images. Low-FA SFs ranged from 1.2 to 1.9 and high-FA SFs ranged from 1.6 to 4.9. The solid and dashed lines indicate the median $\pm$ median absolute deviation. (C-D) Plot showing the range of estimated BD for an example subject calculated with SFs from 1.0 to 6.0, for low- and high-FA data. SFs that were generated for this subject are shown in red. Dashed lines indicate the potential uncertainty for a $\pm 10\%$ error on the SF, resulting in a $\pm 2\%$ uncertainty in estimated BD. BD, breast density; FA, flip angle; SF, scaling factor.

All $F + W_{SF}$ data sets were visually inspected and no errors in determination of SF were apparent. Figure 4A/B shows the range of SFs calculated across subjects (median $\pm$ median absolute deviation for low- and high-FA data were respectively right breasts 1.60 ± 0.15 , 4.15 ± 0.30 for right breasts and 1.90 ± 0.30 , 4.65 ± 0.65 for left breasts). Figure 4C/D demonstrates the steepness of the change in BD with SF for one particular data set; taking a $\pm 10\%$ uncertainty for the determination of the SF produced an absolute error of $\sim \pm 2\%$ in BD. Therefore, the step size (0.1) in our iterative algorithm is expected to be appropriate for accurate correction.

Figure 5A/B illustrates the relation between BD values calculated with and without SFs. For most subjects, the absolute increase in BD was $>5\%$ ($1.6\text{--}9.4\%/2.5\text{--}24.7\%$ for low-/high-FA data, respectively). Figure 5C/D compares BD values calculated from low- and high-FA data; BD estimated from the different T_1 weightings is in closer agreement following SF-correction (particularly for higher values).

DISCUSSION

The use of Dixon water-fat data for BD assessment relies on voxelwise water-fraction estimates, where partial volume effects introduce a source of inaccuracy when parenchyma intensity in water images is not equivalent to the same volume of adipose in fat images. Since this effect is more pronounced in partial volumes, its correction should allow for lower-resolution sequences, bringing increased signal-to-noise ratio and shorter scan times.

Correction of Dixon breast images via uniform SFs has been employed previously,^{7,8} where region of interest-based approaches

were used to estimate the ratio between the intensity of regions of (assumed) pure parenchyma and fat. Our approach, based on minimising the spatial gradient of summed fat-water images, is expected to hold two main advantages over these methods. First, automation removes errors occurring due to intra- and interuser variation. Second, our method does not require that such pure regions in fat and water images are present. Our findings suggest that the automated scaling method presented here is a feasible approach for improving accuracy in BD estimates with Dixon methods.

It should be noted that our approach does not correct for variations in image intensity due to T_1 changes across parenchyma. High-FA images are expected to be less "flat" than low-FA images due to the added T_1 weighting, and this effect cannot be removed via a simple uniform scaling.¹⁵ A potential pitfall of our correction method is that it will generate water and fat images with different respective weightings of noise/noise-floor level. High-FA data will be more strongly affected, where water signal requires a greater degree of scaling, potentially resulting in artificially high BD estimates as noise-floor is rectified. Noise and noise-floor propagation may be minimised, therefore, by use of low-FA acquisitions to yield smaller SFs or low-resolution data-sets with higher signal-to-noise ratios. Therefore, our work adds evidence to the need to employ Dixon sequences with lower T_1 weighting in order to achieve more accurate results.

Although MRI-based BD measurements have been proposed as a gold-standard,^{3,4} our results demonstrate signal differences between water and fat introduce significant systematic errors. We evaluated our method by assuming that corrected low-FA data, with little T_1 weighting, provided the most accurate results.

Figure 5. Scatter plots of BD for values estimated for the different methods: (A, B) BD estimated with SF vs without SF; (C, D) BD estimated using high-FA acquisitions vs low-FA acquisitions, with trendlines indicated in light grey. BD, breast density; FA, flip angle; SF, scaling factor.

It was noted that the three subjects of lowest BD (Figure 4A/B) yielded the lowest SFs; 26–32%/35–68% lower than median SFs for low-/high-FA data, respectively. It is thus possible that our automated method underestimates SFs and hence BD when parenchyma signal is sparse and low. However, these subjects would remain within the low-BD range, for which XRM is an effective screening tool.² Our work will therefore still have high-impact potential benefits for the target population, for whom MRI is the best screening technology: females of high BD. Their risk levels are higher and thus it is more important to employ the most accurate methods for BD assessment.

