



Published in final edited form as:

J Comput Assist Tomogr. 2016 ; 40(3): 428–435. doi:10.1097/RCT.0000000000000372.

Diffusion Weighted Breast MRI: A Semi-automated Voxel Selection Technique Improves Inter-Reader Reproducibility of Apparent Diffusion Coefficient Measurements

Habib Rahbar, MD^a, Brenda F. Kurland, PhD^b, Matthew L. Olson, MS^a, Averi E. Kitsch, BS^a, John R Scheel, MD, PhD^a, Xiaoyu Chai, MS^b, Joshua Usoro, BS^a, Constance D. Lehman, MD, PhD^a, and Savannah C. Partridge, PhD^a

^aUniversity of Washington, Seattle Cancer Care Alliance, Department of Radiology, Breast Imaging Section, 825 Eastlake Avenue East, P.O. Box 19023, Seattle, WA 98109–1023, USA

^bFred Hutchinson Cancer Research Center, Clinical Research Division, 1100 Fairview Avenue North, P.O. Box 19024, Seattle, WA 98109-1024

Abstract

Objective—To determine whether a semi-automated voxel selection technique improves inter-reader reproducibility for breast apparent diffusion coefficient (ADC) measurements.

Methods—Three readers retrospectively performed ADC measurements of 31 breast lesions (16 malignant, 15 benign) and contralateral normal tissue in 26 women both unassisted (manual method) and assisted by a semi-automated software tool that excludes voxels below a dynamically specified signal intensity threshold. Reproducibility between readers for each technique was assessed by Bland-Altman analysis and concordance correlation coefficients (CCCs).

Results—Differences between readers' measured ADCs of lesions were smaller with the semi-automated tool versus the manual method. CCCs for each reader pair were greater with the semi-automated tool for lesions (mean CCC difference= 0.11; CI= 0.04–0.26). For normal tissue, reader agreement was lower than for lesions, and did not differ based on software tools (mean CCC difference= 0.00; 95% CI= –0.14–0.13).

Conclusions—A semi-automated voxel selection tool can improve inter-reader reproducibility of breast lesion ADC measures.

Keywords

Diffusion weighted imaging; breast; apparent diffusion coefficient; reproducibility; region of interest

Corresponding Author: Habib Rahbar, MD, University of Washington, Seattle Cancer Care Alliance, Department of Radiology, Breast Imaging Section, 825 Eastlake Avenue East, P.O. Box 19023, Seattle, WA 98109–1023, USA, ; Email: hrahbar@uw.edu, Phone: 206-288-1257, Fax: 206-288-6573.

Present Address (author who has moved since work was done):

Brenda F. Kurland, PhD, University of Pittsburgh Cancer Institute, UPCI Biostatistics Facility, Sterling Plaza, Suite 325, 201 N Craig St, Pittsburgh, PA 15213

Previous presentations

This study was presented as electronic poster at the 2013 annual ISMRM meeting in Salt Lake City, Utah, USA on April 23rd, 2013.

Introduction

Diffusion weighted magnetic resonance imaging (DWI) has shown promise for improving the accuracy and positive predictive value of standard clinical dynamic contrast-enhanced (DCE) MRI¹. DWI provides indirect information on lesion cellularity through the measurement of water diffusivity by apparent diffusion coefficient (ADC) calculation². On average, malignant breast lesions display lower ADC values than benign lesions³⁻⁸, making ADC a potentially useful parameter for discriminating between benign and malignant pathology. At present, technical challenges such as inconsistent ADC quantification approaches limit widespread use of DWI in routine clinical breast MRI. Breast lesion ADC measurements typically require that a user define a region of interest (ROI) on the DW images and ADC map corresponding to the suspicious MRI finding. However, the use of either manually defined ROIs or pre-defined shaped ROIs to encompass irregularly shaped or less well-defined breast lesions, such as non-mass enhancement (NME), can be technically challenging and prone to sampling error. Similarly, wide variations in the amount and pattern of normal fibroglandular tissue cause reference normal tissue ROI measurements to be overly dependent on ROI positioning, which in turn makes these measurements susceptible to partial volume averaging errors from the unintended inclusion of intervening or adjacent adipose tissue.

As DWI is increasingly gaining acceptance for clinical use for breast imaging applications and currently is being studied in large multi-site trials^{9, 10}, it is critical that a standardized and reproducible method for ADC measurement be established that facilitates consistency across readers. Accordingly, we have developed a semi-automated approach to calculate ADC values for breast lesions and normal breast parenchyma. This technique incorporates interactive thresholding to select appropriate voxels on the DW images based on signal differences between breast lesions, normal parenchyma, and fat. The purpose of this study was to test whether this semi-automated technique improves the reproducibility of quantitative ADC measurements of breast lesions and normal fibroglandular tissue between readers over a standard manual ROI method.

Methods

This study compared inter-reader breast ADC measurement variability between 3 readers, each using both the semi-automated voxel selection and manual ROI techniques.

Subjects

The study was institutional review board (IRB) approved and Health Insurance Portability and Accountability Act (HIPAA) compliant. All patients provided written informed consent as part of a larger prospective DWI breast MRI trial. Included patients must have had at least one suspicious lesion (prospectively reported as Breast Imaging-Reporting and Data System [BIRADS] category 4 or 5) identified on a breast MRI at 3 tesla (T) including both dynamic contrast-enhanced (DCE) and DWI sequences that subsequently underwent core needle and/or surgical sampling. Cases were consecutively selected from eligible study subjects who underwent breast MRI from October 2010 to June 2011 to include a similar proportion of benign and malignant pathologies. In order to accurately measure contralateral normal

breast parenchyma, patients who had a contralateral mastectomy, a contralateral suspicious breast lesion, or fatty breast were excluded.

Image Acquisition

All scans were performed on a Philips 3T Achieva scanner (Philips Healthcare, Best, The Netherlands) using a 16-channel breast coil. MR examinations included both DCE and DWI sequences with images acquired in the axial orientation. DCE was performed using a 3D T1-weighted gradient echo acquisition with parallel imaging, active fat suppression, and the following parameters: repetition time/echo time = 5.9ms/3ms, flip angle = 10°, SENSE reduction factors = 2.7 right-to-left / 2.0 superior-to-inferior, number of averages = 1, matrix size = 420 × 660, field-of-view = 22 cm anterior-to-posterior × 33 cm right-to-left, slice thickness = 1.3 cm, acquired voxel size = 0.5 × 0.5 × 1.3 mm. DCE scan time was 2:57 minutes per acquisition. The DCE sequence included a pre-contrast scan and three post-contrast scans after intravenous injection of 0.1 mmol/kg gadoteridol (Prohance, Bracco Diagnostics Inc., Princeton, NJ). Following DCE, DWI was performed using a single shot spin-echo echo-planar imaging acquisition with the following parameters: repetition time/echo time = 5336ms/61ms, flip angle = 90°, SENSE reduction factor = 3 anterior-to-posterior, number of averages = 2, matrix size = 240 × 240, field-of-view = 36 cm anterior-to-posterior and 36 cm right-to-left, slice thickness = 5 mm, gap = 0, acquired voxel size = 1.5×1.5×5 mm. Diffusion gradients were applied in six directions with b values of 0, 100, and 800 s/mm². The total DWI acquisition time was 3:28 minutes.

Image Analysis

A commercially-available 3D affine transformation algorithm (Diffusion Registration tool, Philips Healthcare, Best, The Netherlands) was first used for DWI image registration. DWI was analyzed with in-house software written in Java language and incorporating open source image analysis tools (ImageJ, National Institutes of Health). Voxel-based ADC maps were calculated by fitting the monoexponential function

$$S = S_0 \times e^{-b \times ADC} \quad (1)$$

where S is the signal intensity on the DW image after application of the diffusion gradient and S_0 is the signal intensity acquired at $b = 0$ s/mm². Both manual and semi-automated ROI measures were performed offline using custom software implemented on the Image J platform.

Measurements were performed by three readers with varying levels of DWI experience who were blinded to lesion pathologies. The readers included two radiologists specializing in breast imaging (H.R., 5 years of MRI experience and J.R.S., 4 years of MRI experience) and a research scientist skilled in quantitative breast MRI measurements (M.O., 2 years of breast DWI experience). Working at separate workstations and using annotated DCE MR images for reference, readers first drew ROIs on the DW images utilizing the manual method to measure ADCs for lesions and contralateral normal tissue for all study subjects, followed by the semi-automated method for all study subjects. For lesions, ROIs were first drawn on the high b (b=800 s/mm²) DW images where lesion contrast (with respect to fibroglandular and

adipose tissues) tends to be highest. For contralateral normal fibroglandular tissue, ROIs were drawn on the unweighted $b=0$ s/mm² images where contrast between fibroglandular and adipose tissues is highest. All ROIs were propagated onto corresponding ADC maps and the mean of voxel ADC values within the ROIs were calculated.

For the manual method, readers drew freehand ROIs outlining the lesion of interest and normal fibroglandular tissue in the contralateral breast in the following fashion and as described previously⁴: First, the location of the lesion on DWI was determined by comparing the appearance of the lesion and fibroglandular tissue landmarks on the DCE-MRI image. The boundaries of each ROI then were drawn so as to include the greatest contiguous area of high DWI signal on the $b = 800$ s/mm² images corresponding to the lesion in question, taking care to avoid hypo-intense voxels that represented either normal fibroglandular or adipose tissue. Readers also avoided regions of high DWI signal due to T2 effects within a lesion (i.e. T2 shine-through, potentially reflecting fluid content from cystic spaces or necrosis) by verifying the appearance of the lesion on the T2-weighted $b = 0$ s/mm² image. Next, in the contralateral breast each reader also carefully drew an ROI encompassing as large a contiguous section of normal tissue as possible at the level of the nipple to measure the ADC values of normal fibroglandular tissue for each patient, taking care to exclude adjacent or interspersed voxels of very low signal intensity likely representing adipose tissue.

To perform semi-automated ADC measurements, a software algorithm was developed that allowed the reader to dynamically adjust the signal threshold on the DW image for lesion measures to mask out voxels exhibiting signal lower than the specified threshold¹¹. This allowed the reader to subjectively exclude normal fibroglandular and/or adipose tissue adjacent to the lesion in real time and prior to drawing an ROI. Subsequently, each reader drew a generous ROI encompassing the entire lesion, and the ROI and mask were then applied to the corresponding ADC map, with only the unmasked voxels within the ROI included in the ADC measurement. The readers repeated this process to perform semi-automated ADC measures of contralateral normal fibroglandular tissue, dynamically adjusting the signal threshold on the $b=0$ s/mm² image and drawing a generous ROI to encompass the maximal region of normal fibroglandular tissue. Examples of the variations in ROI definition between the three readers for a NME, mass, and normal tissue using both manual and semi-automated methods, are provided in Figures 1, 2, and 3, respectively.

Statistical Analysis

Bland-Altman plots were constructed for inter-reader comparison of ADC measures obtained using the manual and semi-automated techniques and visual inspection of ADC reproducibility between readers¹². Agreement between manual and semi-automated ADC measures were assessed by paired Wilcoxon signed-rank tests for both lesions and normal breast tissue. Concordance correlation coefficients (CCC)^{13–15} were calculated for reader pairs to assess inter-reader reproducibility of ADC measurements using manual and semi-automated ROI methods. Analyses were performed using R software (version 3.1.3; R Foundation for Statistical Computing) and SAS version 9.4 (SAS Institute, Inc.).

Results

The study cohort included 31 biopsy-confirmed lesions (15 benign, 16 malignant) in 26 women; five women each had two ipsilateral lesions that were evaluated separately for the study. Lesion characteristics and pathology features are summarized in Table 1. The 15 benign lesions had a median size of 9 (range: 5 – 67) mm, with eight presenting as a mass on DCE MRI and seven as NME. The 16 malignant lesions had a median size of 12.5 (range: 5 – 105) mm with 10 presenting as masses and six as NME. Descriptive summaries of ADC measures for all readers and methods are given in Table 2. Of note, the median and range of measured ADC values were comparable across all readers and methods, particularly when measuring lesions. All three readers subjectively described the semi-automated method as faster and less tedious to perform than the manual method since it allowed inappropriate voxels to be efficiently masked out rather than having to be actively avoided while drawing the ROI.