In conclusion, our results demonstrate the need for scaling in BD estimates, where application of SFs brought high-FA and low-FA data into closer agreement. The highest impact of correction is in high-BD data sets. The method we propose is feasible, and could be included in the workflow for Dixon-based BD measurements, providing a higher level of standardization to longitudinal

studies and multicentre trials. Future work will investigate the clinical value of our correction method.

ACKNOWLEDGMENT

CRUK and EPSRC support to the Cancer Imaging Centre at ICR and RMH in association with MRC & Dept of Health C1060/A10334, C1060/A16464 and NHS funding to the NIHR Biomedical Research Centre and the Clinical Research Facility in Imaging. This report is independent research funded partially by the National Institute for Health Research. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health.

FUNDING

This report is independent research funded partially by the National Institute for Health Research.

REFERENCES

1. Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast tissue composition and susceptibility to breast cancer. *J Natl Cancer Inst* 2010; **102**: 1224–37. doi: <https://doi.org/10.1093/jnci/djq239>
2. Buist DSM, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40–49 years. *JNCI Journal of the National Cancer Institute* 2004; **96**: 1432–40. doi: <https://doi.org/10.1093/jnci/djh269>
3. Berg WA, Zhang Z, Lehrer D, et al. ACRIN 6666 Investigators. detection of breast cancer with addition of annual screening ultrasound, or a single screening MRI to mammography in women with elevated breast cancer risk. *JAMA* 2012; **307**: 1394–404.
4. Kuhl C, Weigel S, Schrading S, Arand B, Bieling H, König R, Heribert B, Roy K, et al. Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. *JCO* 2010; **28**: 1450–7. doi: <https://doi.org/10.1200/JCO.2009.23.0839>
5. Dixon WT. Simple proton spectroscopic imaging. *Radiology* 1984; **153**: 189–94. doi: <https://doi.org/10.1148/radiology.153.1.6089263>
6. Boyd N, Martin L, Chavez S, Gunasekara A, Salleh A, Melnichouk O, et al. Breast-tissue composition and other risk factors for breast cancer in young women: a cross-sectional study. *Lancet Oncol* 2009; **10**: 569–80. doi: [https://doi.org/10.1016/S1470-2045\(09\)70078-6](https://doi.org/10.1016/S1470-2045(09)70078-6)
7. Ledger AEW, Scurr ED, Hughes J, Macdonald A, Wallace T, Thomas K, et al. Comparison of Dixon sequences for estimation of percent breast fibroglandular tissue. *PLoS One* 2016; **11**: e0152152. doi: <https://doi.org/10.1371/journal.pone.0152152>
8. Doran SJ, Hipwell JH, Denholm R, Eiben B, Busana M, Hawkes DJ, et al. Breast MRI segmentation for density estimation: do different methods give the same results and how much do differences matter? *Med Phys* 2017; **44**: 4573–92. doi: <https://doi.org/10.1002/mp.12320>
9. Khazen M, Warren RML, Boggis CRM, Bryant EC, Reed S, Warsi I, et al. A pilot study of compositional analysis of the breast and estimation of breast mammographic density using three-dimensional T1-weighted magnetic resonance imaging. *Cancer Epidemiology Biomarkers & Prevention* 2008; **17**: 2268–74. doi: <https://doi.org/10.1158/1055-9965.EPI-07-2547>
10. Nie K, Chen J-H, Chan S, Chau M-KI, Yu HJ, Bahri S, et al. Development of a quantitative method for analysis of breast density based on three-dimensional breast MRI. *Med Phys* 2008; **35**: 5253–62. doi: <https://doi.org/10.1118/1.3002306>
11. Klifa C, Carballido-Gamio J, Wilmes L, Laprie A, Shepherd J, Gibbs J, et al. Magnetic resonance imaging for secondary assessment of breast density in a high-risk cohort. *Magn Reson Imaging* 2010; **28**: 8–15. doi: <https://doi.org/10.1016/j.mri.2009.05.040>
12. Ertas G, Reed S, Leach MO. Comparison of breast density assessments based on interactive thresholding and automated fast fuzzy C-Means clustering in three-dimensional MR imaging. *FMBE Proc* 2010; **25**: 1893–6.
13. regionGrowing KD. MathWorks file exchange. Available from: <https://uk.mathworks.com/matlabcentral/fileexchange/32532-region-growing-2d-3d-grayscale2011>.
14. Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, Gimpel Smith R, Ho S C, GJ, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* 2006; **31**: 1116–28. doi: <https://doi.org/10.1016/j.neuroimage.2006.01.015>
15. Lui CY, McKenzie CA, Yu H, Brittain JH, Reeder SB. Fat quantification with ideal gradient echo imaging: correction of bias from T1 and noise. *Magn Reson Med* 2007; **58**: 354–64.