Comparison of manual versus semi-automated ADC measures

Bland-Altman plots comparing ADC measures between the manual and semi-automated methods demonstrated high agreement without systematic bias between all three readers for lesions (Figure 4A), regardless of morphology (mass vs. NME, Figure 5). Wilcoxon signed-rank tests, conducted separately for each reader, found no systematic differences in lesion ADC measures from manual and semi-automated methods (median absolute difference range = 0.005 to 0.009×10^{-3} mm²/s, $p > 0.5$, Table 2). The readers also attained excellent agreement between manual and semi-automated methods for measuring normal tissue (Figure 4B).

Inter-reader reproducibility of lesion ADC measures

Bland-Altman plots comparing inter-reader ADC measures demonstrate that the limits of agreement for lesion ADC measures were narrower for the semi-automated method (Figure 6B) compared to the manual method (Figure 6A). The overall concordance correlation coefficient (CCC)¹⁶ among the three readers was greater with the semi-automated method compared to the manual method (Table 3). The overall CCC for the three reader pairs was 0.97 (95% confidence interval [CI] = 0.95 to 0.99) using the semi-automated method versus 0.86 (95% CI = 0.71 to 0.95) for manual ROIs. The mean CCC difference was 0.11 (CI = 0.04 to 0.26), indicating that the semi-automated method was significantly more reproducible for ADC calculations of lesions than the manual method. Inter-reader reproducibility (CCC value) and improvement with the semi-automated method was similar for both lesion subtypes, masses and NMEs (Table 3).

Inter-reader reproducibility of normal tissue ADC measures

For contralateral normal parenchyma ADC measures, Bland-Altman plots (Figure 6C and 6D) demonstrate that agreement between readers was similar for the manual and the semi-automated method. The overall CCC was 0.76 (95% CI = 0.49 to 0.88) for the manual method and 0.76 (95% CI = 0.61 to 0.87) for the semi-automated method (Table 3). Thus, inter-reader agreement for normal tissue ADC measures was lower than that for lesion ADC measures, and was not improved by use of semi-automated software tools.

Discussion

As interest in clinical use of DWI for breast MRI applications increases, a standardized approach to ADC measurement is vital for minimizing variability in ADC measures among different readers. In this study, we assessed the potential improvement in reproducibility of ADC measurements of breast lesions and normal tissue by using a semi-automated DWI signal thresholding tool that allows dynamic exclusion of voxels of lower signal intensity to assist in defining ROIs. We found that breast lesion ADCs obtained with the semi-automated approach are more reproducible than those obtained by a non-assisted manual ROI method, suggesting that this tool could be helpful for standardizing ADC measurements.

Prior work using water^{17, 18} or gel-based¹⁹ phantoms has demonstrated that fairly high reproducibility in measured ADC values can be achieved between MRI scanners when using consistent DWI protocols. However, there are scarce data to date showing the effect of ROI drawing technique on ADC reproducibility for in vivo breast imaging, which may be even more problematic than other organ systems due to anatomic considerations and unique lesion characteristics. In particular, the varying ratios of fibroglandular to adipose tissue in breast parenchyma may contribute to less accurate ADCs for less well-defined lesions, which could be a contributing factor to the recognized relatively poor utility of DWI to discriminate malignant from benign pathologies presenting as NME on DCE MRI⁶. The few previous studies that have examined inter-reader reproducibility of ADC measures have focused on the effect of standardizing ROI shapes (e.g. round or elliptical vs. freehand): Hakulinen et al demonstrated that circular ROIs provide higher repeatability than freehand ROIs but greater variability in ADC measurements of white matter tissue of the brain²⁰, while Tagliafico et al reported “acceptable” levels of inter-reader reproducibility of normal breast parenchyma ADCs by employing a standardized elliptical ROI of predetermined size ($\kappa = 0.81$)²¹.

However, the use of standardized ROI shapes often is not practical for performing ADC measurements in the breast. For example, approximately one third of suspicious MRI findings present as NME²², which often are not easily encompassed in their entirety by standardized ROI shapes. In our study, improvement in ADC reproducibility with the semi-automated technique was as great for NME lesions as for masses, suggesting the semi-automated technique is valuable for lesions of any morphologic type. Furthermore, because the ratio of fat to fibroglandular tissue can vary significantly among women, free-hand ROIs often are preferred to minimize partial volume effects that can artificially lower ADC measurements. Our study found that ADC measurements of lesions taken with freehand ROIs with the assistance of a dynamic semi-automated thresholding tool to exclude non-tumor voxels exhibiting lower signal intensity on DW images allowed readers to produce ADC measures with higher inter-reader reproducibility than the manual unassisted free-hand method. This implies that use of the semi-automated method increases ADC reproducibility by minimizing the impact of inadvertent and sometimes unavoidable inclusion of low signal voxels when defining free-hand ROIs for breast measurements. On the other hand, we found that the semi-automated approach did not improve reproducibility between readers for normal tissue measurements. This finding is not surprising in this study using a two-dimensional approach due to the high subjectivity for selection of a representative ROI of

normal tissue. Future development of a three-dimensional approach using a semi-automated thresholding tool may allow for greater reproducibility of normal tissue measures between readers.

Our study is limited by several factors that could be addressed through further investigation. The ROIs drawn were two-dimensional rather than volumetric, and volumetric measurements could further improve lesion and normal tissue characterization. The threshold level by which lower DWI signal intensity voxels were excluded was dynamically selected by each reader, which itself could introduce measurement variability, particularly for lesions that exhibit lower DWI signal. Additional optimization to automate selection of threshold levels in each case could help to further improve inter-reader reproducibility, particularly for normal tissue. Furthermore, long T2 decay time in some tissues can cause areas of the DW image to appear artificially high in signal (i.e. “T2 shine-through”), which could lead to inappropriate voxels being included in lesion ROIs due to high T2 signal rather than restricted diffusion. This might be particularly problematic in extent-of-disease MR examinations when evaluating cancers that exhibit high T2 signal due to post biopsy change. Further optimizations could reduce the influence of T2 shine-through by cross-referencing with ADC for voxel selection so that areas that are high in ADC are excluded. Finally, we utilized DWI data from a single institution and scanner platform; cross platform data would help to further evaluate the robustness and generalizability of our technique.

In summary, we present a straightforward semi-automated approach of ROI selection through dynamic thresholding of signal on the DW image to allow highly reproducible ADC measurements of breast lesions. Our method produced ADC measurements with improved reproducibility across readers without introducing systematic bias. Development of such a method is timely given the growing body of evidence supporting clinical use of DWI in breast imaging. Further validation of this approach is warranted to confirm its ability to provide efficient, accurate, and reproducible DWI measures across institutions and in multicenter clinical trials (currently underway as part of the American College of Radiology Imaging Network 6702 trial²³).

Acknowledgments

Funding support:

This study was supported by Grants from the following organizations: National Institutes of Health (NIH): R01CA151326 (Partridge, P.I.), UL1TR000423 (Institute of Translational Health Sciences) and 5P30CA047904-24 (UPCI biostatistics shared resource); Radiological Society of North America (RSNA): Research Scholar Grant (Rahbar, P.I.)

References

1. Partridge SC, McDonald ES. Diffusion weighted magnetic resonance imaging of the breast: protocol optimization, interpretation, and clinical applications. *Magn Reson Imaging Clin N Am*. 2013; 21:601–624. [PubMed: 23928248]
2. Le Bihan D, Breton E, Lallemand D, et al. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders. *Radiology*. 1986; 161:401–407. [PubMed: 3763909]

3. Parsian S, Rahbar H, Allison KH, et al. Nonmalignant breast lesions: ADCs of benign and high-risk subtypes assessed as false-positive at dynamic enhanced MR imaging. *Radiology*. 2012; 265:696–706. [PubMed: 23033500]
4. Partridge SC, DeMartini WB, Kurland BF, et al. Quantitative diffusion-weighted imaging as an adjunct to conventional breast MRI for improved positive predictive value. *AJR Am J Roentgenol*. 2009; 193:1716–1722. [PubMed: 19933670]
5. Partridge SC, Demartini WB, Kurland BF, et al. Differential diagnosis of mammographically and clinically occult breast lesions on diffusion-weighted MRI. *J Magn Reson Imaging*. 2010; 31:562–570. [PubMed: 20187198]
6. Partridge SC, Mullins CD, Kurland BF, et al. Apparent diffusion coefficient values for discriminating benign and malignant breast MRI lesions: effects of lesion type and size. *AJR Am J Roentgenol*. 2010; 194:1664–1673. [PubMed: 20489111]
7. Partridge SC, Rahbar H, Murthy R, et al. Improved diagnostic accuracy of breast MRI through combined apparent diffusion coefficients and dynamic contrast-enhanced kinetics. *Magn Reson Med*. 2011; 65:1759–1767. [PubMed: 21254208]
8. Rahbar H, Partridge SC, Eby PR, et al. Characterization of ductal carcinoma in situ on diffusion weighted breast MRI. *Eur Radiol*. 2011; 21:2011–2019. [PubMed: 21562806]
9. Partridge SC, Rahbar H, Chenevert T, et al. American College of Radiology Imaging Network (ACRIN) 6702: A multi-center study evaluating the utility of diffusion weighted imaging for detection and diagnosis of breast cancer. 2014 2014. Available at: <http://www.acrin.org/Portals/0/Protocols/6702/ACRIN6702Protocolwebversion30Oct2013pdf>.
10. Hylton N, Partridge SC, Rosen M, et al. American College of Radiology Imaging Network (ACRIN) 6698: Diffusion Weighted MR Imaging Biomarkers for Assessment of Breast Cancer Response to Neoadjuvant Treatment: A sub-study of the I-SPY 2 TRIAL (Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging And moLecular Analysis). 2014 2012. Available at: http://www.acrin.org/Portals/0/Protocols/6698/Protocol-ACRIN6698_v22912_active_ForOnline.pdf.
11. McDonald ES, Schopp JG, Peacock S, et al. Diffusion-weighted MRI: association between patient characteristics and apparent diffusion coefficients of normal breast fibroglandular tissue at 3 T. *AJR Am J Roentgenol*. 2014; 202:W496–W502. [PubMed: 24758685]
12. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1:307–310. [PubMed: 2868172]
13. Crawford SB, Kosinski AS, Lin HM, et al. Computer programs for the concordance correlation coefficient. *Comput Methods Programs Biomed*. 2007; 88:62–74. [PubMed: 17709153]
14. Williamson JM, Crawford SB, Lin HM. Resampling dependent concordance correlation coefficients. *J Biopharm Stat*. 2007; 17:685–696. [PubMed: 17613648]
15. Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989; 45:255–268. [PubMed: 2720055]
16. Barnhart HX, Haber M, Song J. Overall concordance correlation coefficient for evaluating agreement among multiple observers. *Biometrics*. 2002; 58:1020–1027. [PubMed: 12495158]
17. Malyarenko D, Galban CJ, Lundy FJ, et al. Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. *J Magn Reson Imaging*. 2012
18. 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]
19. Lavdas I, Behan KC, Papadaki A, et al. A phantom for diffusion-weighted MRI (DW-MRI). *J Magn Reson Imaging*. 2013
20. Hakulinen U, Brander A, Ryymin P, et al. Repeatability and variation of region-of-interest methods using quantitative diffusion tensor MR imaging of the brain. *BMC Med Imaging*. 2012; 12:30. [PubMed: 23057584]
21. Tagliafico A, Rescinito G, Monetti F, et al. Diffusion tensor magnetic resonance imaging of the normal breast: reproducibility of DTI-derived fractional anisotropy and apparent diffusion coefficient at 3.0 T. *Radiol Med*. 2012; 117:992–1003. [PubMed: 22580812]

22. Mahoney MC, Gatsonis C, Hanna L, et al. Positive predictive value of BI-RADS MR imaging. *Radiology*. 2012; 264:51–58. [PubMed: 22589320]
23. American College of Radiology Imaging Network (ACRIN) 6702: A Multi-Center Study Evaluating the Utility of Diffusion Weighted Imaging for Detection and Diagnosis of Breast Cancer. <https://www.acrin.org/TabID/879/Default.aspx>.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

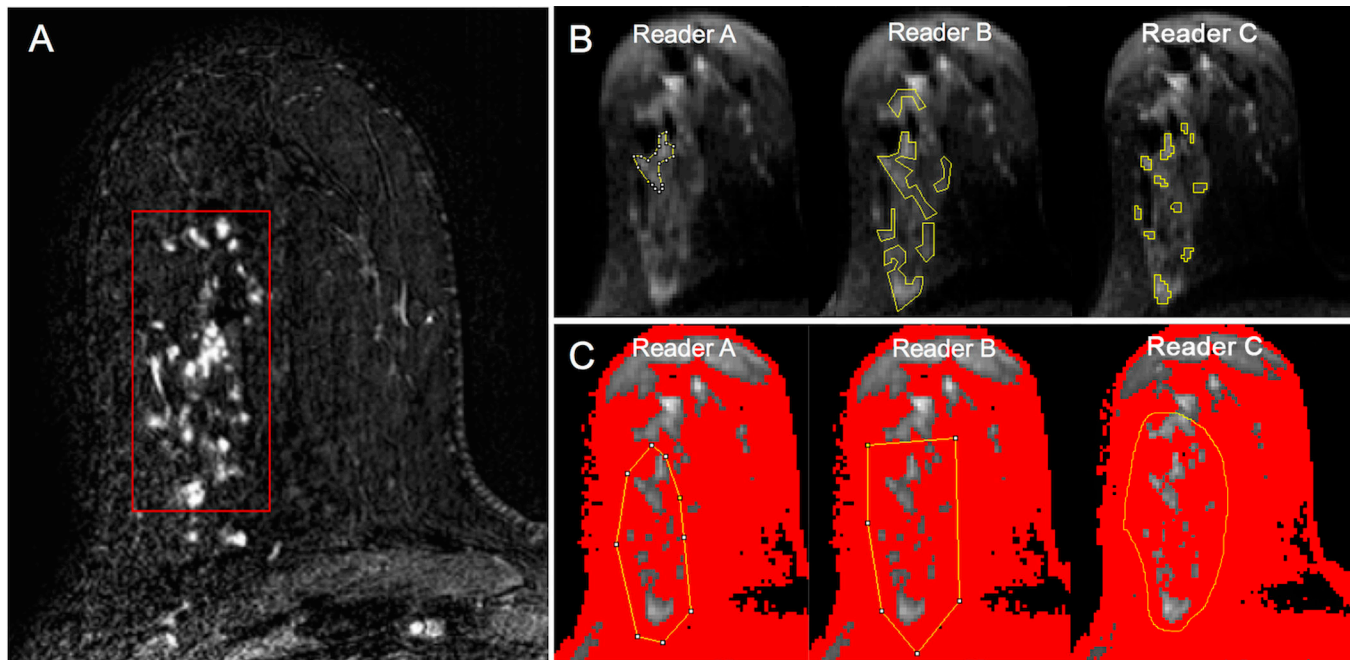


Figure 1.

An example of regions of interest (ROIs) placed on diffusion weighted images ($b = 800$ s/mm^2) by three readers for a suspicious non-mass enhancement (NME) in the right breast as depicted on the subtracted axial post-contrast image (red box, A), biopsy-proven invasive lobular carcinoma with associated lobular carcinoma in situ. All three readers placed ROIs (yellow lines) encompassing the lesion using both manual (B) and semi-automated (C) methods. Using the semi-automated thresholding tool, voxels of lower signal intensity were first masked out by each reader to be excluded from analysis (depicted in red) (C). Note the wide variability in manual ROI size and shape between readers when attempting to include the lesion without including adjacent voxels of fat. By comparison, the voxels selected for analysis using the semi-automated method agreed well across readers.

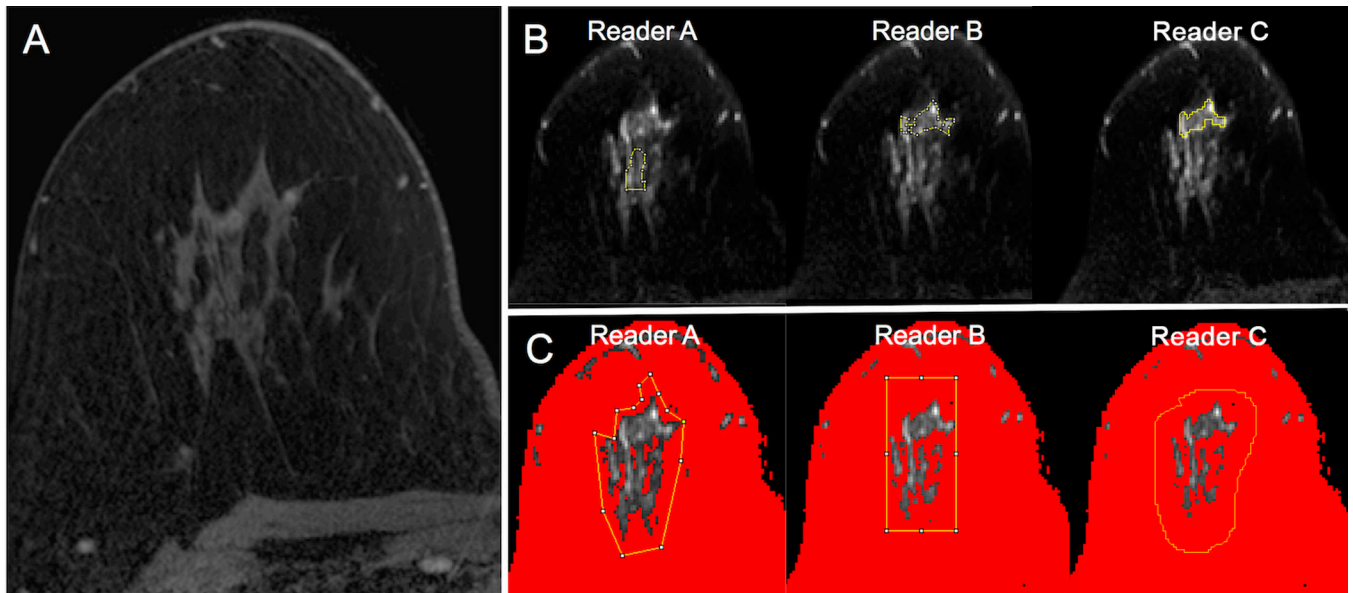


Figure 2.

An example of regions of interest (ROIs) placed on diffusion weighted images ($b = 800 \text{ s/mm}^2$) by three readers for a suspicious mass in the right breast as depicted on the subtracted axial post-contrast image (red box, A), biopsy-proven benign apocrine cystic metaplasia. All three readers placed ROIs (yellow lines) encompassing the lesion using both manual (B) and semi-automated (C) methods. Using the semi-automated thresholding tool, voxels of lower signal intensity were first masked out by each reader to be excluded from analysis (depicted in red) (C). Note the moderate variability in manual ROI size and shape between readers when attempting to include the lesion without including adjacent voxels of fat. By comparison, the voxels selected for analysis using the semi-automated method agreed well across readers.

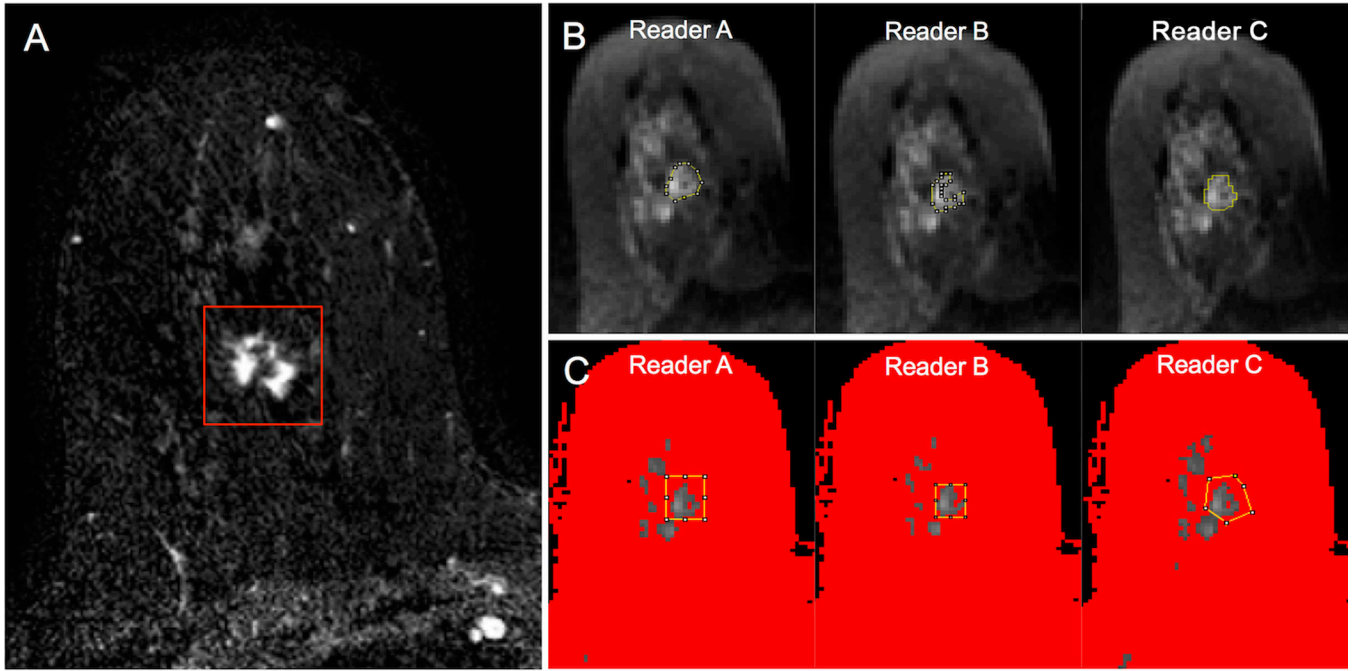


Figure 3.

An example of regions of interest (ROIs) placed on diffusion weighted images ($b = 0 \text{ s/mm}^2$) by three readers on a region of normal tissue in the right breast, as depicted on T1-weighted post-contrast image (A). All three readers placed ROIs (yellow lines) encompassing the normal tissue using both the manual (B) and semi-automated (C) ROI selections methods. Using the semi-automated thresholding tool, voxels of lower signal intensity were first masked out by each reader to be excluded from analysis (depicted in red) (C). Note the wide variability in ROI sizes between readers attempting to include an area of normal tissue without including adjacent voxels of fat, while the voxels selected for analysis using the semi-automated method were relatively comparable.

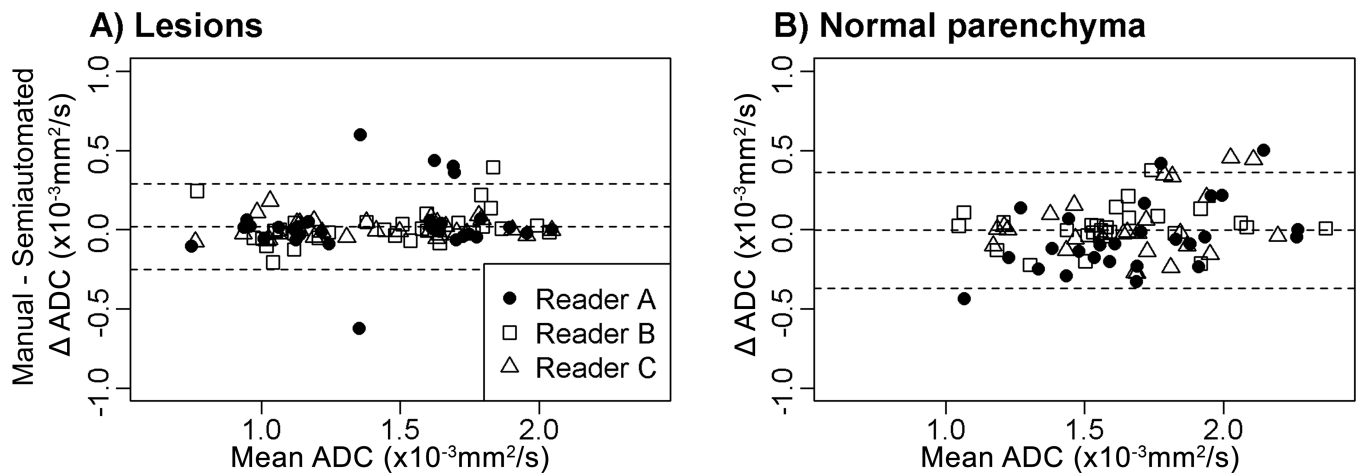


Figure 4. Comparisons of ADC measures ($\times 10^{-3} \text{mm}^2/\text{s}$) by manual and semi-automated methods for lesions (A) and normal tissue (B). For both lesions and normal tissue, the middle horizontal line (the average difference of readers' ADC measures with manual and semi-automated methods) is close to zero, indicating that the use of the semi-automated method did not result in systematic bias.

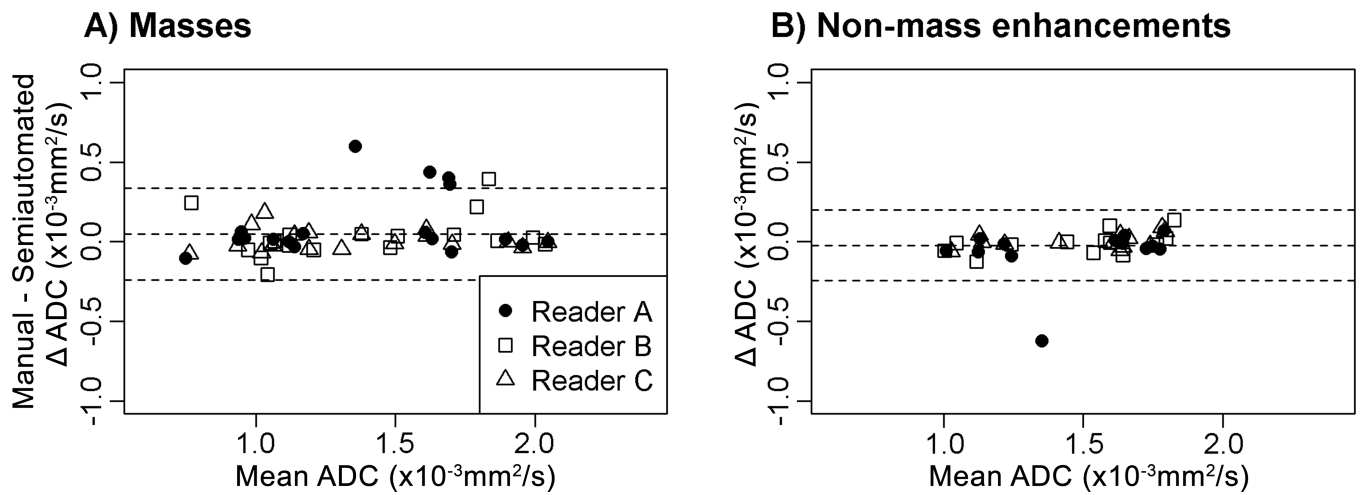


Figure 5. Comparisons of ADC measures ($\times 10^{-3} \text{mm}^2/\text{s}$) by manual and semi-automated methods for masses (A) and non-mass enhancement (B) lesions. For both lesion types, the middle horizontal line (the average difference of readers' ADC measures with manual and semi-automated methods) is close to zero, indicating that the use of the semi-automated method did not result in systematic bias.

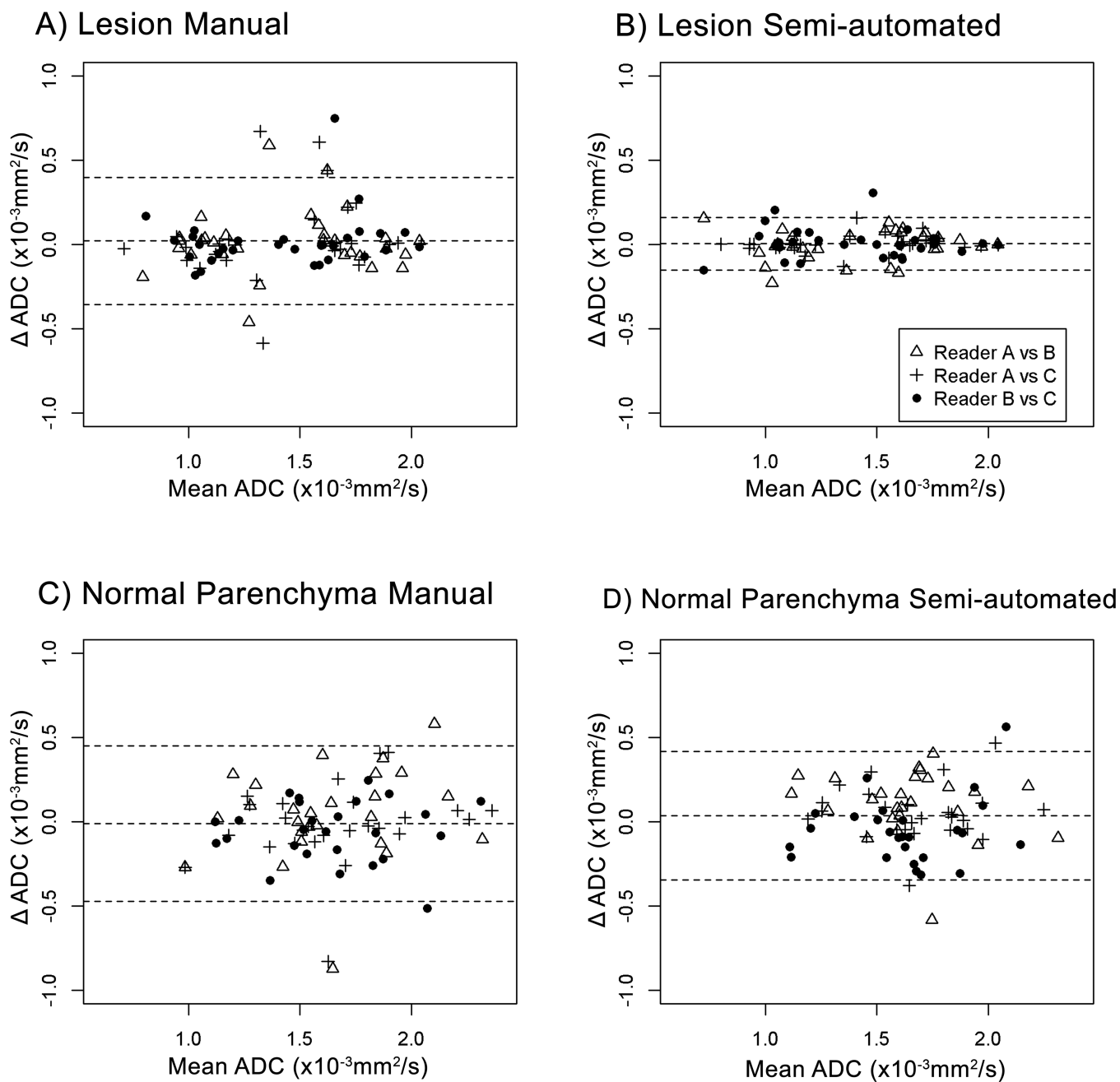


Figure 6.

Bland-Altman plots showing comparisons of ADC measures ($\times 10^{-3} \text{ mm}^2/\text{s}$) among 3 readers for lesions using the manual (A) and semi-automated (B) techniques as well as for normal tissue using the manual (C) and semi-automated (D) techniques. In all panels the middle horizontal line (the average difference of readers' paired ADC measures) is close to zero. The limits of agreement for lesions were narrower for the semi-automated method (B) when compared to the manual method (A), indicating the semi-automated method had greater inter-reader reproducibility.

Table 1

Lesion characteristics.

	Benign lesions (n=15) N (%)	Malignant lesions (n=16) N (%)
MRI Lesion Size (median, range)	9 (5 – 67) mm	12.5 (5 – 105) mm
MRI Morphology		
Mass	8 (53%)	10 (63%)
Non-mass enhancement	7 (47%)	6 (37%)
Histology		
Benign	15 (100%)	-
Invasive ductal carcinoma	-	7 (44%)
Invasive lobular carcinoma	-	4 (25%)
Ductal carcinoma in situ	-	5 (31%)

Comparison of manual and semi-automated ADC measures ($\times 10^{-3} \text{ mm}^2/\text{s}$) performed by each reader.

Table 2

Type	Reader*	Method	N	Median	Min	Max	Difference Median (p-value)**
Lesion							
	A	Manual	31	1.64	0.70	2.05	0.009 (0.62)
		Semi-automated	31	1.49	0.80	2.04	
	B	Manual	31	1.46	0.89	2.03	-0.006 (0.94)
		Semi-automated	31	1.49	0.65	2.04	
	C	Manual	31	1.41	0.72	2.04	-0.005 (0.79)
		Semi-automated	31	1.42	0.80	2.05	
Normal Tissue							
	A	Manual	26	1.57	0.85	2.39	-0.089 (0.14)
		Semi-automated	26	1.67	1.20	2.28	
	B	Manual	26	1.57	1.06	2.37	0.016 (0.28)
		Semi-automated	26	1.56	1.01	2.36	
	C	Manual	26	1.65	1.12	2.33	-0.020 (0.81)
		Semi-automated	26	1.67	1.18	2.21	

* Readers A and B were radiologists specializing in breast imaging. Reader C was a research associate with 2 years of experience with breast MRI.

** P-values calculated by Wilcoxon signed-rank test for paired observations (manual and semi-automated measures).

Table 3

Inter-reader reproducibility: Concordance correlation coefficients (CCC) for ADC measurements.

Reader pair	Manual CCC (95% bootstrap confidence interval)	Semi-automated CCC (95% bootstrap confidence interval)	CCC Difference (Confidence Interval)
Lesion			
A-B	0.88 (0.72 – 0.95)	0.97 (0.94 – 0.98)	0.09 (0.01 – 0.28)
A-C	0.80 (0.54 – 0.93)	0.99 (0.97 – 1.00)	0.19 (0.06 – 0.43)
B-C	0.90 (0.63 – 0.97)	0.96 (0.91 – 0.98)	0.07 (0.01 – 0.29)
Overall	0.86 (0.71 – 0.95)	0.97 (0.95 – 0.99)	0.11 (0.04 – 0.26)
Mass	0.86 (0.64 – 0.96)	0.97 (0.93 – 0.99)	0.11 (0.02 – 0.29)
NME	0.86 (0.45 – 0.96)	0.98 (0.94 – 0.99)	0.12 (0.02 – 0.42)
Normal Tissue			
A-B	0.68 (0.28 – 0.85)	0.72 (0.44 – 0.87)	0.04 (–0.09 – 0.20)
A-C	0.78 (0.35 – 0.92)	0.81 (0.61 – 0.93)	0.04 (–0.18 – 0.28)
B-C	0.83 (0.67 – 0.93)	0.76 (0.56 – 0.88)	–0.07 (–0.29 – 0.09)
Overall	0.76 (0.49 – 0.88)	0.76 (0.61 – 0.87)	0.00 (–0.14 – 0.13)

CCC = concordance correlation coefficient.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript