UCLA
UCLA Electronic Theses and Dissertations

Title

Towards Early Treatment Response Prediction using Longitudinal Distortion-free Diffusion-weighted Magnetic Resonance Imaging

Permalink

https://escholarship.org/uc/item/05w7s7ns

Author

Gao, Yu

Publication Date

2019

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Towards Early Treatment Response Prediction using Longitudinal Distortion-free Diffusion-weighted Magnetic Resonance Imaging

A dissertation submitted in partial satisfaction of the

requirements for degree Doctor of Philosophy

in Biomedical Physics

by

Yu Gao

2019

© Copyright by

Yu Gao

2019

ABSTRACT OF THE DISSERTATION

Towards Early Treatment Response Prediction using Longitudinal Distortion-free Diffusion-weighted Magnetic Resonance Imaging

by

Yu Gao

Doctor of Philosophy in Biomedical Physics

University of California, Los Angeles, 2019

Professor Peng Hu, Co-Chair

Professor Daniel Abraham Low, Co-Chair

Radiotherapy is an effective tool to treat tumors by delivering high-energy photons or charged particles to destroy malignant cancer cells. In the current fractionated radiotherapy treatment workflow, the same radiation plan is delivered to the patient in every fraction during the entire course of treatment, whereas tumor changes and patient response are not taken into consideration. Early patient response prediction is appealing as it offers a window of personalized treatment adaptation for potentially improved treatment outcome.

Diffusion-weighted imaging (DWI) has been shown to be a promising non-invasive biomarker for treatment response assessment. However, the conventional diffusion-weighted single-shot echo-planar-imaging (DW-ssEPI) has strong spatial distortion, which is unacceptable for radiotherapy applications as high geometric accuracy is required for accurate tumor

delineation. To achieve the ultimate goal of response-based adaptive radiotherapy, this dissertation sought to develop distortion-free diffusion sequences, and evaluate the possibility of early treatment response assessment using longitudinal DWI on sarcoma patient.

In Chapter 3, a diffusion-prepared turbo spin-echo (DP-TSE) sequence was programmed and compared with the DW-ssEPI sequence on a 0.35T MRI-guided radiotherapy system. The DW-ssEPI failed the spatial integrity test due to severe distortion and low signal intensity, whereas the DP-TSE passed the test successfully. The diffusion phantom study showed that noise correction must be performed for the DW-ssEPI sequence to avoid apparent diffusion coefficient (ADC) quantification errors, whereas DP-TSE had desirable ADC accuracy and ADC reproducibility. Good geometric fidelity and ADC quantifications were obtained in the patient study involving two glioblastoma (GBM) patients and six sarcoma patients.

Shot-to-shot k-space magnitude inconsistency is a common problem in multi-shot diffusion-prepared imaging. In Chapter 4, a magnitude stabilizer strategy was proposed to convert the malignant magnitude inconsistency to phase inconsistency, which is easier to resolve. We demonstrated that the proposed diffusion-prepared magnitude-stabilized balanced steady-state free precession sequence (DP-MS-bSSFP, abbreviated as DP-MS) had satisfactory ADC accuracy, and significantly improved geometric accuracy on both phantom and volunteers compared to the conventional DW-ssEPI approach.

To meet the requirement of high spatial integrity and high spatial resolution for treatment planning and adaptation, the 2D DP-MS sequence was extended to 3D in Chapter 5. A locally low-rank constrained reconstruction was applied to correct the k-space inconsistency. Similar as Chapter 4, the 3D DP-MS sequence was verified on the diffusion phantom and five healthy

volunteers for geometric fidelity and ADC accuracy. Overall, the 3D DP-MS sequence had submillimeter geometric accuracy and satisfactory ADC accuracy on both phantom and volunteers.

In the second half of this dissertation, we focused on early treatment response prediction using longitudinal DWI on sarcoma patients treated with hypofractionated radiation therapy. Diffusion images were acquired using DW-ssEPI three times through the treatment. In Chapter 6, a radiomics-based approach was used to prediction treatment effect score, which is obtained from the post-surgery pathology. Logistic regression and support vector machine (SVM) models were constructed to predict the treatment effect using radiomics features selected by univariate analysis and sequential forward selection. The SVM model outperformed logistic regression and had an area under the receiver operating characteristic curve (AUC) of 0.91±0.05.

To overcome the small sample size problem in Chapter 6 and further improve the prediction, deep learning-based data augmentation and prediction were implemented in Chapter 7. An ACGAN network was trained to augment the data based on training patients, and a prediction model based on the VGG-19 was trained using the synthesized data and validated on the training patient dataset. This trained model was then tested on the hold-out test patients. Overall, the training, validation, and test accuracies was 94.3%, 90.1%, and 87.7%, indicating good performance of the data augmentation and response prediction models.

The dissertation of Yu Gao is approved.

Yingli Yang

Dan Ruan

Ke Sheng

William Hsu

Danial Abraham Low, Committee Co-Chair

Peng Hu, Committee Co-Chair

University of California, Los Angeles

2019

TABLE OF CONTENT

TABLE OF CONTENT ..... **vi**

LIST OF ACRONYMS ..... **viii**

LIST OF FIGURES ..... **x**

LIST OF TABLES ..... **xvi**

ACKNOWLEDGEMENTS ..... **xvii**

Vita ..... **xx**

Chapter 1 Introduction **1**

1.1 Thesis Motivation 1

1.2 Organization of the thesis 4

Chapter 2 Background **8**

2.1 Diffusion-weighted imaging 8

2.2 Echo planar imaging 10

2.3 DW and DP 12

2.4 Radiomics 14

2.5 Deep Learning 15

Chapter 3 Diffusion-prepared TSE sequence **18**

3.1 Introduction 18

3.2 Methods 20

3.3 Result 26

3.4 Discussion 33

3.5 Conclusion 36

Chapter 4 2D Diffusion-prepared Magnitude-stabilized bSSFP imaging **37**

4.1 Introduction 37

4.2 Methods 40

4.3 Result 47

4.4 Discussion 54
4.5 Conclusion 57

Chapter 5 3D Diffusion-prepared Magnitude-stabilized bSSFP imaging 59
5.1 Introduction 59
5.2 Methods 61
5.3 Result 66
5.4 Discussion 71
5.5 Conclusion 75

Chapter 6 Treatment response prediction: a radiomics approach 76
6.1 Introduction 76
6.2 Methods 78
6.3 Result 84
6.4 Discussion 87
6.5 Conclusion 89

Chapter 7 Treatment response prediction: a deep learning approach 90
7.1 Introduction 90
7.2 Methods 92
7.3 Result 99
7.4 Discussion 102
7.5 Conclusion 105

Chapter 8 Conclusion 106
8.1 Summary of contribution 106
8.2 Future outlook 109

REFERENCE 112

LIST OF ACRONYMS

General

T	-	Tesla
mm	-	millimeter
s	-	second
ms	-	millisecond
ECG	-	Electrocardiogram
RT	-	radiation therapy
MRgRT	-	MRI-guided radiotherapy
STS	-	soft tissue sarcoma
TES	-	treatment effect score

General Magnetic Resonance Imaging

MRI	-	magnetic resonance imaging
2D/3D	-	two/three dimensional
FOV	-	field of view
TE	-	echo time
TR	-	repetition time
SNR	-	signal to noise ratio
RF	-	radiofrequency
ETL	-	echo train length
DWI	-	diffusion-weighted magnetic resonance imaging
CPMG	-	Carr-Purcell-Meiboom-Gill
ADC	-	apparent diffusion coefficient
FA	-	fractional anisotropy
cFA	-	colored fractional anisotropy
bSSFP	-	balanced steady-state free precession
TSE	-	turbo spin echo imaging

FLASH	-	fast low-angle shot imaging
EPI	-	echo-planar imaging
DW-ssEPI	-	diffusion-weighted single-shot echo-planar imaging
DP	-	diffusion-prepared
DW	-	diffusion-weighted
DP-TSE	-	diffusion-prepared turbo spin echo imaging
DP-MS-bSSFP	-	diffusion-prepared magnitude-stabilized bSSFP sequence
TRSE	-	twice-refocused spin-echo

Special Technical Terms

MUSE	-	multiplexed sensitivity-encoding
MS	-	magnitude stabilizer
TRE	-	target registration error
CoV	-	coefficient of variation
ROI	-	regions of interest
PVP	-	polymer polyvinylpyrrolidone
SVM	-	support vector machine
AUC	-	area under the receiver operating characteristic curve
GLCM	-	gray level co-occurrence matrix
GLRLM	-	gray level run length matrix
GLSZM	-	gray level size zone matrix
NGTDM	-	neighboring gray tone difference matrix
GLDM	-	gray level dependence matrix
CNN	-	convolutional neural network
GAN	-	generative adversarial networks
ACGAN	-	auxiliary classifier generative adversarial network

LIST OF FIGURES

Figure 1-1 Outline of technical developments in this work. 5

Figure 2-1 Gradient waveforms for the (a) spin-echo pulse sequence, and (b) gradient-echo pulse sequence. DG represents the diffusion gradient. 10

Figure 2-2 (a) gradient waveform of diffusion-weighted EPI sequence. (b) corresponding k-space trajectory. Gss, Gro, Gpe represent the gradient direction of slice selective, readout and phase encoding respectively..... 11

Figure 2-3 Interleaved EPI (a) and readout-segmented EPI (b). Each color represents one echo train. In both examples, the k-space was filled in with five shots. 12

Figure 2-4 Difference between diffusion-weighted approach and diffusion-prepared approach. (a) sequence diagram of the DW approach. (b) example of phase inconsistency between two different shots in the DW approach. (c) sequence diagram of DP approach. (d) (e) examples of magnitude inconsistency in the DP approach before and after the -90˚ flip back RF pulse. The red and blue arrows represent the signal vectors after the diffusion encoding for two different shots. 13

Figure 2-5 Architectures of (a) GAN, (b)cGAN, and (c) ACGAN. 16

Figure 2-6 Architecture of the VGG-19 network 17

Figure 3-1 Sequence diagram (a) and an example of a 4-shot k-space view ordering (b). The twice-refocused spin-echo scheme was used to generate diffusion weighting while minimizing eddy currents. K-space view ordering was modified such that the central k-space is acquired by the same imaging shot..... 21

Figure 3-2 Spatial Integrity phantom (a) and diffusion phantom (b). The spatial integrity phantom is of dimension 33.02×33.02×10.16cm^3 and consists of 197 cylindrical grids (20×20 with 3 missing for positioning and detecting). The diffusion phantom is 194mm in diameter and has 13 vials with six different diffusivities. Vials with diffusivity less than 0.35×10^{-3} mm^2/s were not included in this study. 22

Figure 3-3 Phantom images (transverse view) acquired at iso-center using the proposed DP-TSE (a) and the DW-ssEPI (b) techniques with b=0. (c) and (d) are corresponding zoomed-in images. The green crosses represent the detected marker locations based on the image, and the blue circles are the ground truth marker locations. The DW-ssEPI image has substantial distortions. The detected location was compared with the ground truth for accuracy. A "NaN" would return if the software failed to detect the three missing markers for localization..... 27

Figure 3-4 Raw images of DW-ssEPI (MeR) and DP-TSE (HiR) at b=0 and b=800. 28

Figure 3-5 Effect of noise subtraction on ADC calculation based on DW-ssEPI (a, c) and DP-TSE (b, d) techniques at 0°C (a, b) and room temperature (c, d). 29

Figure 3-6 ADC Accuracy after noise subtraction for DW-ssEPI (MeR) and DP-TSE (HiR) at 0°C (a) and room temperature (b). 30

Figure 3-7 A comparison from two representative patients. (a)-(h) are the CT, clinical MR image from the standard ViewRay bSSFP sequence, DW-ssEPI (b=0), DP-TSE(b=0), DW-ssEPI (b=500), DP-TSE(b=500), and ADC maps calculated from DW-ssEPI and DP-TSE on one GBM patient, respectively. (i)-(p) are corresponding images for a sarcoma patient. The blue and green arrows point to signal pile-up and distortion artifacts in the DW-ssEPI image. The red arrows and red dotted lines indicate chemical shift artifacts of the bone marrow and subcutaneous fat. For each patient, the four diffusion-weighted raw images were displayed at the same window level..... 31

Figure 3-8 TRE measurements based on DW-ssEPI and DP-TSE..... 32

Figure 4-1 (a) Sequence diagram of one shot. Blue rectangles represent the magnitude stabilizers. (b) Illustration of magnetizations after the diffusion gradients for two different shots (blue and red arrows). (c-f) Magnetizations for the two shots at time II-V. (c) The magnetization is evenly dispersed in the transversal plane after the MS gradient. (d) Half of the magnetization is flipped to the z-direction regardless of the signal phase angles, while the remaining transversal magnetization is spoiled. (e) The magnetization after the RF pulse. For simplicity, a 90° readout RF pulse is used in the illustration. (f) After the rewinding MS, the magnetization is restored. DG, diffusion gradient; SP, spoiler; CP, catalyzation pulse..... 42

Figure 4-2 Examples of landmark locations used in the geometric accuracy quantification for slice containing the lateral ventricle and slice at the base of the brain. Ten landmarks were selected for each volunteer. 46

Figure 4-3 Simulation of signal magnitude with respect to phase error for the 1st, 20th and 40th readout echo without and with MS. The phase error does not affect the signal magnitude of the DP-MS-bSSFP approach; whereas it causes signal oscillation in the DP-bSSFP approach. 47

Figure 4-4 Phantom images. TSE was acquired as the geometric reference. Images at b-value = 0 and 800 s/mm^2 from DW-ssEPI, DP-bSSFP, and DP-MS-bSSFP were shown. The red dots on the TSE image indicate the four landmarks (12 o'clock, 3 o'clock, 6 o'clock and 9 o'clock) of the phantom and the four landmarks of vial 1 which were part of the 24 landmarks used for geometric accuracy quantification. Red arrows point to distortion artifact in DW-ssEPI, and blue arrows point to signal void artifact in DP-bSSFP. 48

Figure 4-5 Magnitude and phase images of five repetitions from DW-ssEPI (abbreviated as EPI), DP-bSSFP (abbreviated as DP) without ECG triggering, DP with ECG trigger delay of 0ms,

and 200ms, DP-MS-bSSFP (abbreviated as DP-MS) without ECG triggering, and DP-MS with ECG trigger delay of 0ms. 50

Figure 4-6 CoV maps calculated from the ten repetitions of DW-ssEPI, DP without ECG triggering, DP with ECG trigger delay of 0ms, and 200ms, DP-MS without ECG triggering, and DP-MS with ECG trigger delay of 0ms. 51

Figure 4-7 Multi-shot in vivo results from two volunteers. Images at b-value of 0 s/mm^2 and 800 s/mm^2, and cFA maps from DW-ssEPI and DP-MS-bSSFP were shown. Red arrows point to distortion and signal pile-up artifacts in DW-ssEPI. 52

Figure 4-8 In vivo TRE measurement for the six volunteers. The error bar indicates the standard deviation. For volunteer 1,2,4,6, images were acquired at slice containing the lateral ventricle; for volunteers 3 and 5, images were acquired at the base of the brain. 53

Figure 4-9 Bland-Altman plot for the white matter ADC (a), CSF ADC (b), and white matter FA (c). The p-value is from the paired t-test. 53

Figure 4-10 Images from b=0 and b=800 s/mm^2, and cFA maps for DP-MS-bSSFP and DP-MS-FLASH. Same window and level were used for each comparison. DP-MS-bSSFP provided higher SNR and improved contrast than DP-MS-FLASH. The quality of cFA map from DP-MS-bSSFP was better 55

Figure 4-11 SNR comparison of DP-bSSFP, DP-MS-bSSFP, and DP-MS-FLASH on a water ball phantom. 15 repetitive measurements were performed to calculate the SNR map. SNR of DP-bSSFP was approximately 2.5 times higher than DP-MS-bSSFP, and SNR of DP-MS-bSSFP was about 1.9 times higher than DP-MS-FLASH..... 57

Figure 5-1 Sequence diagram of one example shot. Blue rectangles represent the magnitude stabilizers. In this study, the navigator was acquired in each shot during the 16 linear catalyzation pulses..... 62

Figure 5-2 Phantom image comparison. Images from left to right are: TSE, DW-ssEPI (b=800 s/mm^2), and DP-MS (b=800 s/mm^2). The blue dotted lines and arrow point to distortion and artifact in the DW-ssEPI image..... 67

Figure 5-3 ADC accuracy at ice temperature and room temperature. The x-axis shows the labels of the thirteen vials. The number represents the percent concentration of PVP solutions (higher PVP, lower diffusivity), and c (center), i (inner), o (outer) represent the relative location of the vial. The y-axis is the reference and measured ADC values in 10^{-3}mm^2/s. The error bar represents the standard deviation. 68

Figure 5-4 Comparison of b=800 s/mm^2 images acquired using one shot per kz, two shots per kz, and 4 shots per kz. The blue triangles points to the blurring of the grey matter structure..... 69

Figure 5-5 Representative transversal plane images from two volunteers. The anatomical reference image, images at b-value of 0 s/mm^2 and 800 s/mm^2 from DW-ssEPI and DP-MS were shown. The blue triangles point to distortion and signal pile-up artifacts in DW-ssEPI..... 70

Figure 5-6 Three-plane view of the ADC map from one volunteer. 71

Figure 5-7 Bland-Altman plot of the white matter, cerebellum and CSF ADC between DW-ssEPI and DP-MS. Good agreements were observed for the ADC values from the two sequences. 71

Figure 5-8 Comparison of b=800s/mm^2 images acquired at a 3T scanner using the 1 shot, 2 shots, and 4 shots per kz plane protocols. 74

Figure 6-1 Example of DWI image and the ADC map. Tumor is contoured in red. 81

Figure 6-2. Prediction performance of SVM using only mean ADC or delta ADC 84

Figure 6-3 Model AUC with respect to different feature numbers from 1 to 10 for logistic regression and SVM. The error bar indicates the standard deviation. The best feature number k was determined by inspecting the AUC curve. In this case, three features were identified for LR and four features were selected for SVM 85

Figure 6-4 Prediction performance using different time points. The error bar indicates the standard deviation. The star (*) implies a significant difference with the performance of F_{all}. ... 86

Figure 7-1 (a) Architecture of the ACGAN network. (b) Details of the discriminator. 96

Figure 7-2 Architecture of the VGG-19 based prediction network structure 96

Figure 7-3 Overall workflow of ACGAN-VGG network. 98

Figure 7-4 Examples of longitudinal patient data and ACGAN synthesized data for each class. Two generated examples were shown. For each case, ADC maps from the three time points were displayed. 99

Figure 7-5 Training, validation, and test accuracy over the five rounds 100

LIST OF TABLES

Table 3-1 ADC accuracy results at 0°C and room temperature 30

Table 3-2 In vivo ADC measurement results 33

Table 4-1 Sequence parameter in the in vivo study. 45

Table 4-2 Phantom ADC for the 13 vials 49

Table 5-1 TRE measurement of DW-ssEPI and DP-MS on four volunteers. 70

Table 6-1 Patient characteristics. (N=30) 80

Table 6-2 Features used in this study. 82

Table 6-3 Prediction performance using features from all three time points and corresponding delta radiomics. 85

Table 7-1 Patient characteristics 94

Table 7-2 Confusion matrix of the slice-based prediction results on the test sets. 101

Table 7-3 Confusion matrix of the patient-based prediction results on the test sets. 101

Table 7-4 Prediction results on the test sets for each patient. Slice column: correctly predicted slice number over the total number of slices. Patient column: F - incorrect classification; T - accurate classification 102

ACKNOWLEDGEMENTS

Chapter 3 is a version of: Gao Y, Han F, Zhou Z, Cao M, Kaprealian T, et al. Distortion-free diffusion MRI using an MRI-guided Tri-Cobalt 60 radiotherapy system: Sequence verification and preliminary clinical experience. Med Phys. 2017 Oct 1;44(10): 5357–66.

Chapter 4 is a version of: Gao Y, Fei H, Zhou Z, Zhong X, Bi X, et al. Multishot diffusion-prepared bSSFP imaging with magnitude stabilizer. Magn Reson Med. 2019; 81(4): 2374-2384.

Chapter 5 is a version of: Gao Y, Zhou Z, Fei H, Zhong X, Yang Y, Hu P. Three-dimensional multi-slab distortion-free diffusion-prepared magnitude-stabilized bSSFP imaging at 1.5 Tesla. (ready to submit)

Chapter 6 is a version of: Gao Y, Kalbasi A, Shao J, Fu J, Ruan D et al. Treatment effect prediction for sarcoma patients treated with preoperative radiotherapy using radiomics features from longitudinal diffusion MRI: a preliminary study (ready to submit)

Chapter 7 is a version of: Gao Y, Ghodrati V, Kalbasi A, Fu J, Ruan D, et al. Preliminary study of deep learning-based treatment response prediction using longitudinal diffusion MRI on sarcoma patients (ready to submit)

I would like to sincerely express my appreciation for my advisor, Dr. Peng Hu for his guidance and support throughout my Ph.D. journey. I feel fortunate to have him as my advisor. He gave me the liberty to work on the projects that I have truly enjoyed and has been very supportive when I encountered difficulties in research. His door is always open to answer every question I might have. I believe the philosophy and methodology he taught me about how to tackle challenges would benefit my whole life.

I would like to thank Dr. Yingli Yang, who is like my second advisor. She provided tremendous help in clinical data acquisition in the Radiation Oncology department. She not only provided many insightful advices on my research, but also shared her wisdom on work, life, and work-life balance. It has been a joy working with her.

I would also like to thank my thesis committee, Dr. Danial Low, Dr. Dan Ruan, Dr. Ke Sheng and Dr. William Hsu, as well as other faculties in UCLA: Dr. Minsong Cao, Dr. Anusha Kalbasi, Dr. Daniel Ennis, Dr. Kyung Sung and Dr. Holden Wu for their inspiring discussions and constructive suggestions. Those feedbacks I got have helped strengthen my work, and the knowledge I learned from them have expanded my scope.

It is my pleasure to work with all the current and past members and staff of the Magnetic Resonance Research Lab (MRRL): Felix, Ziwu, Vahid, Jiaxin, Da, Shams, Fadil, Chang, Caroline, Mark, Hengjie, Zhengyang, Take, Novena, Eric, Kevin, Jessica, Mike, Patrick, Nyasha, Xinran, Thomas, Jiahao, Andres, Tess, Le, Xinzhou, Zhaohuan, and more. I would like to thank them for being so kind, and always willing to be my volunteers for the research MRI scans. I would like to especially thank Felix and Ziwu for their generous help in my research. They are my mentors to the MRI field, and I have learned many essential skills such as sequence programming and image reconstruction from them. Without their kind help, I could not finish many projects so smoothly. I also want to extend my thanks to Heather for her efforts in making the MRRL lab clean and nice to work in.

I am lucky to be a member of the Physics and Biology in Medicine program family. It is such a warm family, and I would always remember the happy moments we had. In particular, I would

like to thank Dr. Michael McNitt-Gray, Reth Im, and Alondra Correa Bautista for always being so helpful and supportive.

Last but not least, I would like to express my gratitude to my family. My Ph.D. study would not have been possible without their continuous support. I would like to thank my husband Hanning for the love and joy he gave me during my Ph.D. journey. I would like to thank my little angle Aiden. I am fully recharged whenever I see his lovely smile. Lastly, I would like to thank my parents for always being so supportive and give me the freedom to pursue my goal.

Vita

Education:

- Ph.D. student, Biomedical Physics, University of California, Los Angeles, CA, 2014–2019
- B.S., Applied Mathematics, Xi'an Jiaotong University, Xi'an, China, 2009-2013

Awards:

- Dissertation Year Fellowship, UCLA, 2018-2019
- Young Investigator Award Finalist, AAPM Annual Meeting, 2018
- TEAM Best/AAPM Travel Grant, AAPM Annual Meeting, 2018
- Magna cum laude, ISMRM Annual Meeting, 2017, 2018
- Young Investigator Award Finalist, MR in RT Symposium, 2017
- Best Poster Award, PBM Research Colloquium, UCLA, 2016
- Graduate Dean's Scholar Award, UCLA, 2014

First Author Peer-Reviewed Publications:

1. **Gao Y**, Kalbasi A, Shao J, Fu J, Ruan D et al. Treatment effect prediction for sarcoma patients treated with preoperative radiotherapy using radiomics features from longitudinal diffusion MRI: a preliminary study (ready to submit)
2. **Gao Y**, Ghodrati V, Kalbasi A, Fu J, Ruan D, et al. Preliminary study of deep learning-based treatment response prediction using longitudinal diffusion MRI on sarcoma patients (ready to submit)
3. **Gao Y**, Zhou Z, Fei H, Zhong X, Yang Y, Hu P. Three-dimensional multi-slab distortion-free diffusion-prepared magnitude-stabilized bSSFP imaging at 1.5 Tesla. (ready to submit)
4. **Gao Y**, Fei H, Zhou Z, Zhong X, Bi X, et al. Multishot diffusion-prepared bSSFP imaging with magnitude stabilizer. Magn Reson Med. 2019; 81(4): 2374-2384.
5. **Gao Y**, Zhou Z, Fei H, Cao M, Shaverdian N, et al. Accelerated 3D bSSFP Imaging for Treatment Planning on an MRI-Guided Radiotherapy System. Med Phys. 2018 Jun;45(6):2595–602.

6. **Gao Y**, Han F, Zhou Z, Cao M, Kaprealian T, et al. Distortion-free diffusion MRI using an MRI-guided Tri-Cobalt 60 radiotherapy system: Sequence verification and preliminary clinical experience. Med Phys. 2017 Oct 1;44(10): 5357–66.

Co-Author Peer-Reviewed Publications

7. Zhou Z, Han F, Ghodrati V, **Gao Y**, Yin W, et al. Parallel imaging and convolutional neural network combined fast MR image reconstruction: Applications in low-latency accelerated real-time imaging. Med Phys. 2019 May 28;
8. Rashid S, Han F, **Gao Y**, Sung K, Cao M, et al. Cardiac balanced steady-state free precession MRI at 0.35 T: a comparison study with 1.5 T. Quant Imaging Med Surg. 2018 Aug;8(7):627–36.
9. Han F, Zhou Z, Du D, **Gao Y**, Rashid S, et al. "Respiratory motion-resolved, self-gated 4D-MRI using rotating Cartesian k-space (ROCK): Initial clinical experience on an MRI-guided radiotherapy system", Radiother Oncol J Eur Soc Ther Radiol Oncol. 2018;127(3):467–73.
10. Ginn JS, Agazaryan N, Cao M, Baharom U, Low DA, Yang Y, **Gao Y**, et al. Characterization of spatial distortion in a 0.35 T MRI-guided radiotherapy system. Phys Med Biol. 2017 Jun 7;62(11):4525–40.
11. Han F, Zhou Z, Han E, **Gao Y**, Nguyen K-L, et al. Self-gated 4D multiphase, steady-state imaging with contrast enhancement (MUSIC) using rotating cartesian K-space (ROCK): Validation in children with congenital heart disease. Magn Reson Med. 2017 Aug;78(2):472–83.
12. Yang Y, Cao M, Sheng K, **Gao Y**, Chen A, et al. Longitudinal diffusion MRI for treatment response assessment: Preliminary experience using an MRI-guided tri-cobalt 60 radiotherapy system. Med Phys. 2016 Mar 1;43(3):1369–73.
13. Yang Y, Cao M, Kaprealian T, Sheng K, **Gao Y**, et al. Accuracy of UTE-MRI-based patient setup for brain cancer radiation therapy. Med Phys. 2016 Jan 1;43(1):262–7.
14. Liu J, Zhang X, Zhang X, Zhao H, **Gao Y**, et al. 5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography. Inverse Probl. 2015;31(11):115007.
15. Gao H, Qi XS, **Gao Y**, Low DA. Megavoltage CT imaging quality improvement on TomoTherapy via tensor framelet. Med Phys. 2013 Aug;40(8):081919.

Chapter 1 Introduction

1.1 Thesis Motivation

Cancer, after heart disease, is the second leading cause of death globally. It is estimated that over 1.7 million new cases will be diagnosed in the United States in 2019, and over 0.6 million Americans will die from cancer in 2019[1]. Radiation therapy (RT) along with surgery and chemotherapy, is one of the common treatments for cancer patients. By delivering high-energy photon or charged particle to cancerous tumor, it destroys the DNA of malignant cancer cells, and hence kill cancer.

In the current fractionated radiation therapy treatment workflow, patients are treated daily with the same plan during a 1-2 month course of treatment, whereas tumor changes at the macro-morphological and micro-environment levels are not taken into consideration. However, treatment plans are usually generated based on population-based guidelines, and may not be well suited for each individual patient. Therefore, treatment efficacy is compromised if the tumor response is not taken into consideration. Using the patient-specific response as feedback to modify the treatment, such as dose-escalation for radioresistant tumor subregions or de-escalation for well-responding subregions, or canceling the post-therapy surgery if necessary, becomes appealing in improving locoregional tumor control and reducing critical structure toxicity[2].

Effort in customizing the treatment plan based on tumor response is limited largely due to the lack of tumor response information. Diffusion-weighted imaging (DWI) has been shown great promise in tumor diagnosis and early response assessment[3–6]. By quantifying the random Brownian motion of water molecules at voxel level, diffusion MRI provides tissue morphofunctional information including cellular density which is more sensitive than traditional

tumor size/morphology-based information[7,8]. With the commercialization of MRI-guided radiotherapy systems[9,10], monitoring tumor response during the course of treatment through functional MRI, and modifying treatment plan based on tumor feedback becomes logistically feasible. It is, therefore, encouraging to modify treatment strategy based on diffusion MRI, once the relationship between apparent diffusion coefficient (ADC) changes and tumor response is confirmed.

Despite promises, there are challenges to adopt diffusion imaging to response assessment and adaptive radiotherapy. Clinical diffusion MR images are acquired almost exclusively using the diffusion-weighted single-shot echo-planar-imaging (DW-ssEPI) sequence. Since an entire 2D slice is acquired in one shot after the diffusion encoding, DW-ssEPI has the advantages of fast acquisition speed and robustness to motion artifacts. However, at the same time, the DW-ssEPI sequence is sensitive to field inhomogeneity and magnetic susceptibility due to low bandwidth in the phase-encoding direction, which could lead to several geometric distortion[11]. This is particularly problematic for adaptive radiotherapy because any geometric inaccuracy directly translates to miscalculated radiation dose and potentially radiation target miss. Therefore, the development of a reliable, accurate and distortion-free diffusion MRI technique is crucial for longitudinal tumor response evaluation and adaptive treatment management.

Many techniques have been developed to alleviate EPI-related distortion, including the reduced field of view diffusion imaging, interleaved EPI, and readout segmented EPI[12,13]. However, the reduced field of view technique is not suitable for radiotherapy applications, as a large spatial coverage is required for planning purpose. In segmented EPI techniques, the residue distortion and susceptibility-related artifacts are inversely related to the number of segments. To achieve desirable imaging quality, the number of segments should be high, which means the scan time

could be very long and may not be efficient for the busy clinic schedule. Therefore, the first purpose of this dissertation is to develop an efficient diffusion sequence that has high geometric fidelity for radiation therapy purposes.

The second half of this dissertation focuses on preliminary study of using longitudinal DWI for treatment response prediction on sarcoma patients. Soft tissue sarcomas (STS) are a rare group of tumors and are heterogeneous with regard to location, histologic type, and prognosis[14]. Deep soft parts of the limbs and abdomen are the most frequent sites. More than 50 types and subtypes of sarcomas have been recognized by the World Health Organization (WHO). According to the American Cancer Society, approximately 12,750 new STS cases will be diagnosed in the United States for 2019 and approximately 5,270 Americans are expected to die from it[15]. The prognosis and overall survival of STSs are mainly affected by local recurrence and distant metastasis. Five-year distant metastases rate ranges from less than 30% to greater than 50%, and the lung is the main site of metastases[16–19].

As the biology and histology of sarcoma patients varies a lot, their response to radiation treatment could also vary significantly. Therefore, for soft tissue sarcoma patients treated with pre-operative radiotherapy, predicting radiation therapy response provides a window for treatment adaptation, and is essential in customizing the follow-up treatment and surgery timing. Core needle biopsies and fine-needle aspiration are the most widely used approaches to assess tumor grading. However, those methods are invasive and prone to sampling bias[20]. Surgical biopsy provided more accurate estimation, whereas the surgery could be non-suitable for some patients and the information is only available after the tumor removal. Therefore, an accurate and non-invasive treatment response estimation method is crucial in achieving personalized treatment.

Studies have shown that baseline ADC or change of ADC is predictive of the final response on sarcoma patients[21–24]. However, in most of the previous studies, the mean ADC value was the only feature extracted from the ADC map, whereas other potentially more predictive features were not explored. In addition, most of the previous works only performed simple univariate correlation to show the correlation between ADC and treatment outcome instead of building a predictive model that could provide additional information for clinical decisions. Moreover, DWI was only acquired pre- and/or post- RT in previous studies. It is promising to acquire more data throughout the treatment to capture tumor dynamic response to the radiation for improved outcome prediction, and to investigate the optimal imaging time for response study. Therefore, in the second half of this dissertation, we will focus on extracting more information from longitudinal diffusion data and establishing predictive models for improved treatment response prediction on soft tissue sarcoma patients. Radiomics and deep learning approaches were explored to predict treatment response using longitudinal DWI.

1.2 Organization of the thesis

In chapter 2, background and concepts that are relevant to the subsequent chapters were introduced. These include: background about diffusion-weighted MRI, single-shot EPI readout and its associated artifacts, interleaved and segmented EPI readout and its pros and cons, the difference between diffusion-weighted strategy and diffusion-prepared strategy, the concept of radiomics, and architecture of the auxiliary classifier generative adversarial network (ACGAN) network and the VGG-19 network.

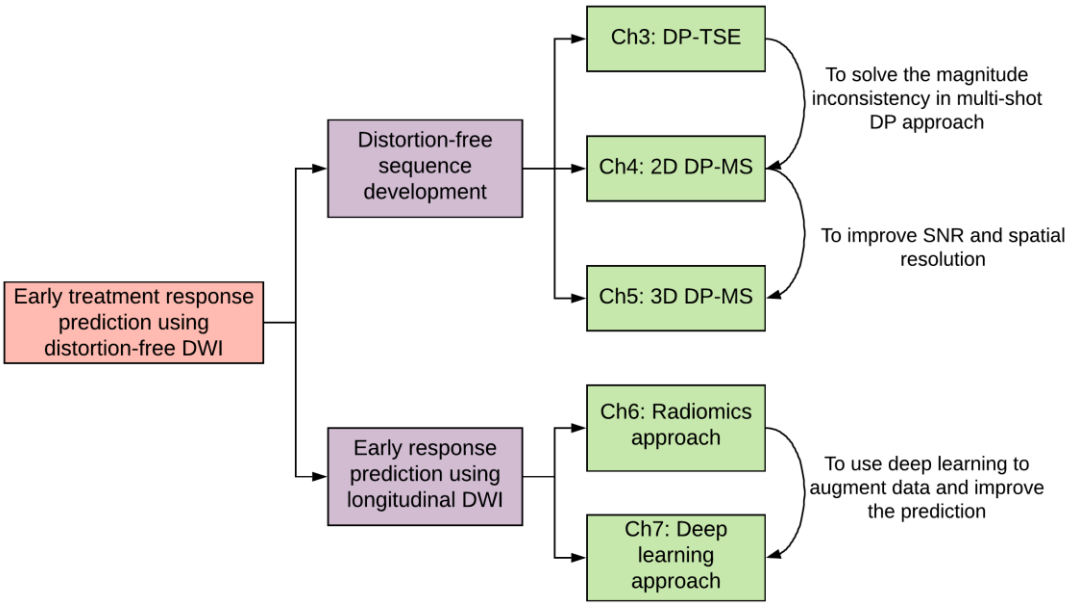

Figure 1-1 Outline of technical developments in this work.

The conventional DW-ssEPI sequence suffers from distortion, susceptibility artifact, and limited resolution problem, which prevent its clinical usage in the radiation therapy field for treatment planning, treatment adaptation and response assessment. Figure 1-1 shows the outline of this work to achieve the ultimate goal of early treatment response prediction using distortion-free DWI. Novel sequence development and treatment response prediction using conventional DWI were investigated in parallel. We first focused on the development of distortion-free diffusion sequences. In chapter 3, a diffusion-prepared turbo spin-echo readout sequence (DP-TSE) was developed and compared with the DW-ssEPI sequence on a 0.35T MRI-guided radiotherapy system. Repeated measurements were conducted on a diffusion phantom to verify the ADC accuracy and repeatability of the proposed sequence. However, multi-shot diffusion-prepared imaging suffers from magnitude inconsistency, which has not been solved in the literature. In Chapter 3, the k-space view ordering was modified to mitigate the shot-to-shot inconsistency. However, the inconsistency still exists and would affect the quantification if motion is strong.

Therefore, in chapter 4, we focused on solving the k-space magnitude inconsistency problem in multi-shot diffusion-prepared imaging. A novel magnitude stabilizer approach was proposed to covert the malignant magnitude inconsistency into phase inconsistency, which is easier to be managed. Combined with existing phase correction techniques, the proposed diffusion-prepared magnitude-stabilized balanced steady-state free precession sequence (DP-MS-bSSFP) provided distortion-free high-quality 2D diffusion-weighted and diffusion tensor images. The DP-MS-bSSFP sequence was then modified to 3D to achieve high-resolution high-quality diffusion imaging in Chapter 5. 3D imaging has the potential of achieving higher signal to noise ratio (SNR) and higher resolution, which is beneficial to overcome partial volume effect and provide more accurate tumor detection and delineation. The proposed 3D diffusion sequence was validated on the diffusion phantom and healthy volunteers in chapter 5.

Chapter 6 and chapter 7 focused on preliminary study of using longitudinal diffusion MRI for treatment response assessment on sarcoma patients treated with preoperative radiotherapy. As discussed in the motivation session, most of the existing studies rely on univariate analysis of ADC with patient outcome without exploring other quantitative features or models to predict treatment response. Whereas a predictive model enables early detection of responder vs. non-responder, which opens a window for treatment adaption. Therefore, radiomics and deep learning were investigated for response prediction. Longitudinal diffusion MRI was acquired three times throughout the hypofractionated RT treatment on soft tissue sarcoma patients to capture the tumor cellularity change during the treatment. In Chapter 6, radiomics features were extracted and the most predictive features were identified. A logistic regression model and a support vector machine model were built to predict the treatment response. Optimal imaging acquisition timing for response assessment was also investigated. In Chapter 7, to alleviate the small data size nature of

medical data, an ACGAN network was implemented to synthesize longitudinal ADC maps. A pre-trained convolutional neural network VGG-19 was modified to predict treatment response.

Finally, Chapter 8 concludes the dissertation with a summary of the technical development and outlook of future research.

Chapter 2 Background

In this Chapter, background about diffusion-weighted imaging, EPI readout, diffusion-weighted and diffusion-prepared approach, radiomics, ACGAN network, and VGG-19 network are introduced. This chapter is not meant to be a thorough summary of all the topics, but to familiarize the readers with relevant aspects of the subsequent chapters.

2.1 Diffusion-weighted imaging

Diffusion-weighted imaging is one special MRI technique that reflects the rate of random Brownian motion of water molecules in the extra-cellular space hence reveals tissue cellularity information. It has been widely used in neurology and oncology, and has shown great potential in stroke imaging, lesion detection and characterization, and tumor response assessment[6,25,26].

A diffusion imaging sequence is composed of the diffusion encoding part and the readout part. In the diffusion encoding component, strong diffusion gradients are applied to cause attenuation of the proton signal. The degree of the signal level decreases exponentially with the diffusion coefficient D, which is related to the intrinsic diffusivity and the b-value, which is determined by the diffusion gradient waveform. Therefore, the resultant signal S can be expressed as:

$$S_b = S_0 e^{-bADC}$$

, where S_0 is the signal intensity without diffusion gradients, and ADC is the apparent diffusion coefficient.

Therefore, ADC map could be calculated if at least 2 different b-values were acquired:

$$ADC = \frac{\ln (S_2/S_1)}{b_1 - b_2}$$

For a gradient waveform $G(t)$, if the gradient is applied along a single axis, the b-value can be calculated as:

$$b = \gamma^2 \int_0^{TE} \left[\int_0^t G(t')dt' \right]^2 dt$$

, and γ is the gyromagnetic ratio.

The most commonly used diffusion gradient waveform is the spin-echo pulse sequence, or the Stejskal-Tanner gradient. (Shown in Figure 2-1 (a)). In this waveform, two gradients with equal size and polarity are placed before and after the refocusing pulse to create diffusion weighting. The b-value with a pair of trapezoidal gradients is:

$$b = \gamma^2 G^2 [\delta^2 \left(\Delta - \frac{\delta}{3} \right) + \frac{\xi^3}{30} - \frac{\delta \xi^2}{6}]$$

, where G is the gradient amplitude, δ^2 is the duration of diffusion gradient (ramp time plus plateau duration), ξ is the ramp time, and Δ is the time interval between the start time of the two gradient lobes.

Another type of gradient waveform is the gradient-echo pulse sequence, where the gradients are of the same size but opposing polarity. (Figure 2-1 (b)) The corresponding b-value is:

$$b = \gamma^2 G^2 [\frac{2\delta^3}{3} + \delta^2 \xi + \frac{\xi^3}{30} - \frac{\delta \xi^2}{6}]$$

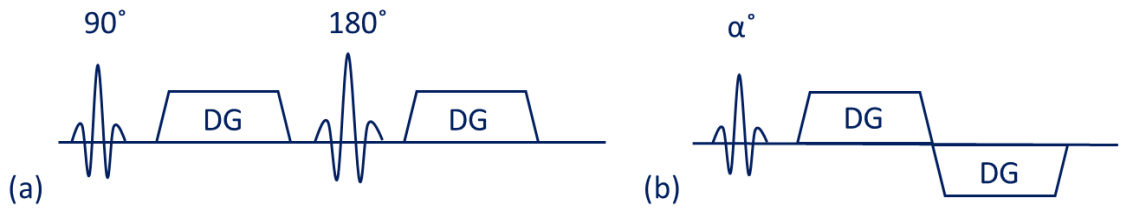

Figure 2-1 Gradient waveforms for the (a) spin-echo pulse sequence, and (b) gradient-echo pulse sequence. DG represents the diffusion gradient.

2.2 Echo planar imaging

Echo planar imaging (EPI) is one of the fastest imaging techniques. In the single-shot EPI sequence, the spatial encoding of an entire k-space can be completed after a single radiofrequency (RF) excitation. Therefore, it is capable of generating an image in a few tens of milliseconds. Because of the rapid acquisition speed and capability to reduce motion artifacts, EPI has been used as the readout module in diffusion imaging.

A standard gradient waveform of the diffusion-weighted sequence with blipped EPI readout was shown in Figure 2-2 (a). The readout direction starts with a pre-phasing gradient, then followed by the readout gradient train with alternating polarity to acquire the signal. The phase encoding direction also starts with a pre-phasing gradient that moves the k-space sampling to the desired ky location. Then a series of gradient blips with the same polarity and area are played to sample the whole ky space. An entire k-space can be acquired in this fashion. Figure 2-2 (b) shows the corresponding k-space trajectory.

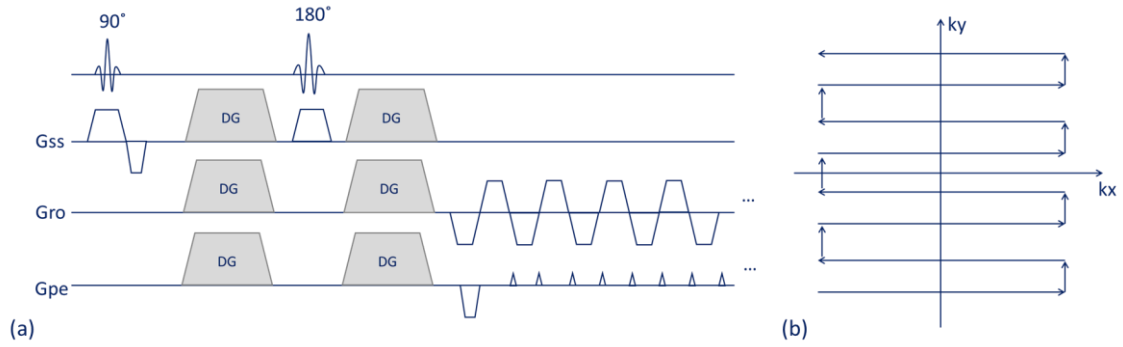

Figure 2-2 (a) gradient waveform of diffusion-weighted EPI sequence. (b) corresponding k-space trajectory. Gss, Gro, Gpe represent the gradient direction of slice selective, readout and phase encoding respectively.

Although the fast acquisition speed and robustness to motion is appealing, there are several artifacts associated with the single-shot EPI technique. The first common artifact is the Nyquist ghosting. Because the blipped EPI technique acquired the whole k-space in a zig-zag acquisition, eddy currents and other imperfection could cause k-space data displacement between even and odd echoes. This will lead to a ghost image with a shift with N/2 pixels relative to the desired image location. To overcome this, reference lines with zero phase encoding were acquired to estimate and compensate the phase difference between even and odd echoes[11].

Another drawback of the EPI readout is the spatial distortion. Since the whole k-space is acquired in one shot, the effective phase encoding bandwidth is very low. Therefore, any off-resonance effects, such as field inhomogeneity, eddy currents, and chemical shift will lead to severe distortion or chemical shift artifact in the phase-encoding direction. Various correction methods have been proposed to address the distortion problem including field mapping correction, multi-reference technique, reversed gradient correction, point spared function based correction, and image processing based techniques that do not require additional scan, etc[27–36].

Another approach to resolve the distortion and chemical shift artifacts is to change the single-shot readout into multi-shot. This also brings the advantages of improved spatial resolution and reduced T2*-related blurring. In the multi-shot EPI approach, the whole k-space is filled in separately with several echo trains instead of one echo train to increase the bandwidth in the phase encoding direction. There are two main strategies of k-space segmentation: one in the phase encoding direction and the other in the readout direction. (See Figure 2-3). The multi-shot EPI approach greatly reduced the distortion and improved the spatial resolution of the single-shot EPI technique. However, the acquisition time is proportional to the number of shots. In addition, multi-shot leads to a k-space inconsistency problem in diffusion imaging. This will be discussed in the next session.

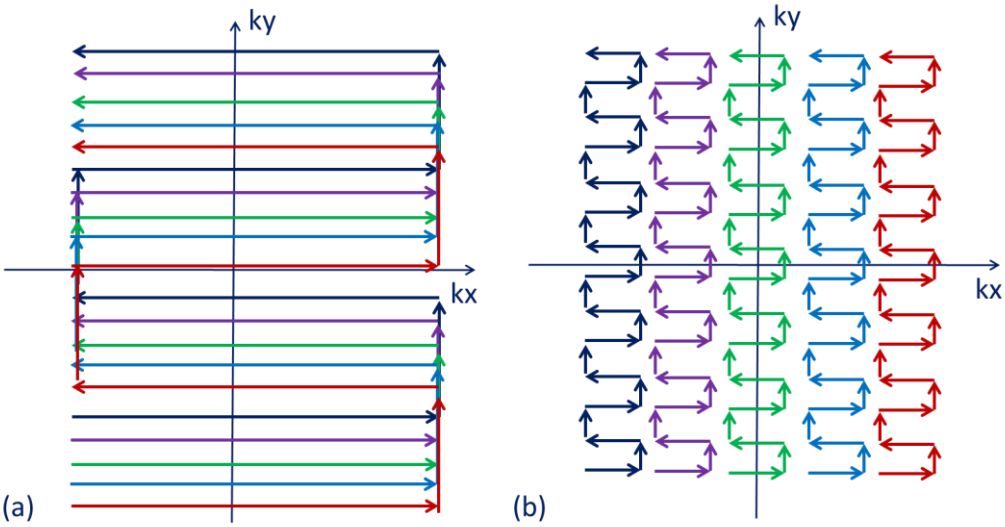

Figure 2-3 Interleaved EPI (a) and readout-segmented EPI (b). Each color represents one echo train. In both examples, the k-space was filled in with five shots.

2.3 DW and DP

Depending on whether the diffusion-sensitized information is stored in the transversal or longitudinal direction prior to the imaging readout, diffusion imaging sequences can be classified

as either diffusion-weighted (DW) or diffusion-prepared (DP) acquisitions. In the diffusion-weighted imaging approach, the imaging readout, usually EPI, immediately follows the diffusion encoding module to acquire the diffusion sensitized transverse magnetization (Figure 2-4 (a)). Whereas in the diffusion-prepared approach, a -90˚ pulse is applied to flip the transversal diffusion information into the longitudinal direction. Spoil gradients were followed to spoil the imperfect components (Figure 2-4 (c)). This diffusion-prepared approach has the flexibility of combining with many types of readout including bSSFP, which has the promise of achieving fast 3D distortion-free diffusion imaging[37,38].

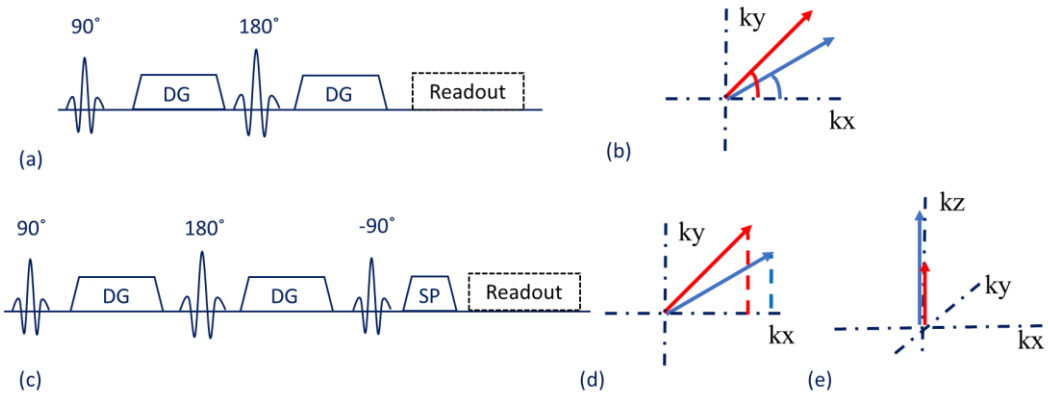

Figure 2-4 Difference between diffusion-weighted approach and diffusion-prepared approach. (a) sequence diagram of the DW approach. (b) example of phase inconsistency between two different shots in the DW approach. (c) sequence diagram of DP approach. (d) (e) examples of magnitude inconsistency in the DP approach before and after the -90˚ flip back RF pulse. The red and blue arrows represent the signal vectors after the diffusion encoding for two different shots.

In multi-shot diffusion imaging, the use of strong diffusion gradients magnifies motion and system imperfections and created strong Eddy currents, which then leads to different signal phase accrual for different k-space segments. As shown in Figure 2-4 (b), this different signal phase accruals are exhibited as a k-space phase inconsistency problem in the DW approach. Various reconstruction methods have been proposed to mitigate the phase variation problem[39–43]. In the DP approach, only the sine or cosine competent is flipped to the longitudinal direction, leading to a

shot-to-shot magnitude inconsistency (Figure 2-4 (d)(e)). This magnitude variation is more challenging to be alleviated due to the complementary parts have already been spoiled and lost. High-order motion compensation, advanced pulse, and twice refocused scheme were frequently adopted in DP approach to mitigate k-space inconsistency originated from rigid motion, B1 inhomogeneity, and Eddy currents[38,44]; however, the inconsistency problem originated from non-rigid motion and other imperfections still exists and could result in inferior image quality and inaccurate quantification if left uncorrected.

2.4 **Radiomics**

Radiomics has become an emerging translational field for quantitative imaging analysis. Instead of treating images as qualitative pictures, the images are viewed as meaningful and mineable data[45]. High-dimensional quantitative features are extracted from images, and those features potentially contain predictive information that is not obvious to human eyes.

In the radiomics workflow, regions or volumes of interest (ROIs or VOIs), such as the tumor region, are first segmented. Quantitative features are then extracted based on this segmented region. Many software and open-source toolboxes are available to perform the feature extraction. Such as PyRadiomics[46], IBEX[47], MaZda[48], and more[49,50]. Statistical analysis and model building are then performed to construct diagnostic and predictive models for detection, classification, and clinical outcome prediction.

There are several categories of radiomics features. The first group is shape features, which describe the geometric properties of the interested region. This includes volume, surface area, sphericity, compactness, etc. The second group is first-order statistics features, which is purely based on the histogram-based signal intensity of the interested region. Examples of first-order

features include mean, median, maximum, minimum, skewness, kurtosis, entropy, etc. The third group of features is second-order statistics features, which is also called texture features. To calculate texture features, a specific matrix that reflects the spatial relation of pixel intensity is first obtained. Features are then calculated based on this new matrix. Those matrixes include Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Neighbouring Gray Tone Difference Matrix (NGTDM), and Gray Level Dependence Matrix (GLDM) etc. Details about the matrix and feature definitions can be found at the reference manual provided by the Imaging Biomarker Standardization Initiative (IBSI)[51]. The last group of radiomics features is the higher-order statistics features, which are obtained after applying a filter, such as wavelet transform, to the original images. Depending on the choice of filter, those features can help to identify repetitive patterns, suppress noise, or highlight details[52].

2.5 Deep Learning

The past decade has seen the exponential growth of deep learning in medical image analysis. It has become the method of choice for image segmentation, detection, classification, and registration for various disease sites[53–55]. Several studies have applied deep learning to predict treatment response using CT images on lung, rectal, and bladder patient[55–57]. Promising results outperforming the conventional linear regression or support vector machine were obtained. In this session, two networks that are used in Chapter 7 will be briefly introduced.

GAN

Generative adversarial network (GAN) is a deep learning framework developed by Ian Goodfellow et al[58]. It achieved state-of-the-art performance in image generation and has been used

as a data augmentation tool for many medical applications such as image segmentation and lesion classification[59–62].

As shown in Figure 2-5(a), in the GAN network, two networks were trained in an adversarial process: one discriminator network was trained to discriminate between real and fake images, whereas the other network was trained to generate fake images to fool the discriminator. High-quality realistic images were synthesized via this competing process. The difference between GAN and conditional GAN (cGAN)[63] is that the image generator and discriminator of the latter are conditional on the class label (Figure 2-5(b)). Therefore, images of a given class can be generated. In the Auxiliary classifier GAN (ACGAN)[64], the class label information is feed into the generator, and an auxiliary node is added to the discriminator to output the class label (Figure 2-5(c)).

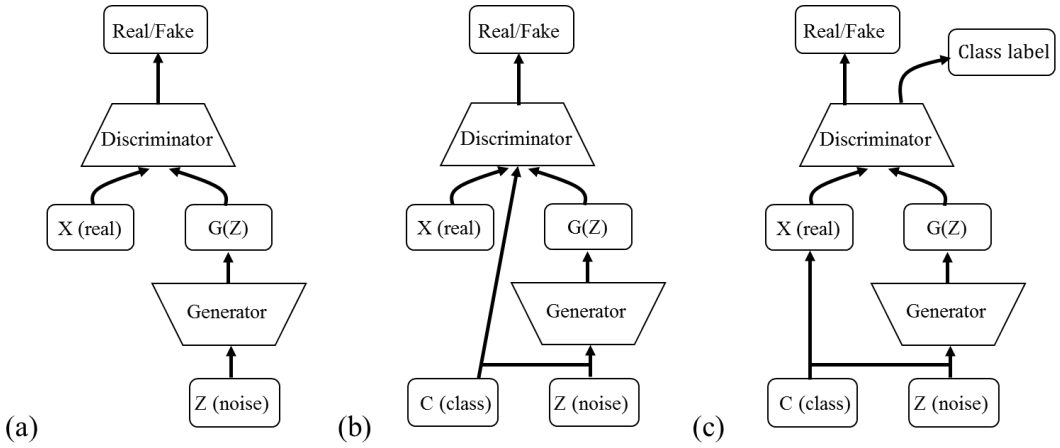

Figure 2-5 Architectures of (a) GAN, (b)cGAN, and (c) ACGAN.

VGG-19

VGGNet is a group of deep convolution network developed by the Visual Geometry Group for image recognition[65]. It won the first and second places in the 2014 ImageNet Challenge for localization and classification respectively.

In this work, we used VGG-19 network, which consists of 19 layers that the weights are learned. The architecture of the network is shown in Figure 2-6. The network has 16 convolutional layers and 3 dense layers. The network only uses 3x3 filters (conv3) but with multiple layers to cover the desired area but with a reduced number of parameters. The number of filters increases from conv3-64, to conv3-128, to conv3-256 to conv3-512 and conv3-512, and max-pooling layers are applied at different steps to extract informative features. The three fully connected dense layers have 4096, 4096, and 1000 nodes respectively. The VGG-19 network has a total of 144 million of parameters. It was trained on over a million 224x224 RGB images from the ImageNet dataset.

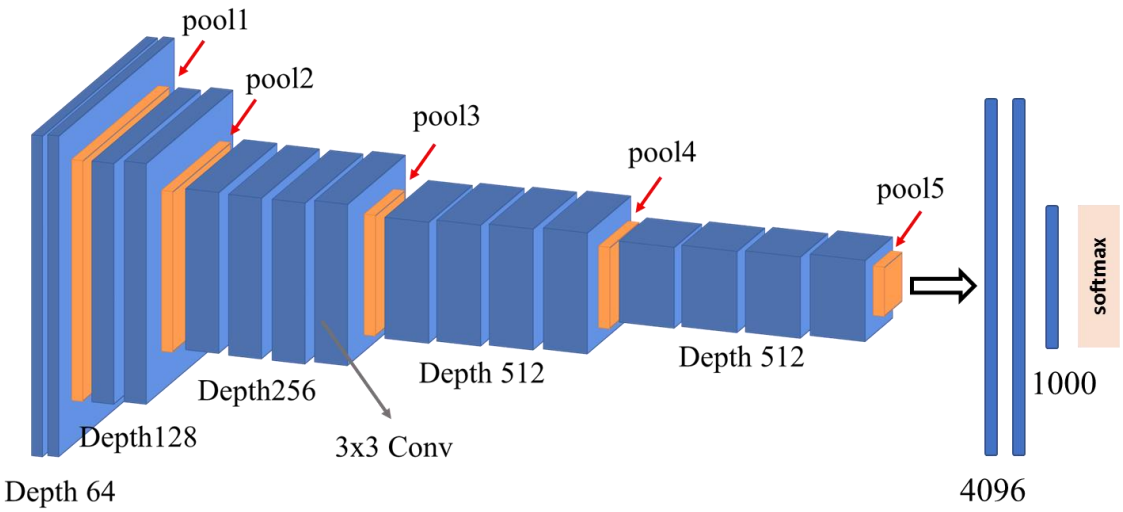

Figure 2-6 Architecture of the VGG-19 network

Chapter 3 Diffusion-prepared TSE sequence

Monitoring tumor response during the course of treatment and adaptively modifying treatment plan based on tumor biological feedback may represent a new paradigm for radiotherapy. Diffusion MRI has shown great promises in assessing and predicting tumor response to radiotherapy. However, the DW-ssEPI technique suffers from limited resolution, severe distortion, and possibly inaccurate ADC quantification at low field strength. Therefore, in this chapter, we sought to develop a reliable, accurate and distortion-free diffusion MRI technique that is practicable for longitudinal tumor response evaluation and adaptive radiotherapy on a 0.35T MRI-guided radiotherapy system. A diffusion-prepared TSE sequence was programmed and tested on a spatial integrity phantom, a diffusion phantom, and eight patients to evaluate its geometric fidelity and ADC accuracy.

A version of this chapter has been published in the Medical Physics Journal: Gao Y, Han F, Zhou Z, Cao M, Kaprealian T, et al. Distortion-free diffusion MRI using an MRI-guided Tri-Cobalt 60 radiotherapy system: Sequence verification and preliminary clinical experience. Med Phys. 2017 Oct 1;44(10): 5357–66.

3.1 Introduction

Diffusion MRI is a promising imaging technique for assessing and predicting tumor response to radiotherapy[6], often well before macroscopic changes in tumor size and morphology[7,8]. The changes in the ADC during the course of radiotherapy have been shown to correlate with tumor control and treatment outcome[4,66,67]. Therefore, diffusion MRI is an important imaging biomarker for tumor response assessment for which the treatment plan can be adaptively altered during

therapy. As such, functional MRI-based adaptive radiotherapy therapy strategy has the potential of improving treatment outcome.

To enable the widespread use of diffusion MRI-guided adaptive therapy, the optimal timing for imaging needs to be established. A number of previous studies of diffusion MRI for predicting tumor response to radiotherapy focused on performing diffusion MRI at selected time points before, during and after the radiotherapy[4,67–71]. More recently, a longitudinal diffusion MRI approach has been proposed by Yang et al. [72]. In this approach, diffusion MRI is acquired every 2-5 days throughout the course of radiotherapy using an MRI-guided radiotherapy system, and a longitudinal ADC curve is constructed based on the diffusion measurements at multiple time points. Despite its promise, most of existing studies used the DW-ssEPI sequence, which suffers from low spatial resolution and severe geometric distortion due to its single-shot EPI readout. This is particularly problematic for radiotherapy as any geometric inaccuracy could directly translate to miscalculated radiation dose and potential radiation target error if adaptive treatment planning is based on these diffusion images. In addition, at low field strength, the conventional DW-ssEPI becomes inaccurate when high b-value is used due to SNR loss associated with the lower field strength, which may be problematic for fully characterizing tumors with low diffusivity. It is therefore highly desirable to develop a distortion-free diffusion MRI technique with higher SNR for tumor response evaluation and potentially adaptive radiotherapy.

Various techniques have been developed to increase the resolution or alleviate the distortion associated with DW-ssEPI. One approach uses reduced field-of-view diffusion imaging[12], which is not applicable to treatment response evaluation and plan modification since a large spatial coverage is necessary for radiotherapy purposes. A diffusion-weighted line-scan technique has been used to mitigate distortion; however, it suffers from low SNR and limited resolution[73]. The

reserved gradient method uses two images with reversed gradient encoding to calculate the distortion-free image[74]. The problems of this method include limited resolution and the requirement to identify the boundary which could be challenging at low field and high b-value scenarios. Readout-segmented EPI has been proposed to partially mitigate distortion and improve resolution[13]; however, it still suffers from distortion due to the phase error along with the EPI readout.

In the current study, we sought to develop a reliable, accurate and distortion-free diffusion sequence based on a turbo spin-echo acquisition that is practicable and specifically designed for assessment of tumor response to radiotherapy.

3.2 Methods

3.2.1 DP-TSE Pulse Sequence Design

The diffusion sequence was programmed for an MRI-guided radiotherapy (MRgRT) system (MRIdian, ViewRay, Mountain View, CA) with a pulse sequence programming software (IDEA version VB19, Siemens Medical Solutions, Erlangen, Germany). As illustrated in Figure 3-1 (a), the diffusion sequence used in this study combines a diffusion preparation module with a segmented turbo spin-echo readout module, where similar approaches have been previously explored in diagnostic imaging at higher field strength (1.5T or higher)[38,75,76]. The diffusion module, which was inserted at the beginning of each k-space readout segment, uses a twice-refocused spin-echo (TRSE) with diffusion gradients to generate diffusion weighting. A second refocusing pulse was added in the diffusion preparation module and the timing for each gradient was optimized to reduce eddy current effects from the diffusion gradients[77]. Crusher gradients were placed before and after each 180° RF pulse to spoil simulated echoes from imperfect

refocusing. At the end of diffusion encoding, the diffusion-encoded transverse magnetization was flipped back to the longitudinal direction by a -90° RF pulse, which is immediately followed by spoilers to eliminate the remaining transverse magnetization. This strategy only modulated the longitudinal magnetization with diffusion weighting and no residual phase exist during the TSE readout so that the Carr-Purcell-Meiboom-Gill (CPMG) condition of TSE sequence is not violated[76]. To mitigate the signal inconsistencies between different imaging shots within the k-space, a problem frequently seen in multi-shot diffusion imaging, k-space view ordering was modified such that the central k-space is acquired by the same imaging shot [78]. Figure 3-1 (b) is an example of a 4-shot k-space view ordering scheme where arrows from left to right indicate the acquisition order. Lines with the same color were acquired with the same shot.

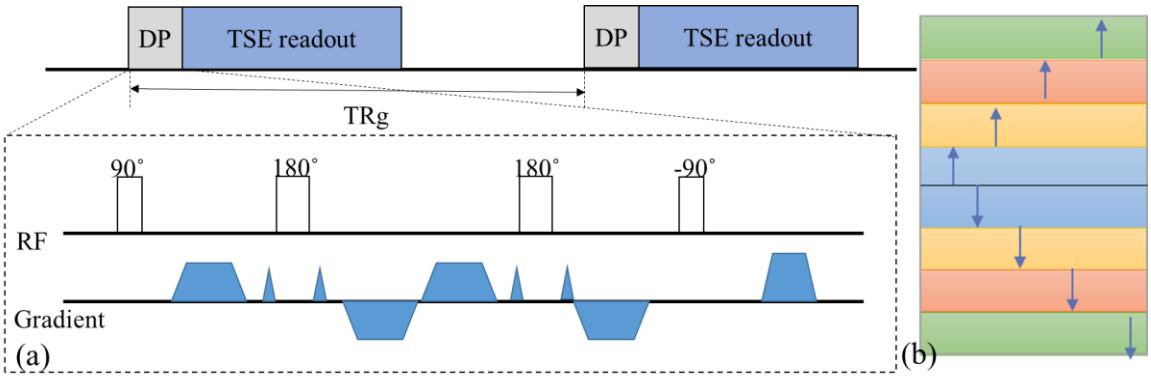

Figure 3-1 Sequence diagram (a) and an example of a 4-shot k-space view ordering (b). The twice-refocused spin-echo scheme was used to generate diffusion weighting while minimizing eddy currents. K-space view ordering was modified such that the central k-space is acquired by the same imaging shot

3.2.2 **Quantitative Phantom Experiment**

The spatial integrity, ADC accuracy and ADC reproducibility of the proposed DP-TSE technique were evaluated and compared with the standard DW-ssEPI using the 0.35T ViewRay MR scanner with 12-channel surface coils.

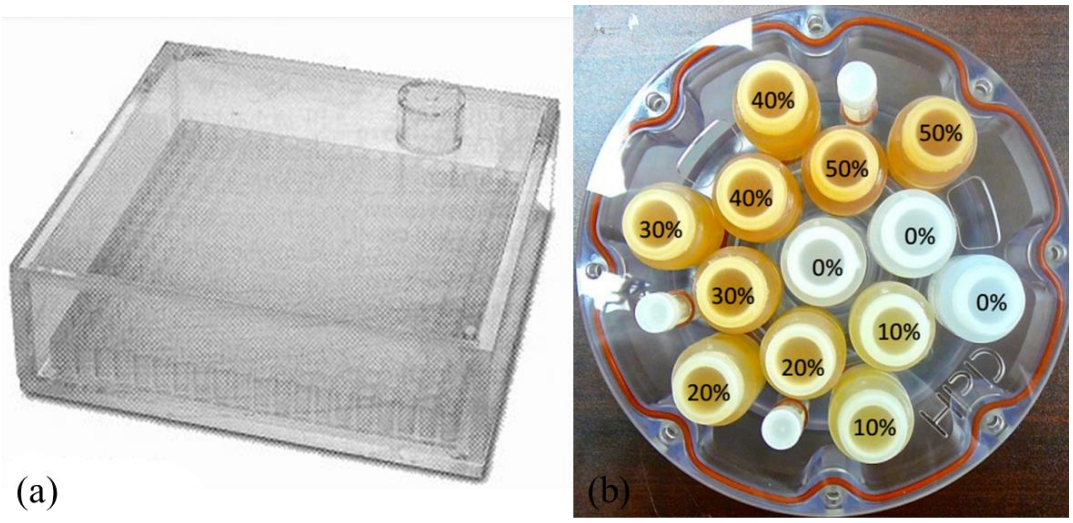

Figure 3-2 Spatial Integrity phantom (a) and diffusion phantom (b). The spatial integrity phantom is of dimension 33.02×33.02×10.16cm^3 and consists of 197 cylindrical grids (20×20 with 3 missing for positioning and detecting). The diffusion phantom is 194mm in diameter and has 13 vials with six different diffusivities. Vials with diffusivity less than 0.35×10^{-3} mm^2/s were not included in this study.

Spatial Integrity

A uniformity and linearity phantom (Fluke Biomedical, model 76-907, Figure 3-2 (a)) was used to assess the spatial integrity of the proposed technique under three orthogonal orientations: transverse, coronal, and sagittal. DP-TSE and DW-ssEPI images on all three orientations were acquired with FOV =400×400mm^2. Parameters for the DW-ssEPI/DP-TSE sequences were: TR=2000/2000ms; TE=160/115ms; matrix size=128/192; number of averages=10/2, readout bandwidth =752/789 Hz/pixel, echo train length (ETL) was 24 for DP-TSE (8 shots). A spatial integrity analysis tool provided by ViewRay was used to analyze the images by detecting the center location of each cylinder in the phantom based on the image and comparing the image-based location measurements with the known ground truth. Automatic rotation and translation were detected to account for setup imperfection. The geometric integrity passing criteria for each cylinder is 1mm within a 100mm radius and 2mm within a 175mm radius.

ADC Accuracy

ADC accuracy of the proposed technique was compared with the conventional DW-ssEPI using a commercially available diffusion phantom (High Precision Device, Inc. Model 128). The phantom is 194mm in radius and consists of 13 vials filled with aqueous solutions of polymer polyvinylpyrrolidone (PVP) at 0%, 10%, 20%, 30%, 40% and 50% concentration levels (Figure 3-2 (b), abbreviated as vial 1 to vial 6). The weighted mean ADC of each vial at 0°C was provided by the manufacturer as a reference.

Both 0°C and room temperature studies were conducted to cover a relatively wide diffusivity range. In the 0°C study, the phantom was filled with crushed ice one day before the measurement and stored in a refrigerator overnight. The phantom was in the ice water bath during the imaging experiment to ensure 0°C temperature. In the room temperature (~21.0°C) study, the phantom was placed in the scanner room at least two days before the scan to allow the phantom to reach room temperature. The temperature of the phantom was measured before and after each study using an extra-long-stem thermometer through the top fill port. Only the first nine vials and the first eleven vials were analyzed in the 0°C and room temperature studies, respectively. Vials with diffusivity lower than 0.35×10^{-3} mm^2/s, a number that is below the range of physiologically possible tumor ADC values, were not included.

In each study, the DW-ssEPI data were acquired at a medium resolution (MeR) of $2.34\times2.34\times5mm^3$ (FOV=$300\times300mm^2$, matrix size=128) using TR/TE=2000/160ms, bandwidth=752Hz/pixel and 17 averages. The DP-TSE data were acquired at both the same medium resolution and a higher resolution (HiR) of $1.56\times1.56\times5mm^3$ (FOV=$300\times300mm^2$, matrix size=192). The DP-TSE sequence parameters included: TR/TE=2000/115ms, 2 averages, bandwidth =789Hz/pixel, ETL=16 for the medium resolution and 24 for the high-resolution. For both DW-ssEPI and DP-TSE, four b-values (0, 200, 500, 800 s/mm^2) were applied along the

readout direction and a single TE was used for all b-values within the same sequence to remove the influence of T2 on diffusion quantification. The total scan time was 140s for DW-ssEPI and 144s for DP-TSE. For both sequences, image averaging was performed in magnitude only as complex averaging of diffusion data in k-space has been shown to cause signal loss and artifacts[79]. To obtain the reference ADC values at room temperature, the phantom was scanned with a 3T scanner (Prisma, Siemens Medical Solution) using the standard DW-ssEPI sequence with FOV=300×300mm^2, matrix size=160, TR/TE=4200/76ms, bandwidth=1564Hz/pixel, b=0, 200, 500, 800, and 8 averages.

The noise subtraction method proposed by Dietrich et al.[80] was performed before ADC fitting to remove the influence of noise floor in ADC quantification at low SNR scenario:

$$S_{nc} = \sqrt{\left| S_n^2 - \frac{2}{\pi} N^2 \right|}$$

, where S_n and S_{nc} are noisy and noise-corrected images, N represents the average background noise signal intensity. ADC map was then calculated using the least-squares fitting of the mono-exponential curve:

$$S_{nc}(b) = S_0 e^{-b \times ADC}$$

, where $S_{nc}(b)$ is noise-corrected signal intensity, b is the b-value controlled by the gradient waveform, S_0 and ADC are parameters to be fitted which correspond to signal intensity without diffusion and apparent diffusion coefficient value respectively.

ADC Reproducibility

To verify the reproducibility of the proposed sequence, which is essential for the robustness of the ADC-based tumor response assessment, ten room temperature studies using the high-resolution DP-TSE protocol were carried out on five different days over three weeks with the same phantom setup. Temperatures of the phantom were recorded before and after each scan. ADC maps were generated after the noise correction, and ADC values of the first five diffusivity levels were recorded.

3.2.3 **In vivo Study**

The in vivo study was approved by our institutional review board and each subject provided written informed consent. Eight patients were recruited, including two glioblastoma (GBM) patients and six sarcoma patients. Imaging was acquired immediately after the patient's treatment (on ViewRay or other treatment systems). The DW-ssEPI sequence parameters included: FOV=350-400×350-400 mm^2, matrix size=128 × 128, b=0, 200, and 500 s/mm^2, 6 mm slices, 5 averages, and total acquisition time=32s. The DP-TSE were acquired at ~3 times finer resolution with the following sequence parameters: FOV=350-400×350-400 mm^2, matrix size=192 × 192, b=0, 200, and 500 s/mm^2, 5 mm slices, 2 averages, and total acquisition time=108s. Diffusion gradients were only applied along the readout direction at the current sequence verification stage. 12-channel surface coils were used in the sarcoma patients, and 10-channel head-and-neck coils were used in the GBM patients.

Quantitative geometric accuracy assessment was evaluated using target registration error (TRE)[81]. The reference images, namely patient simulation CT images for the two GBM patients and imaging-day clinical bSSFP MR images for the six sarcoma patients, were rigidly registered to the diffusion imaging plane prior to the analysis. Seven to twelve landmarks were selected for each patient on the reference images, the DW-ssEPI images, and the DP-TSE images, respectively,

by a radiation oncologist. These landmarks were easily identifiable structures such as eyeball, CSF, skull, and tumor in GBM patients; muscle group, bone marrow, and tumor in sarcoma patients. Four corners of large-size structures were identified to evaluate shape distortions. TRE was calculated as the Euclidean distance difference between the locations of landmarks on the diffusion images and on the reference images:

$$\mathrm{TRE} = \sqrt{\left(x_i - X_i\right)^2 + \left(y_i - Y_i\right)^2}$$

, where x_i and y_i are the coordinates of the i-th landmark on the diffusion image, and X_i and Y_i are corresponding coordinates on the reference image.

ADC accuracy was compared between DW-ssEPI and DP-TSE. In the two GBM patients, ROIs were drawn inside the surgical cavity, the CSF, and the white matter in the DW-ssEPI and DP-TSE images, respectively. For the six sarcoma patients, ROIs were drawn inside the tumor region. ADCs estimated from the two techniques were compared. For patients not treated, only imaged, on ViewRay, a tube phantom containing diluted gadolinium contrast and agarose gel was placed next to the patient when they were brought to ViewRay for imaging. The phantom was not placed on patients treated on ViewRay to avoid interference with treatments. This phantom served as quality control to monitor the stability of the longitudinal diffusivity measurements in a different study, and is used in this study to show the agreement in ADC measurement between the two sequences.

3.3 Result

3.3.1 Quantitative Phantom Experiments

Spatial Integrity

DP-TSE passed the spatial integrity test of <1mm discrepancy within a 100mm radius and <2mm discrepancy within a 175mm radius in all three orientations while DW-ssEPI did not. The errors between the detected cylinder center locations based on DP-TSE and the known ground truth locations were: 0.474mm±0.355mm in the axial plane, 0.475mm±0.287mm in the coronal plane, and 0.546mm±0.336mm in the sagittal plane. The maximum errors in the three orientations were 1.993mm, 1.281mm, and 1.388mm respectively. As shown in Figure 3-3, due to severe distortion and low signal intensity of the DW-ssEPI images, the software failed to detect the cylinder markers in any of the three orientations using DW-ssEPI images.

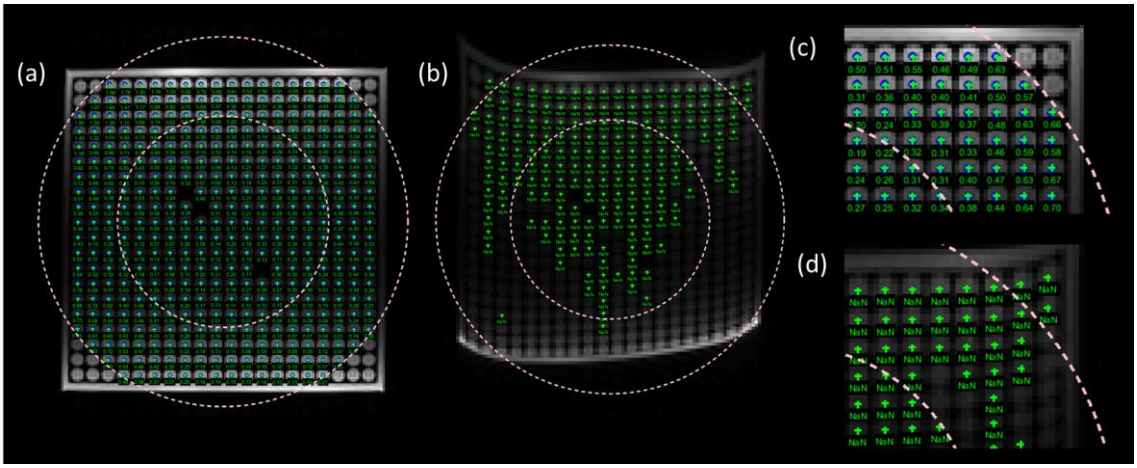

Figure 3-3 Phantom images (transverse view) acquired at iso-center using the proposed DP-TSE (a) and the DW-ssEPI (b) techniques with b=0. (c) and (d) are corresponding zoomed-in images. The green crosses represent the detected marker locations based on the image, and the blue circles are the ground truth marker locations. The DW-ssEPI image has substantial distortions. The detected location was compared with the ground truth for accuracy. A "NaN" would return if the software failed to detect the three missing markers for localization.

ADC Accuracy

Figure 3-4 shows the b=0 and b=800 images of DW-ssEPI and DP-TSE under room temperature study. Same window level was used for b=0 and b=800 respectively. The average background signal intensity of DW-ssEPI was about four times higher than that of DP-TSE, and the signal from the central pure water vial approached to the noise floor level at b=800. Compared

with DW-ssEPI, the proposed DP-TSE had higher resolution, lower background noise level, and less distortion.

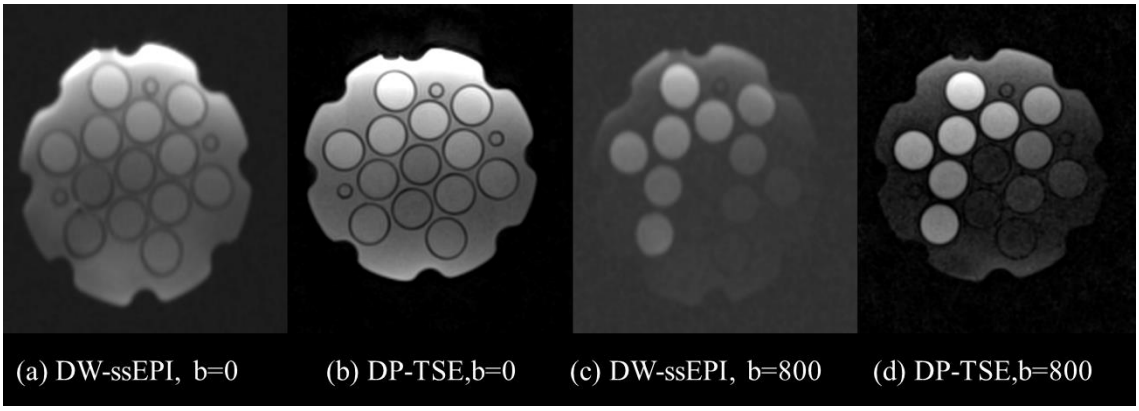

Figure 3-4 Raw images of DW-ssEPI (MeR) and DP-TSE (HiR) at b=0 and b=800.

Figure 3-5 demonstrated the effect of noise correction on the central water vial at 0°C and 21°C. Because the signal of the EPI approached the noise floor at a modest b-value of 500-800, which usually does not happen at higher field strengths, the data points before noise subtraction (blue line with circle markers in (a) and (c)) deviated from a straight line in the logarithm scale when the b-value gets into the range of 500-800 s/mm^2. This subsequently resulted in substantial underestimation of the ADC if the mono-exponential fit is used without noise correction. Following noise subtraction (green line with triangle markers in (a) and (c)), the data points displayed improved linearity. For the TSE case in the plot (b) and (d), the data before and after noise reduction were both fairly linear, mainly because the TSE signal did not approach the noise floor until much higher b-values of >1400 s/mm^2.

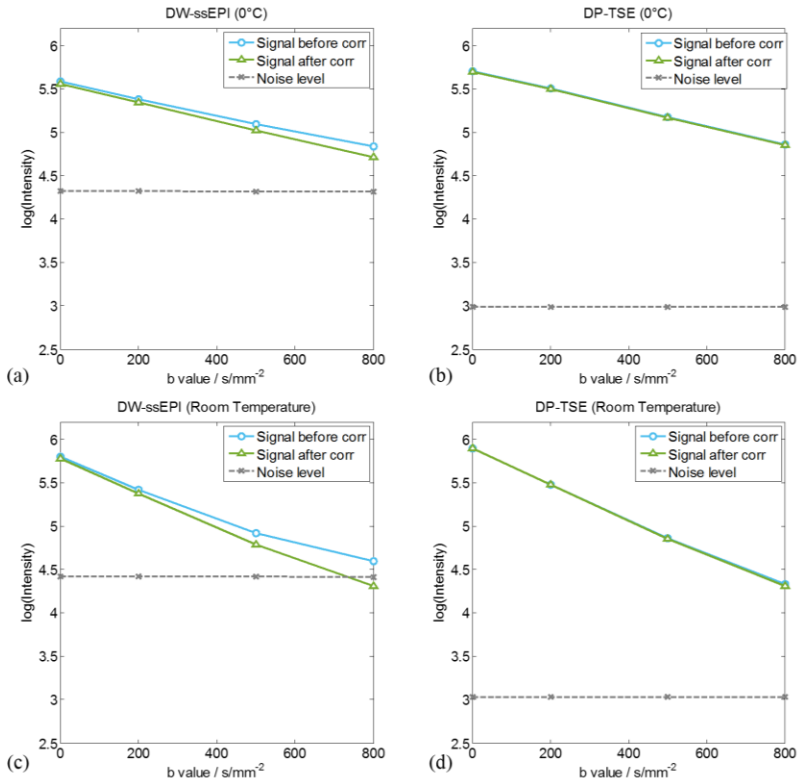

Figure 3-5 Effect of noise subtraction on ADC calculation based on DW-ssEPI (a, c) and DP-TSE (b, d) techniques at 0°C (a, b) and room temperature (c, d).

Figure 3-6 shows the ADC before and after noise correction at 0°C and room temperature (21.0°C in the ViewRay study and 21.8°C in the 3T study). Noise subtraction must be performed in DW-ssEPI, otherwise, there will be a 12% underestimation and 25% underestimation for vial 1 at 0°C and room temperature, respectively, using our fit model without noise correction. ADC elevations through noise correction were within 1.5% for DP-TSE, therefore the noise correction step is optional for DP-TSE.

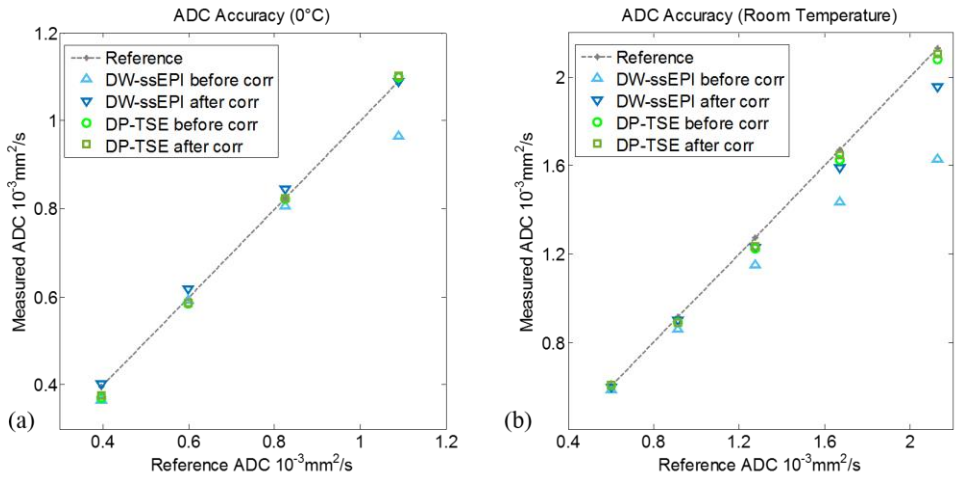

Figure 3-6 ADC Accuracy after noise subtraction for DW-ssEPI (MeR) and DP-TSE (HiR) at 0°C (a) and room temperature (b).

ADC accuracy results after noise subtraction are summarized in Table 3-1. Under 0°C, the difference between measured ADCs and reference ADCs were within 3% for both sequences. For the room temperature study, DP-TSE had higher accuracy than DW-ssEPI: discrepancies between ADCs from DP-TSE and the reference were within 4%, but were as high as 8% for DW-ssEPI. However, ADCs from DP-TSE had a higher standard deviation than DW-ssEPI, probably due to lower number of averages used.

Table 3-1 ADC accuracy results at 0°C and room temperature

	0°C				Room Temperature			
	Ref	EPI MeR	DP-TSE MeR	DP-TSE HiR	Ref	EPI MeR	DP-TSE MeR	DP-TSE HiR
Vial1	1.09	1.09±0.02	1.11±0.03	1.10±0.06	2.13	1.96±0.04	2.13±0.08	2.10±0.09
Vial2	0.82	0.85±0.02	0.83±0.02	0.82±0.04	1.67	1.59±0.03	1.64±0.05	1.63±0.07
Vial3	0.60	0.62±0.02	0.59±0.02	0.59±0.04	1.27	1.23±0.02	1.22±0.04	1.20±0.06
Vial4	0.40	0.40±0.03	0.39±0.03	0.38±0.04	0.91	0.90±0.03	0.89±0.03	0.89±0.05
Vial5					0.60	0.60±0.04	0.61±0.05	0.61±0.05

In the reproducibility study, the pre-scan temperature of the phantom was 20.68±0.28°C, and temperature changes after the scans were all within 0.2°C. Mean ADCs across the ten measurements were: $2.10\pm0.04\times10^{-3}$ mm^2/s for vial 1; $1.62\pm0.03\times10^{-3}$ mm^2/s for vial 2; $1.19\pm0.02\times10^{-3}$ mm^2/s for vial 3; $0.88\pm0.03\times10^{-3}$ mm^2/s for vial 4; and $0.61\pm0.03\times10^{-3}$ mm^2/s for

vial 5. The maximum coefficient of variation was below 5% indicating acceptable reproducibility of the proposed technique.

3.3.2 **In vivo Study**

Figure 3-7 shows a comparison between DW-ssEPI and DP-TSE from two representative patients. Geometric distortion (green arrows), chemical shift artifacts of the bone marrow and subcutaneous fat (red arrows and dotted lines), and susceptibility related signal pile up (blue arrows) were apparent in the DW-ssEPI images. DP-TSE provided considerably improved geometric accuracy and eliminated chemical shift from fat.

Quantitative landmark TRE measurement results are shown in Figure 3-8. The mean TRE was within 1.6mm for DP-TSE across all eight patients, whereas this value was as high as 12mm for DW-ssEPI. DW-ssEPI also had a high TRE standard deviation due to the fact that distortion is more apparent along the phase encoding direction and less severe in the readout direction so that the measurements had variations depending on the landmark locations.

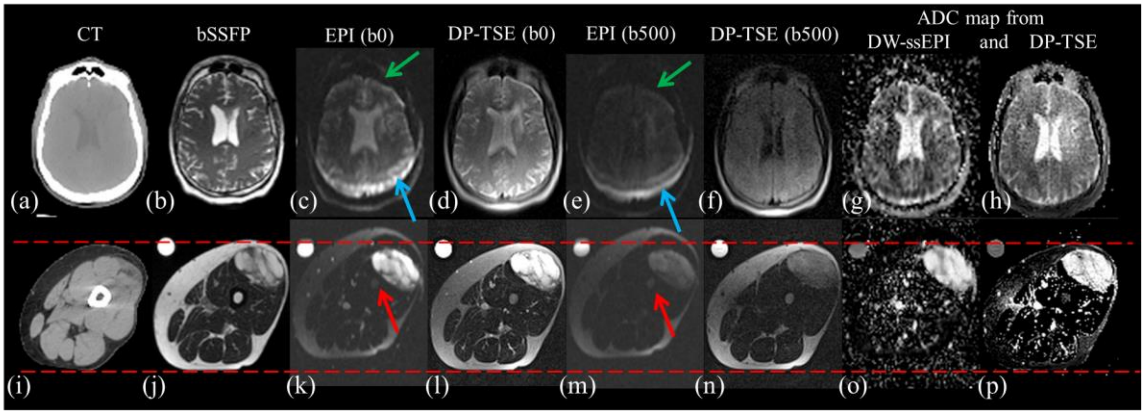

Figure 3-7 A comparison from two representative patients. (a)-(h) are the CT, clinical MR image from the standard ViewRay bSSFP sequence, DW-ssEPI (b=0), DP-TSE(b=0), DW-ssEPI (b=500), DP-TSE(b=500), and ADC maps calculated from DW-ssEPI and DP-TSE on one GBM patient, respectively. (i)-(p) are corresponding images for a sarcoma patient. The blue and green arrows point to signal pile-up and distortion artifacts in the DW-ssEPI image. The red arrows and red dotted lines indicate chemical shift artifacts of the bone marrow and subcutaneous fat. For each patient, the four diffusion-weighted raw images were displayed at the same window level.

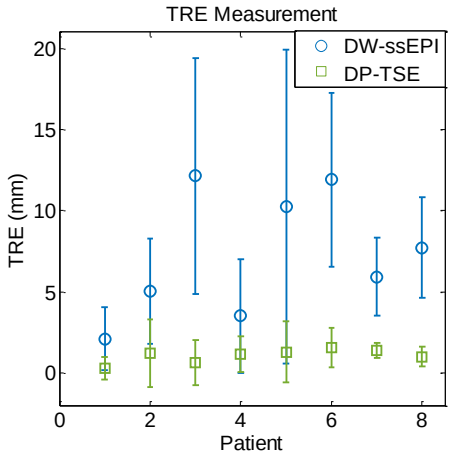

Figure 3-8 TRE measurements based on DW-ssEPI and DP-TSE.

Table 3-2 is a summary of the ADC measurements on all eight patients. In the two GBM patients, ADC values estimated from DP-TSE sequence for CSF and white matter were within literature ranges, which are 2.40×10^{-3} - 4.40×10^{-3} mm^2/s for CSF, and 0.60×10^{-3} -1.05×10^{-3} mm^2/s for white matter[82]. White matter ADCs from DW-ssEPI were around the lower limit of the literature value because the low T2 value of white matter aggravated the signal loss, which translated to ADC quantification inaccuracy. The tumor ADCs from DW-ssEPI were about 12% and 15% higher than those from DW-ssEPI. Among the six sarcoma patients, the ADCs difference were within 8%. For the reference tube that had a relatively low diffusivity, the two techniques provided matched measurements.

Table 3-2 In vivo ADC measurement results

ADC Measurement($\times 10^{-3}$ mm^2/s)			
		DW-ssEPI	DP-TSE
PT1, GBM	Tumor	2.02±0.48	2.29±0.65
	CSF	2.90±0.50	3.17±0.66
	White matter	0.56±0.13	0.75±0.33
PT2, GBM	Tumor	1.04±0.24	1.22±0.33
	CSF	2.88±0.43	3.03±0.48
	White matter	0.62±0.22	0.76±0.13
PT3-8, Sarcoma	Tumor	2.60±0.17	2.82±0.45
	Tumor	2.34±0.34	2.53±0.72
	Tumor	2.67±0.21	2.75±0.37
	Tumor	1.28±0.45	1.29±0.60
	Tumor	2.70±0.32	2.76±0.44
	Tumor	0.65±0.21	0.68±0.28
Vial		1.13±0.04	1.18±0.06

3.4 Discussion

To the best of our knowledge, the current work is the first study of distortion-free diffusion imaging sequence development on the MRgRT system. Our results show that the proposed DP-TSE technique provides good ADC quantification accuracy and substantially improved geometric fidelity when compared with the conventional single-shot EPI approach.

Distortion associated with phase accumulations during the EPI readout is a major problem for applying conventional EPI-based diffusion techniques to MRI-guided radiotherapy workflow, even at a low field strength of 0.35T. TSE, on the other hand, is insensitive to B_0 and off-resonance related artifacts due to the use of 180° refocusing pulses. As shown by the spatial integrity test, our proposed technique provided excellent geometric accuracy with a mean distortion of less than 0.6mm. The improved geometric fidelity is crucial for future adaptive radiotherapy techniques based on longitudinal onboard ADC measurements[72].

Diffusion imaging requires sufficient SNR in the k-space data to avoid ADC quantification errors, which occurs when the signal approaches the noise floor level after diffusion attenuation[80].

The problem is further aggravated when the diffusion data are acquired using a low field system like ours, which lowers the magnetic field strength to reduce electron returning effect[83]. In addition, our system has a maximum gradient amplitude of 18mT/m, and hence requires a long TE for moderate b-values, which consequently reduces SNR. Although TE can be reduced by using a higher readout bandwidth, our preliminary results showed that this led to an even lower SNR and inaccurate ADC quantification. As a consequence, the DW-ssEPI technique required a noise subtraction step before quantifying the ADC. For a low field strength of 0.35T, the low SNR using the EPI approach might limit the maximum achievable b-value. Our DP-TSE approach was able to achieve a shorter TE (115ms for DP-TSE v.s. 160ms DW-ssEPI) due to the use of a diffusion preparation module. It uses a segmented k-space sampling strategy to further boost the signal level compared to the single-shot EPI approach. Therefore, our DP-TSE signal does not approach the noise floor for a typical range of b-values of 100-1000 s/mm^2. The noise subtraction step is not strictly needed when using our DP-TSE approach. Besides the noise subtraction method implemented in this paper, there are several other methods that handle the noise in diffusion data, such as including the noise in the fitting[84].

Concomitant fields are generally of concern at low field strengths, and they could lead to errors in ADC quantification[85,86]. In this study, our sequences have a standard vendor-provided concomitant field correction module and the imaging slice was prescribed at the iso-center, which should reduce the concomitant field effects. However, we did not assess the effects of any remaining concomitant field on ADC accuracy.

It is well known that k-space segmented diffusion acquisitions are vulnerable to signal inconsistencies between different k-space segments because these segments are acquired at different time points and are subject to different signal phase accrual due to motion and system

imperfections. Several studies have been proposed to deal with motion-related artifacts in segmented TSE-based diffusion MRI. Pipe et al. implemented a diffusion-weighted PROPELLER TSE sequence, where phase cycling was used to mitigate the magnitude oscillation caused by the non-CPMG components[87]. The PROPELLER trajectory is robust to bulk motion and is self-navigated for phase correction, however, at the expense of about 50% increase of acquisition time compared with Cartesian sampling[87]. In a separate study, a self-navigated interleaved spiral diffusion MRI method with TSE readouts was proposed[88]. Although the 3D acquisition enabled higher resolution and SNR, only in-plane phase inconsistency was corrected and the CPMG condition was violated in its spiral TSE readouts. A more recent multi-shot Cartesian TSE diffusion study by Zhang et al.[89] used a preparation method proposed by Alsop[90] where an additional 90˚ pulse and dephasing gradients were used to eliminate the non-CPMG component at the expense of losing half of the signal. A multiplexed sensitivity encoding (MUSE)-based reconstruction[43] was used to estimate and correct the k-space phase inconsistencies. One disadvantage of the MUSE-based method it that the number of k-space segments is limited because each segment needs to have a sufficient number of k-space lines for the purpose of accurately estimating the phase variation maps.

In the current study, we chose the traditional Cartesian k-space sampling due to its widespread utility in clinical and research studies. A major difference between our TSE technique and most previously described techniques[87–89] is that we used a stand-alone diffusion preparation module before the imaging readout while most previous techniques used a diffusion-weighted strategy whereby diffusion encoding is a part of the imaging readout. The diffusion preparation strategy eliminates any violation of the CPMG condition in the TSE readout. In our TSE diffusion implementation, we changed the k-space view ordering to ensure that the central k-space lines

were acquired in a single shot, which appears to have mitigated ghosting artifacts in our patient cohort.

Our DP-TSE technique has limitations. Similar to other segmented diffusion MRI techniques, the proposed DP-TSE technique requires a longer scan time when compared with conventional single-shot EPI diffusion techniques. The trade-off between scan time and improved SNR and/or geometric accuracy needs to be balanced for different applications. The current work mainly focuses on phantom verification and in vivo feasibility demonstration. Larger patient cohort studies are warranted to demonstrate the value of this technique in clinical radiotherapy workflow.

3.5 Conclusion

A diffusion-prepared TSE-based diffusion technique with excellent geometric fidelity, accurate and highly reproducible ADC measurement was proposed and verified for longitudinal tumor response assessment using an MRI-guided radiotherapy system.

Chapter 4 2D Diffusion-prepared Magnitude-stabilized bSSFP imaging

In the previous chapter, the k-space magnitude inconsistency was mitigated by changing the k-space ordering such that the central k-space was covered by the same shot. This mitigates the ghosting artifact, however, the inconsistency still exists. This will consequently affect the ADC quantification if motion is strong. In this chapter, we aim to solve the k-space magnitude inconsistency in the multi-shot diffusion prepared approach. A magnitude stabilizer strategy was proposed to convert malignant magnitude inconsistency to benign phase inconsistency. The proposed diffusion-prepared magnitude-stabilized bSSFP sequence the on a diffusion phantom, and six healthy volunteers.

A version of this chapter has been published in the Magnetic Resonance in Medicine: Gao Y, Fei H, Zhou Z, Zhong X, Bi X, et al. Multishot diffusion-prepared bSSFP imaging with magnitude stabilizer. Magn Reson Med. 2019; 81(4): 2374-2384.

4.1 Introduction

Diffusion-weighted single-shot EPI MRI has been widely used for neurological and oncological applications, from imaging stroke to predicting tumor response to therapy [6,25,26]. The combination of EPI readout and single-shot acquisition mode enables fast acquisition speed and reduces motion-related image artifacts. However, DW-ssEPI often suffers from geometric distortions, susceptibility related artifacts, and limited spatial resolution. All these impede its acceptability in applications that require high spatial integrity such as tumor delineation and adaptive radiotherapy. Multi-shot diffusion-weighted imaging techniques have been proposed to

address these issues [13,87,91–93]. However, in multi-shot acquisitions, the use of strong diffusion gradients can cause significant variations in signal between imaging shots due to motion, physiological status changes, eddy currents, and other system imperfections. Direct combination of the inconsistent k-space segments obtained from the multi-shot acquisition could lead to destructive signal loss and ghosting artifacts in the reconstructed images. This k-space inconsistency is frequently observed and cannot be easily resolved by low-order motion compensation[39,94,95].

As introduced in Chapter 2.3, depending on how the diffusion-sensitized information is stored prior to the imaging readout, diffusion imaging sequences can be categorized as either diffusion-weighted or diffusion-prepared acquisitions. In a multi-shot diffusion-weighted imaging strategy, the imaging readout immediately follows the diffusion encoding module to acquire the diffusion sensitized transverse magnetization. In such a strategy, the k-space inconsistencies are evident as the k-space phase variations between different imaging shots. Various reconstruction methods have been proposed to mitigate the phase variation problem and these methods fall into two categories: navigator-based and navigator-free approaches. In the navigator-based phase correction approach, separate or built-in low-resolution images are acquired in each imaging shot to estimate the spatially varying phase, which is then used to correct the corresponding k-space segment[39–42]. Spiral, PROPELLER, and SNAILS are popular trajectories due to their inherent low-resolution navigator echoes[87,88,96–98]. Navigator-free phase corrections are increasingly investigated for improved scan efficiency. The multiplexed sensitivity-encoding (MUSE) technique and its extensions rely on parallel imaging reconstruction to estimate the phase from each individual imaging shot[43,99,100]. The estimated phase variations were then formulated in the reconstruction step to avoid image artifacts. The structured low-rank matrix completion

(MUSSELS) scheme was later proposed to recover artifact-free images without explicit phase estimation[101].

Another multi-shot diffusion imaging strategy is diffusion prepared imaging, which uses a -90˚ pulses to store the diffusion sensitized transverse magnetization information into the longitudinal direction and spoiler gradients to spoil the imperfect transverse components. This diffusion-prepared approach is compatible with many types of readout including turbo spin echo[102–104], and balanced steady-state free precession[37,38], enabling the flexibility of optimizing image acquisition type, SNR, contrast, and resolution. However, in diffusion-prepared approaches, the use of spoilers after the -90˚ tip-up pulse induces a shot-to-shot magnitude inconsistency, which consequently leads to signal void and ghosting artifacts, as well as inaccurate ADC quantification[105]. This magnitude variation mixed with the diffusion-induced signal attenuation is more challenging to correct because the conjugate signals are spoiled and lost. The aforementioned correction techniques are only applicable to address the phase inconsistency, but cannot be applied directly to correct the magnitude inconsistency. High-order motion compensation, advanced pulses, and twice refocused scheme were frequently adopted in DP diffusion MRI approaches to mitigate k-space inconsistencies originating from rigid motion, B1 inhomogeneity, and eddy currents[38,44]; however, the inconsistency problem originated from non-rigid motion, B0 inhomogeneity, and other imperfections still exists and could result in inferior image quality and inaccurate quantification if uncorrected. Orthogonal refocusing techniques, where images are acquired with 90°_{-x} tip-up pulse and 90°_{-y} tip-up pulse and the square root of the sum of squares of the two images, have been proposed to eliminate artifacts caused by eddy currents[37,106]. However, these methods suffer from doubled scan time and only works when the imperfections between the two shots are identical.

To address the magnitude inconsistency problem, we sought to incorporate a magnitude stabilizer strategy[90,102] to a DP-bSSFP technique. This magnitude stabilizer approach converts the magnitude problem back to the well-studied phase problem, such that the images can be reconstructed using previously described phase-correction methods for multi-shot diffusion imaging.

4.2 Methods

4.2.1 Imaging sequence

Figure 4-1 (a) shows the sequence diagram of the proposed diffusion-prepared magnitude-stabilized bSSFP sequence (DP-MS-bSSFP). The Stejskal-Tanner diffusion scheme is used in the diffusion prepared module[107]. The total duration of each diffusion gradient is 14ms (capable of creating a b-value of approximate 1200s/mm^2 at 80mT/m gradient amplitude). One 4π dephasing gradient is placed along the slice-selective direction before the 90°_{-y} pulse to evenly disperse the transversal magnetization. Consequently, half of the magnetization is flipped to the longitudinal direction by the 90°_{-y} pulse, while the remaining half is spoiled. The magnitude stabilizer (MS) gradients could be applied in any individual directions or combined directions. In this study. the slice-selective direction is chosen because the slice thickness is frequently larger than the in-plane resolution so that a lower gradient is sufficient to create the 4π dephasing. The total duration of the diffusion prepared module is 40ms. The diffusion module is followed by a bSSFP readout, which acquires the diffusion-sensitized signal after the steady-state catalyzation pulses. Rephasing and dephasing gradients are inserted before and after the echo during each readout to recall the echo and disperse the magnetization, respectively. During the readout, any T1 regrown signal will not be refocused by the rephasing gradient, so the signal is less sensitive to T1 contamination[108]. This

leads to an increased T2-weighting in the DP-MS-bSSFP images. The sequence without the magnitude stabilizers is the conventional diffusion-prepared bSSFP sequence (DP-bSSFP).

Figure 4-1 (b) is a simplified illustration of magnetizations for two shots with different phase angles φ before the tip-up pulse (blue and red arrow, respectively). Due to motion, system B0 and B1 imperfection, and eddy currents, the magnetization after the diffusion gradients has different phase angle with respect to the expected x-axis. Without MS, only the cos(φ) component is flipped to the z-direction, while the conjugate y component is spoiled, causing magnitude differences between the two segments if the phase angles are different. With MS gradient, the magnetization is evenly dispersed in the transversal plane (Figure 4-1 (c)), and then half of the magnetization is flipped to the z-direction regardless of phase angle before the MS gradient (Figure 4-1(d)). The remaining transversal magnetization is dephased by the spoiler gradients. The phase difference is stored and will be recalled after rewinding the magnitude stabilizer during each readout (Figure 4-1 (f)). Therefore, the magnitude stabilizer gradients convert the k-space magnitude inconsistency back to phase inconsistency in diffusion-prepared acquisition at the cost of half of the signal loss.

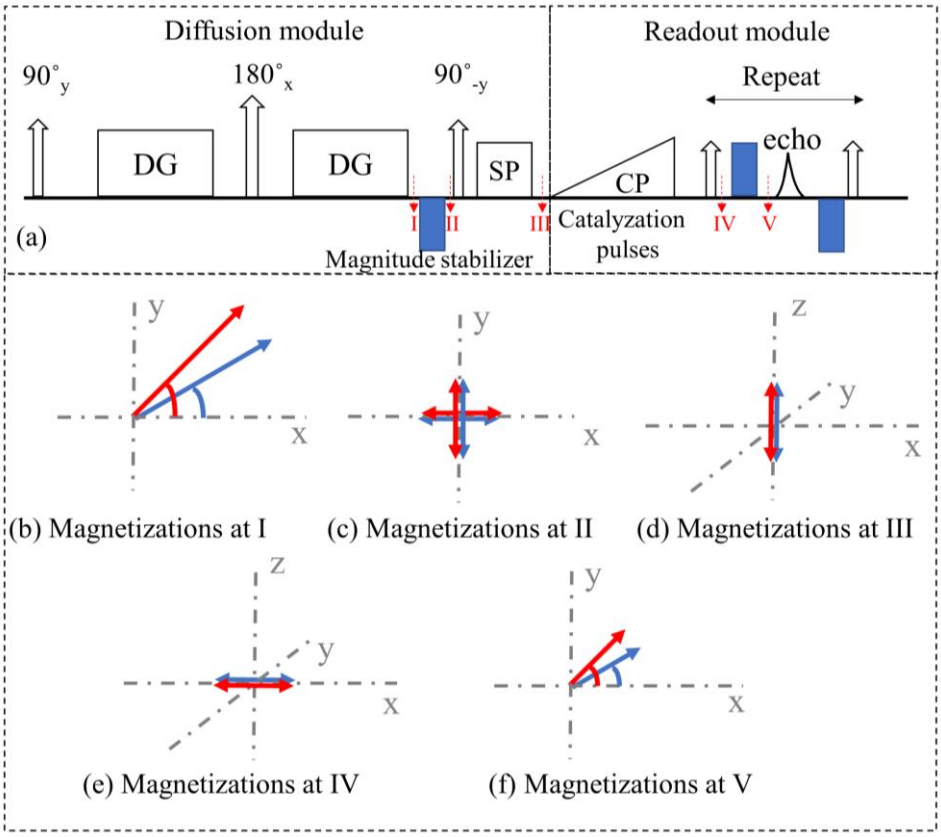

Figure 4-1 (a) Sequence diagram of one shot. Blue rectangles represent the magnitude stabilizers. (b) Illustration of magnetizations after the diffusion gradients for two different shots (blue and red arrows). (c-f) Magnetizations for the two shots at time II-V. (c) The magnetization is evenly dispersed in the transversal plane after the MS gradient. (d) Half of the magnetization is flipped to the z-direction regardless of the signal phase angles, while the remaining transversal magnetization is spoiled. (e) The magnetization after the RF pulse. For simplicity, a 90˚ readout RF pulse is used in the illustration. (f) After the rewinding MS, the magnetization is restored. DG, diffusion gradient; SP, spoiler; CP, catalyzation pulse

4.2.2 **Numerical Simulations**

Bloch simulation was performed to simulate the effect of adding MS gradients with respect to different phase errors. 2000 spins were simulated, with a uniform phase error ranging from -π to π added before the 90˚-y tip-up pulse. Each magnitude stabilizer created a 2π phase dispense among the 2000 spins in the transversal plane. The simulated T1 and T2 relaxation times were 800ms and 80ms, which are close to the white matter T1 and T2 value at 3T[109]. Flip angle of 60˚ was used for the bSSFP readout after ten linear steady-state catalyzation pulses[110]. TR was 2s and the duration

of the diffusion preparation module was 40ms. Readout echo spacing was 3.7ms for DP-bSSFP and 4.7ms for DP-MS-bSSFP. Those simulation parameters were the same as the parameters used in the actual imaging protocol. Five shot was simulated with the first shot discarded due to the signal level did not reach to a steady level. Simulation code can be found online (https://github.com/ygaogithub/DP-MS-Simulation).

4.2.3 **Phantom Study**

To evaluate the geometric reliability and ADC accuracy of the proposed DP-MS-bSSFP approach, a diffusion phantom (High Precision Device, Inc. Model 128. Boulder, USA) was scanned using a 3T scanner (MAGNETOM Prisma, Siemens Healthineers, Erlangen, Germany). The phantom consists of thirteen vials with varying concentrations of PVP from 0 to 50% w/w. Reference ADC values at zero degree were provided by the phantom vendor. In this study, the diffusion phantom was scanned at room temperature to cover a relatively large diffusivity range. TSE and DW-ssEPI were acquired to serve as the geometric reference and ADC reference, respectively; DP-bSSFP was acquired to evaluate the benefits of adding stabilizers. Scanning parameters for TSE, DW-ssEPI, DP-bSSFP, and DP-MS-bSSFP were: resolution=1.2x1.2x4mm^3 for all four sequences, TE=84/86/79/89ms, number of shot=12/1/4/4, and TR=3ms for all four sequences. Three b-value of 0, 500 and 800 s/mm^2 along the readout direction were acquired in the three diffusion sequences. Centric-out k-space ordering scheme was applied in DP-bSSFP and DP-MS-bSSFP acquisitions.

ADC maps of DW-ssEPI and DP-MS-bSSFP were obtained using the mono-exponential fitting $S(b) = S_0 e^{-b \times ADC}$, where S(b) is the signal intensity at b-value equals to b, S_0 is the signal intensity without diffusion, and ADC is the apparent diffusion coefficient. ADCs of all thirteen

vials were recorded and compared. Geometric accuracy was quantified using the target registration error. TSE image was used as the geometric reference, and b-value=0 s/mm^2 images of the DW-ssEPI and DP-MS-bSSFP were used as the test image to calculate the TRE respectively. The geometric fidelity of DP-bSSFP was not evaluated since it should have the same performance as DP-MS-bSSFP. Four landmarks (12 o'clock, 3 o'clock, 6 o'clock and 9 o'clock) were picked for the phantom and five selected vials (vial 1, 3, 5, 9 11. See Figure 4-4) respectively, resulting to a total of 24 landmarks on each image. Paired t-test was used to evaluate whether TREs from DW-ssEPI and DP-MS-bSSFP were statistically significantly different at the level of $p<0.01$.

4.2.4 **In vivo study**

Six healthy volunteers were recruited and imaged at a 3T scanner (MAGNETOM Prisma, Siemens Healthineers, Erlangen, Germany) under the Institutional Review Board approval. Written informed consent was obtained from each subject. Twenty-channel head and neck coils were used for all the studies.

To demonstrate the phase and magnitude variations caused by diffusion gradients, repeated single-shot measurement was performed in the volunteers. DW-ssEPI without electrocardiogram (ECG) triggering, DP-bSSFP without and with ECG triggering, and DP-MS-bSSFP without and with ECG triggering were acquired ten times with b-value=500s/mm^2 in a single-shot manner at the same resolution (1.8x1.8x4mm^3). TR=3s was used in the non-triggered acquisition. TR of three cardiac cycles and trigger delay of 0ms or 200ms was used in the triggered acquisitions. Magnitude and phase images were reconstructed for all acquisitions. The relative magnitude image variability was assessed by the coefficient of variation (CoV) map, which was calculated as σ/μ (σ: standard deviation of the pixel over the 10 repeated scans; μ mean pixel signal intensity of the repeated images).

To evaluate the multi-shot performance of the proposed method, DW-ssEPI and 4-shot DP-MS-bSSFP (no ECG triggering) with centric-out k-space ordering scheme were acquired using parameters listed in the Table 4-1. TSE was acquired to serve as the geometric reference. Two slice locations were prescribed and evaluated in this study: one contains the lateral ventricle where the pulsatile motion is strong, and the other at the base of the brain where susceptibility-related artifacts are often a problem. Three b-values (0, 500, 800s/mm^2) along the three orthogonal directions were acquired to evaluate the ADC accuracy of the proposed method. In addition, images at b-value of 800s/mm^2 along 20 diffusion weighting directions were acquired to test the capability of achieving high-quality diffusion tensor imaging. All multi-shot DP-MS-bSSFP images were reconstructed using the MUSE technique[43].

Table 4-1 Sequence parameter in the in vivo study.

	TSE	**DW-ssEPI**		**DP-MS-bSSFP**	
FOV (mm^2)	230x230	230x230		230x230	
Base resolution	192	192		192	
In-plane Resolution(mm^2)	1.2x1.2	1.2x1.2		1.2x1.2	
Slice thickness (mm)	4.0	4.0		4.0	
TRg/TE (ms)	4000/84	2000/86		2000/89	
Echo spacing (ms)	9.3	0.93		4.7	
Partial Fourier	-	4/5		-	
GRAPPA	-	2		-	
Echo train length	16	53		48	
Number of shot	12	1		4	
Bandwidth (Hz/px)	223	1240		789	
Average	1	4		1	
		DWI	DTI	DWI	DTI
B-value (s/mm^2)	-	0,500,800	0, 800	0,500,800	0,800
Diffusion Directions	-	3	20	3	20
Scan time (s)	0:54	1:04	2:56	1:02	2:50

To quantify the geometric distortion, ten landmarks were selected for each volunteer. Figure 4-2 shows examples of landmark locations. The locations of the landmarks on the DW-ssEPI and DP-MS-bSSFP images were compared with the coordinates on the TSE images for the TRE

calculation. To compare the ADC agreement between DW-ssEPI and DP-MS-bSSFP, mean diffusivity map was generated as the mean ADC over the three orthogonal directions. Four ROIs in the white matter region and two ROIs in the CSF region (if present) were drawn on the ADC map of each volunteer. Fractional anisotropy (FA) map and colored FA (cFA) map were generated with the OsiriX plug-in (Pixmeo, Geneva, Switzerland) using the non-diffusion-weighted image and the 20 independent diffusion-weighted images. In the cFA map, the major eigenvector directions are indicated by different colors (red: left/right, blue: superior/inferior, and green: anterior/posterior) and weighted by the FA value. To compare the FA value accuracy of the proposed DP-MS-bSSFP with DW-ssEPI, four ROIs were selected in the white matter region. Bland-Altman plot was used to evaluate the white matter ADC, CSF ADC, and white matter FA value agreement between DW-ssEPI and DP-MS-bSSFP.

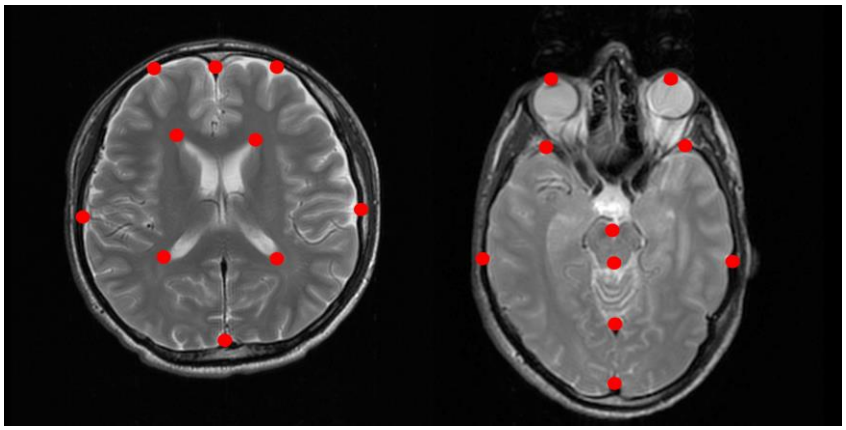

Figure 4-2 Examples of landmark locations used in the geometric accuracy quantification for slice containing the lateral ventricle and slice at the base of the brain. Ten landmarks were selected for each volunteer.

4.3 Result

4.3.1 Simulation

The signal magnitude of the 1st, 20th and 40th readout echo for the second imaging shot with and without magnitude stabilizers are shown in Figure 4-3. For all readout echoes, the signal magnitude exhibited a strong dependence on the phase error φ without the MS gradient. At certain phase error (around ±π/2 for the first echo), no diffusion sensitized transversal magnetization was restored to the longitudinal direction in the DP-bSSFP approach. Adding magnitude stabilizers provided stable signal magnitude regardless of the phase error. However, the signal magnitude was markedly reduced. For φ =0, the magnitude ratio between DP-bSSFP and DP-MS-bSSFP was 2.30, 2.88 and 3.73 for the 1st, 20th, and 40th echo, respectively.

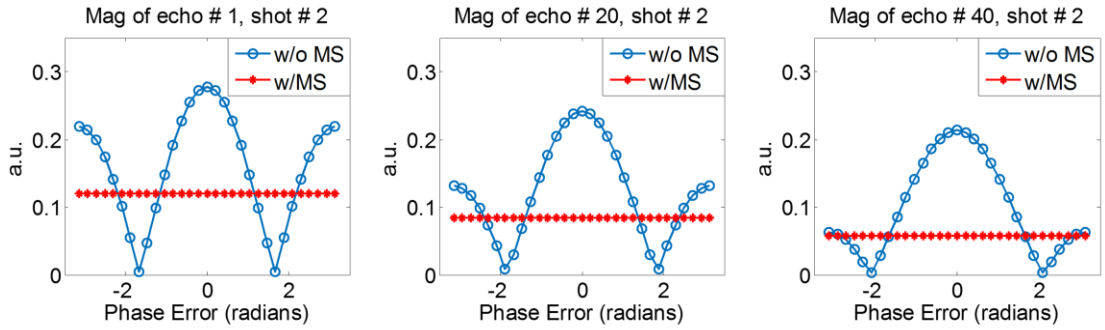

Figure 4-3 Simulation of signal magnitude with respect to phase error for the 1st, 20th and 40th readout echo without and with MS. The phase error does not affect the signal magnitude of the DP-MS-bSSFP approach; whereas it causes signal oscillation in the DP-bSSFP approach.

4.3.2 Phantom study

Phantom images at b-value=0 s/mm^2 and 800 s/mm^2 are shown in Figure 4-4. Compared to the reference TSE image, severe distortion and susceptibility-related artifacts were observed on DW-ssEPI images as indicated by the red arrows. The DP-bSSFP approach reduced the distortion;

however, susceptibility related artifacts at the tube/phantom boundary were observed on the b=800 s/mm^2 image as pointed by the blue arrows. The proposed DP-MS-bSSFP method provided satisfactory image quality with no observable artifacts.

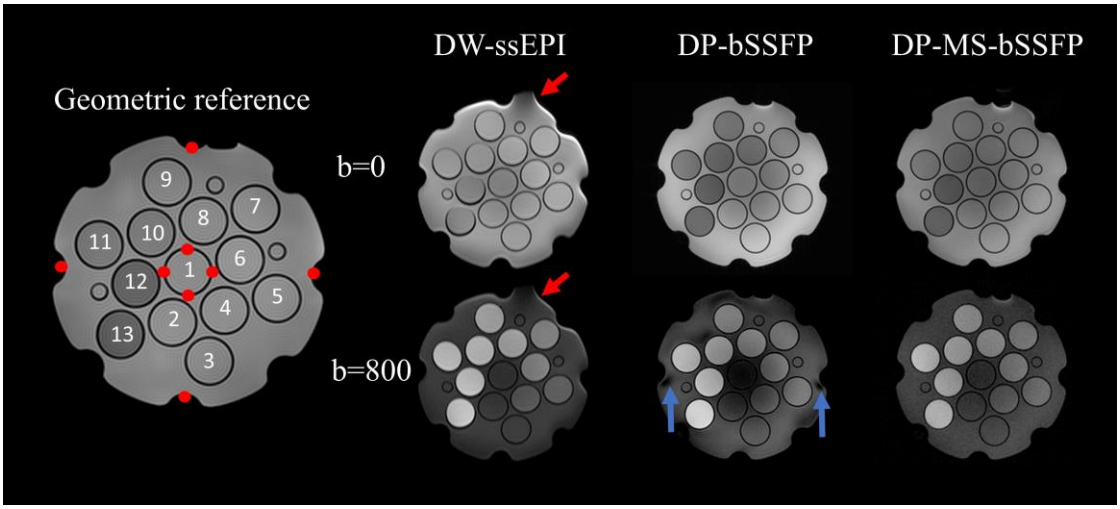

Figure 4-4 Phantom images. TSE was acquired as the geometric reference. Images at b-value = 0 and 800 s/mm^2 from DW-ssEPI, DP-bSSFP, and DP-MS-bSSFP were shown. The red dots on the TSE image indicate the four landmarks (12 o'clock, 3 o'clock, 6 o'clock and 9 o'clock) of the phantom and the four landmarks of vial 1 which were part of the 24 landmarks used for geometric accuracy quantification. Red arrows point to distortion artifact in DW-ssEPI, and blue arrows point to signal void artifact in DP-bSSFP.

Using the four landmarks of the phantom and four landmarks of the five selected vials (vial 1, 3, 5, 9 11 in Figure 4-4), the quantitative TRE measurement for DW-ssEPI was 2.31±1.02mm. The maximum landmark coordinate difference between DW-ssEPI and TSE was 4.29mm. TRE for DP-MS-bSSFP was 0.51±0.31mm, with the maximum coordinate difference of 1.20mm. TREs from DW-ssEPI were significantly larger than that of DP-MS-bSSFP ($p<0.01$).

ADC accuracy is summarized in Table 4-2. The proposed DP-MS-bSSFP had satisfactory ADC accuracy with less than 5% difference compared with ADCs from DW-ssEPI across all vials.

Table 4-2 Phantom ADC for the 13 vials

ADC ($x10^{-3}$mm/s)	DW-ssEPI	DP-MS-bSSFP
v1	2.19 ± 0.05	2.20 ± 0.14
v2	2.20 ± 0.04	2.23 ± 0.12
v3	2.21 ± 0.02	2.30 ± 0.10
v4	1.74 ± 0.02	1.75 ± 0.10
v5	1.73 ± 0.02	1.78 ± 0.07
v6	1.31 ± 0.02	1.31 ± 0.07
v7	1.32 ± 0.01	1.33 ± 0.06
v8	0.96 ± 0.02	0.94 ± 0.06
v8	0.96 ± 0.02	0.98 ± 0.05
v10	0.64 ± 0.01	0.62 ± 0.05
v11	0.64 ± 0.01	0.62 ± 0.04
v12	0.39 ± 0.01	0.37 ± 0.05
v13	0.38 ± 0.01	0.39 ± 0.04

4.3.3 **In vivo study**

Magnitude and phase images from five representative repetitions are shown in Figure 4-5 respectively. Results from DW-ssEPI, DP-bSSFP without ECG triggering, DP-bSSFP with ECG trigger delay of 0ms, and 200ms, DP-MS-bSSFP without ECG triggering, and DP-MS with ECG trigger delay of 0ms were shown. Corresponding CoV maps calculated from all ten repetitions were shown in Figure 4-6. For the standard DW-ssEPI sequence, although no strong variation was observed in magnitude, there existed obvious inconsistency among the phase images. Without MS or ECG triggering, this phase inconsistency resulted in magnitude inconsistency in the DP-bSSFP approach, leading to severe signal void artifacts in several measurements. Those magnitude variations were difficult to compensate for due to the irreversible loss of the remaining transverse magnetization immediately after the tip-up pulse of the DP module. An appropriate ECG trigger delay (0ms in this case) mitigated the signal magnitude variations in the DP-bSSFP approach; however, magnitude variations still existed and led to elevated CoV values compared with DW-ssEPI. An inappropriate trigger delay (200ms in this case) gave rise to signal void artifacts in all

repetitions. Adding the MS provided stable signal magnitude across all repetitions without ECG triggering. The CoV map from DP-MS-bSSFP was similar to DW-ssEPI. The benefit of adding ECG triggering was minimal for DP-MS-bSSFP.

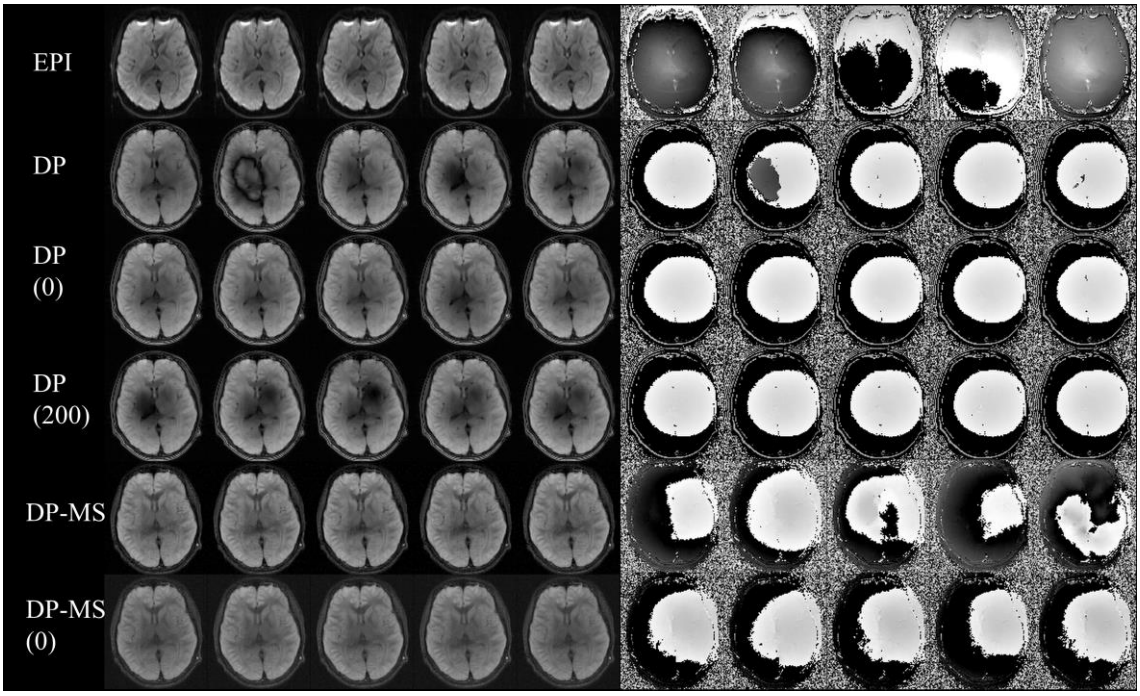

Figure 4-5 Magnitude and phase images of five repetitions from DW-ssEPI (abbreviated as EPI), DP-bSSFP (abbreviated as DP) without ECG triggering, DP with ECG trigger delay of 0ms, and 200ms, DP-MS-bSSFP (abbreviated as DP-MS) without ECG triggering, and DP-MS with ECG trigger delay of 0ms.

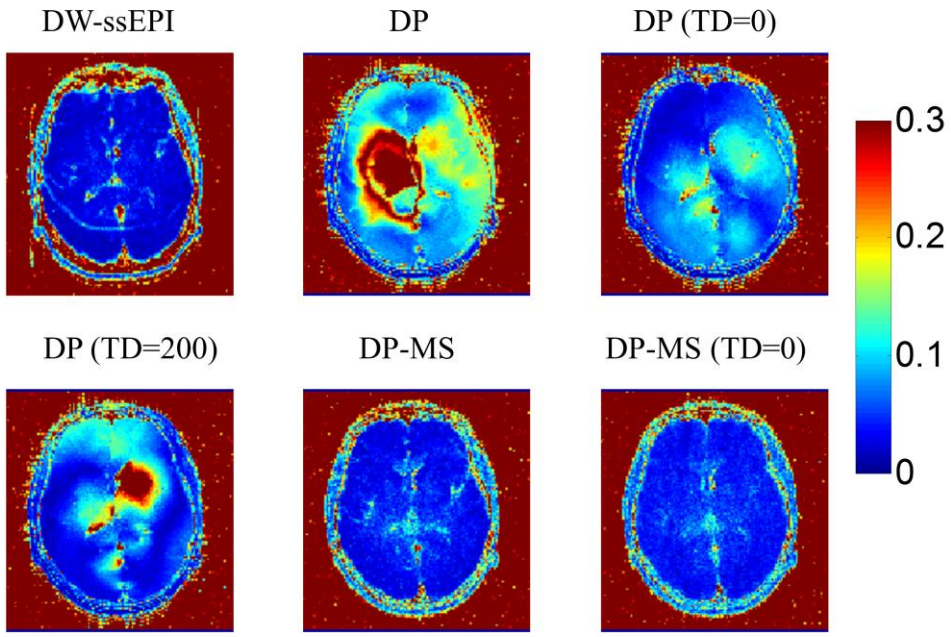

Figure 4-6 CoV maps calculated from the ten repetitions of DW-ssEPI, DP without ECG triggering, DP with ECG trigger delay of 0ms, and 200ms, DP-MS without ECG triggering, and DP-MS with ECG trigger delay of 0ms.

In vivo comparisons of DW-ssEPI and 4-shot DP-MS-bSSFP from 2 volunteers are shown in Figure 4-7. The TSE image, b=0 and 800 s/mm^2 images using the two diffusion sequences, and the corresponding cFA maps are displayed. As pointed by the red arrows, strong distortion and signal pile-up artifacts were found in DW-ssEPI images, particularly at the base of the brain where susceptibility problem is more severe. In comparison, after phase correction, the proposed DP-MS-bSSFP method provided distortion-free high-quality diffusion images and cFA map.

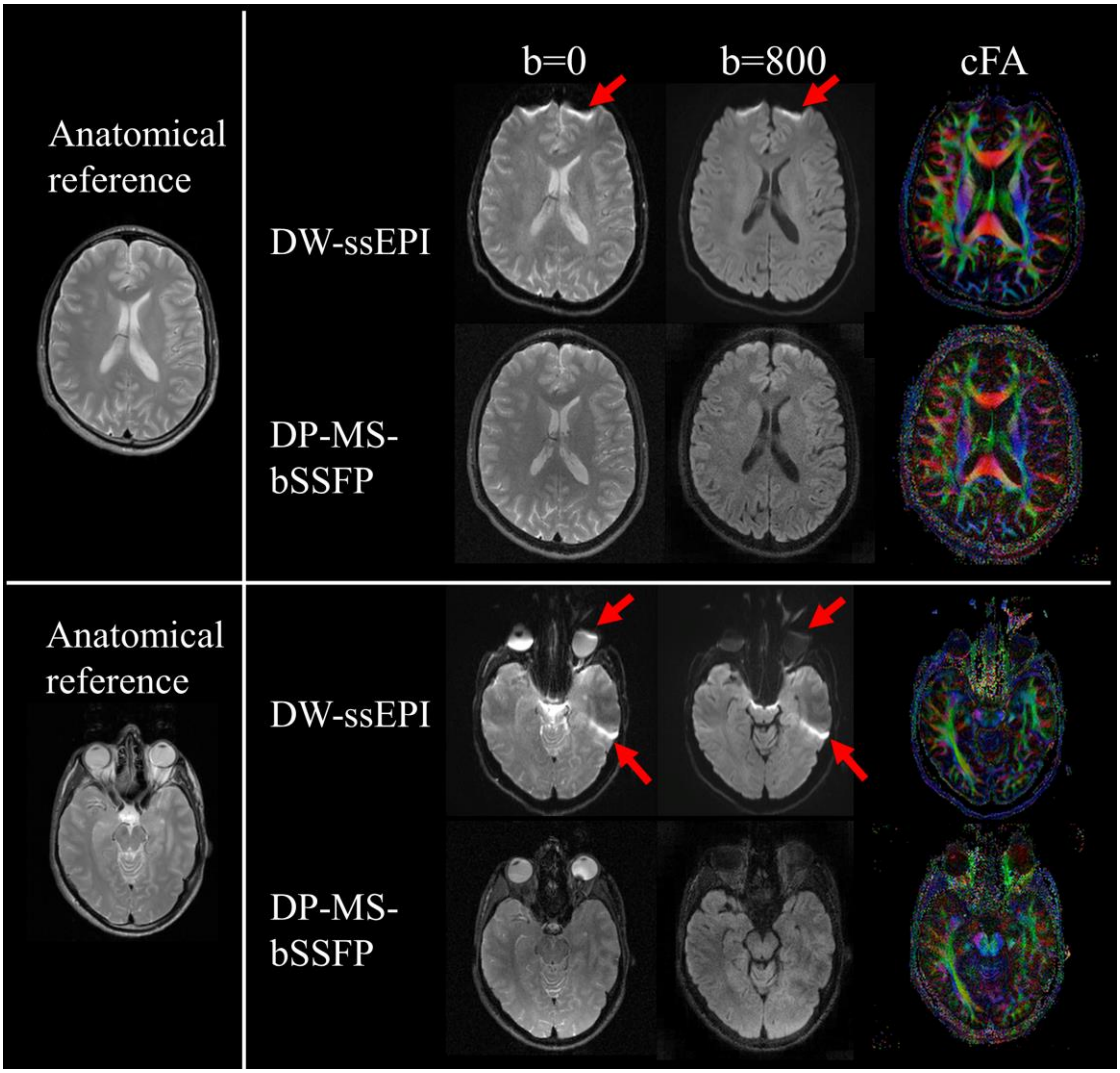

Figure 4-7 Multi-shot in vivo results from two volunteers. Images at b-value of 0 s/mm^2 and 800 s/mm^2, and cFA maps from DW-ssEPI and DP-MS-bSSFP were shown. Red arrows point to distortion and signal pile-up artifacts in DW-ssEPI.

Figure 4-8 shows the TRE measurements for the six volunteers. DP-MS-bSSFP provided substantially reduced distortion compared with DW-ssEPI. The mean TRE was 2.89 ± 2.63 mm for DW-ssEPI, and 0.59 ± 0.46 mm for DP-MS-bSSFP. The maximum landmark coordinate difference between DW-ssEPI and TSE was 10.80 mm, which was approximately 6.4 times higher than the DP-MS-bSSFP approach (1.69mm).

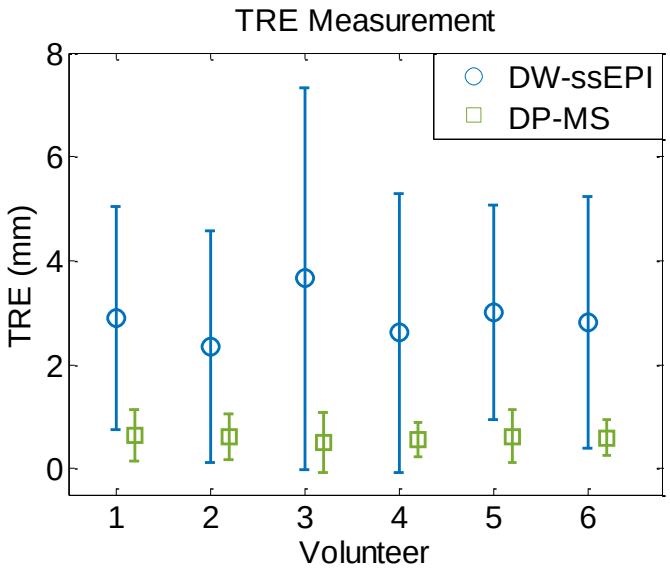

Figure 4-8 In vivo TRE measurement for the six volunteers. The error bar indicates the standard deviation. For volunteer 1,2,4,6, images were acquired at slice containing the lateral ventricle; for volunteers 3 and 5, images were acquired at the base of the brain.

Bland-Altman plots for selected ROIs on the white matter and the CSF were shown in Figure 4-9. Good agreements of white matter ADC, CSF ADC, and white matter FA value between DW-ssEPI and DP-MS-SSFP were observed. P-value from the paired t-test was 0.97, 0.35, and 0.68 for white matter ADC, CSF ADC, and white matter FA value, respectively.

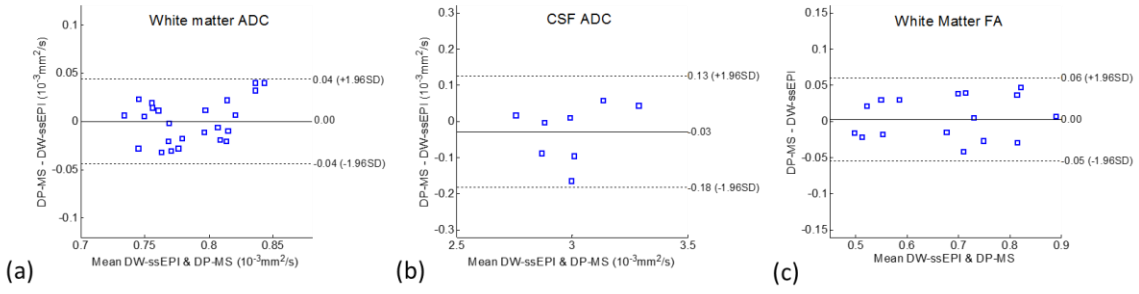

Figure 4-9 Bland-Altman plot for the white matter ADC (a), CSF ADC (b), and white matter FA (c). The p-value is from the paired t-test.

4.4 Discussion

The developed DP-MS-bSSFP sequence uses a magnitude stabilizer strategy to convert the magnitude problem in the diffusion-prepared approach back to the well-studied phase problem so that the shot-to-shot phase inconsistency can be solved by existing phase correction methods either with navigator echoes[39–42,88,96–98] or without navigator echoes [43,99–101]. Phantom and in vivo studies verified that the proposed DP-MS-bSSFP approach provided high-quality multi-shot images with no observable distortion or signal loss artifacts. High quality, distortion-free ADC map and cFA map were obtained using the proposed method for the brain. This study demonstrated that the proposed DP-MS-bSSFP approach is a promising candidate for applications that require high geometric fidelity such as tumor delineation and adaptive radiotherapy.

The magnitude stabilizer approach was inspired by phase insensitive preparation method proposed by Alsop, which uses a 90˚ pulse after the diffusion module and several dephasing and rephasing gradients throughout the sequence to eliminate the non-CPMG component in single-shot diffusion-weighted TSE imaging[90]. Similar strategies have been applied to achieve diffusion imaging with fast low-angle shot (FLASH) readout[108,111] and TSE readout[102–105,112]. However, in all multi-shot studies, only Zhang et al. and Cervantes et al. implemented phase correction to deal with the ghosting artifacts originated from multi-shot phase inconsistency[103,112], while the others did not have further correction methods, leaving the image vulnerable to shot-to-shot phase inconsistency. In this work, although the magnitude stabilizer strategy was similar to previous works, the motivation and purpose of adding those stabilizer gradients are different. In previous studies, the stabilizer gradients were added to only refocus the stimulated echo and avoid T1-dependent loss of diffusion information in the FLASH readout[108,111], or to eliminate the non-CPMG component in the TSE readout[102–105,112]. However, in our work, the stabilizer approach was

developed to transform signal magnitude inconsistencies into phase inconsistencies in DP diffusion MRI.

To our knowledge, this is the first work that implements the magnitude stabilizer approach for diffusion-prepared imaging with a multi-shot bSSFP readout. There are several potential benefits of using bSSFP as the readout. Compared with the EPI readout, bSSFP provides substantially reduced distortion, which is essential for oncological application where high spatial integrity is required. Compared with TSE readout, the bSSFP readout has less SAR concern and shorter echo spacing hence less T2 blurring. Compared with FLASH readout, improved contrast and higher SNR, which is critical for the MUSE reconstruction, are obtained using bSSFP readout as shown in Figure 4-10. Therefore, the bSSFP readout is a promising strategy for achieving 3D high-quality distortion-free diffusion imaging.

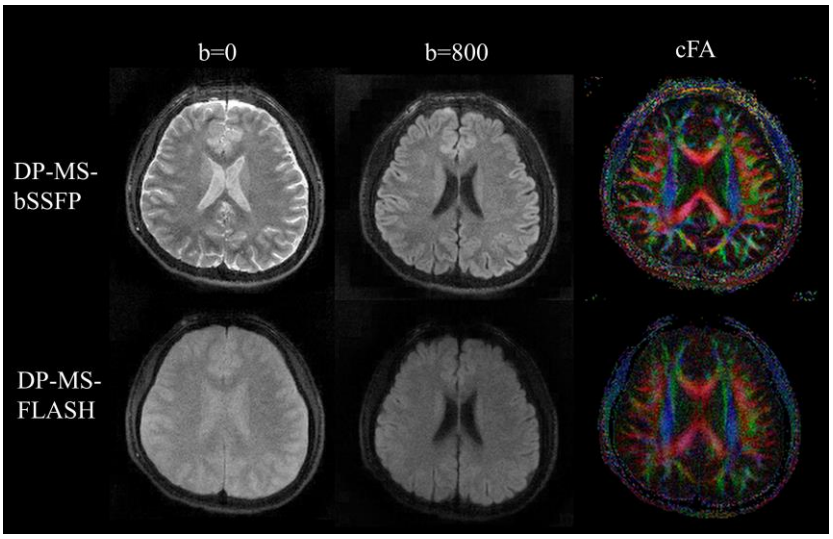

Figure 4-10 Images from b=0 and b=800 s/mm^2, and cFA maps for DP-MS-bSSFP and DP-MS-FLASH. Same window and level were used for each comparison. DP-MS-bSSFP provided higher SNR and improved contrast than DP-MS-FLASH. The quality of cFA map from DP-MS-bSSFP was better

In this work, MUSE reconstruction was implemented to reconstruct the multi-shot brain images. Desirable image quality and ADC and FA accuracy indicate that the MUSE method is

sufficient to handle the 4-shot case. Several studies have demonstrated that for a large number of shots, the MUSE technique might fail due to the difficulty for accurate phase estimation from highly under-sampled data. To further improve the image quality, particularly if a higher number of imaging shots is applied, other reconstruction methods, such as POCSMUSE, self-feeding MUSE, and MUSSEL, may be adopted[99–101]. On the other hand, navigators could be acquired for direct phase correction. For the DP-MS-bSSFP approach, the navigator could be added for free during the steady-state catalyzation pulses, resulting in no scan efficiency penalty. Further investigation is needed to evaluate whether the signal oscillation during the ramp-up pulse will affect the effectiveness of navigator correction.

Although a good trigger delay could reduce signal magnitude variation in the repeated single-shot DP-bSSFP study, its performance is still inferior to DP-MS-bSSFP. As shown in Figure 4-5, selecting an appropriate ECG trigger delay is crucial in mitigating magnitude variation in the DP-bSSFP approach; otherwise, the image artifacts could be magnified. DP-MS-bSSFP, on the other hand, had stable performance without ECG triggering. Therefore, no addition ECG setup is required for the scan.

A disadvantage of the proposed DP-MS-bSSFP approach is the signal loss associated with the added magnitude stabilizers. Theoretically, there is a 50% signal reduction by introducing the MS gradients. The added MS gradients also induce a longer TE and addition dephasing, which could cause a further signal reduction. As shown in Figure 4-3, the magnitude ratio between DP-bSSFP and DP-MS-bSSFP could be as high as 4 for the last few echoes. Therefore, long echo train length is not suggested and a centric-out reordering scheme should be used to ensure the k-space center region is acquired while the diffusion-related signal level is high. For the protocol used in this manuscript, SNR of DP-bSSFP was approximately 2.5 times higher than SNR of DP-MS-bSSFP

(see Figure 4-11). Based on our experience, b-value of 800 s/mm^2 at 1.2x1.2x4mm^3 resolution can be achieved in a single average at 3 Tesla without compromise of image quality or ADC and FA accuracy. However, the signal level might drop to a critical level at higher resolution, lower field strength, or with higher b-value applied. Scan averaging may be used to increase the SNR; whereas the acquisition time is proportionally increased with the number of averages, limiting the application of the proposed approach. Extending the DP-MS-bSSFP approach to the 3D acquisition would naturally increase the SNR, therefore compensating for the signal loss associated with the magnitude stabilizers. Future studies are warranted to optimize the design of k-space trajectory for effective k-space sampling and navigator acquisition so that the total acquisition time is practically reasonable.

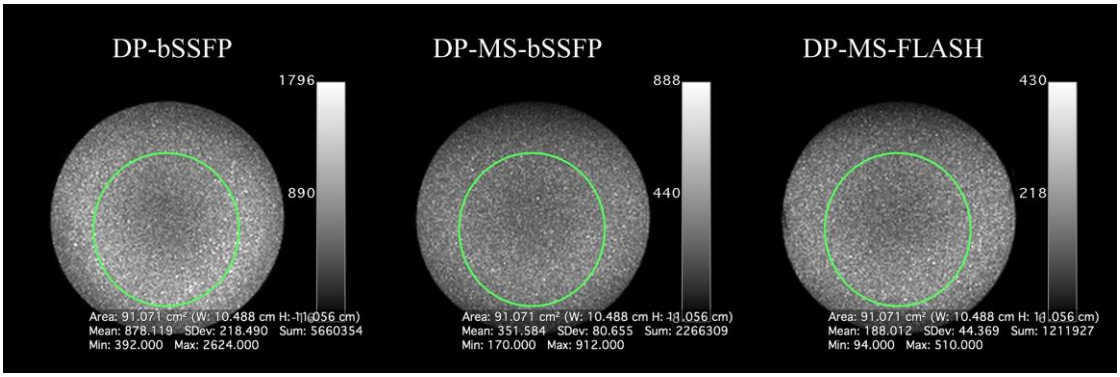

Figure 4-11 SNR comparison of DP-bSSFP, DP-MS-bSSFP, and DP-MS-FLASH on a water ball phantom. 15 repetitive measurements were performed to calculate the SNR map. SNR of DP-bSSFP was approximately 2.5 times higher than DP-MS-bSSFP, and SNR of DP-MS-bSSFP was about 1.9 times higher than DP-MS-FLASH

4.5 **Conclusion**

A magnitude stabilizer strategy was used in the diffusion-prepared bSSFP imaging to convert the magnitude inconsistency into phase inconsistency. The proposed DP-MS-bSSFP approach provided high-quality distortion-free multi-shot diffusion images, and is a promising candidate for applications that require high geometric accuracy.

Chapter 5 3D Diffusion-prepared Magnitude-stabilized bSSFP imaging

The diffusion-prepared magnitude stabilizer bSSFP sequence proposed in Chapter 4 was designed for 2D imaging. The method does not need a navigator for phase estimation. However, for radiotherapy purpose, a larger coverage with thin slice thickness is usually required. 3D imaging has the advantages of improved SNR and RF profile for high-resolution imaging. Therefore, in this chapter, the DP-MS-bSSFP sequence was modified to 3D to achieve high-resolution distortion-free diffusion imaging. Navigator echoes was acquired for phase estimation. The proposed 3D sequence was evaluated on a diffusion phantom and five healthy volunteers.

A version of this chapter is ready to submit: Gao Y, Zhou Z, Fei H, Zhong X, Yang Y, Hu P. Three-dimensional multi-slab distortion-free diffusion-prepared magnitude-stabilized bSSFP imaging at 1.5 Tesla.

5.1 Introduction

MRI has been playing an emerging role in radiation therapy. MR-based or MR-only simulation has become the standard practice in many institutes to take advantage of the superb and versatile soft-tissue contrast for tumor and organ at risk delineation[113]. According to the American Association of Physicists in Medicine (AAPM) Task Group 101 (TG101), MRI should be used as the gold standard for visualization in brain neoplasms stereotactic body radiation therapy (SBRT) treatment, and is increasingly used in prostate, spinal tumors, chest, and solid abdominal tumors SBRT applications[114]. Two MRI-guided radiotherapy systems are now commercially available to explore gated-treatment and online adaptive radiotherapy for improved treatment efficacy[9,10]. One

multi-institutional study showed that inoperable pancreatic cancer patients treated with dose-escalated adaptive MRI-guided radiotherapy had improved overall survival than those with standard treatment[115].

Diffusion-weighted imaging provides a unique contrast that reflects water diffusivity, and has been shown exceptional value in tumor delineation and response assessment[116]. It provides tumor cellularity information for gliomas characterization, which is not available with other conventional MR images[117]. DWI has become a standard protocol in prostate cancer patient imaging, and has improved tumor detectability in the transition zone and the peripheral zone than the conventional T2w imaging[118–120]. As successful treatment will lead to apoptosis and necrosis of the tumor cell, DWI is able to detect those cellularity changes well before anatomical changes for response assessment. Studies have shown the value of using pre-treatment DWI or inter-treatment DWI as an early treatment response biomarker for many disease sites including brain, head and neck, liver, pancreas, and prostate[66,121–125]. This enables individualized response-based treatment adaptation for improved tumor control.

A major obstacle that prevents applying diffusion imaging for treatment planning and online adaptive radiotherapy is the strong spatial distortion, which is problematic for accurate target delineation. Due to low readout bandwidth in the phase encoding direction, the conventional DW-ssEPI sequence is very sensitive to off-resonance effects such as field inhomogeneity, eddy current, and chemical shift. Those field inhomogeneities can lead to significant distortion in the phase encoding direction. Multi-shot diffusion techniques have been extensively investigated to address the distortion problems associated with the DW-ssEPI approach[13,87,91,93]. Instead on fill in the whole k-space in a single readout echo train, the k-space is segmented in the readout or the phase encoding direction, and each segment/shot is covered with a separate echo train to reduce the

effective bandwidth in the phase encoding direction. This also reduces T2*-related blurring and has the potential of achieving higher resolution.

However, most of those techniques still rely on the EPI readout, so the distortion reduction is proportional to the number of imaging shot. To achieve desirable geometric accuracy, a large shot number, consequently long scan time, is required. Therefore, it is beneficial to explore other readout modules, such as turbo spin echo and balanced steady-state precession, to achieve higher spatial integrity in a reasonable time. In addition, most of the diffusion research focuses on 2D imaging, which has limited spatial resolution in the slice direction due to low SNR of reduced voxel size. High spatial resolution is crucial to overcome partial volume effect and provide more accurate target delineation, especially for SBRT and stereotactic radiosurgery (SRS) treatments. 3D diffusion imaging has the potential of achieving higher resolution benefited from the intrinsic high SNR of 3D imaging. The bSSFP readout, in particular, is a promising 3D readout candidate due to its minimal distortion compared to EPI readout and less SAR concern compared with TSE readout.

Therefore, the goal of this work is to develop a 3D diffusion sequence to achieve high-resolution distortion-free diffusion imaging for treatment planning and treatment response assessment of the brain.

5.2 Methods

5.2.1 Sequence design

A previously proposed diffusion-prepared magnitude-stabilized bSSFP (DP-MS-bSSFP, abbreviated as DP-MS) sequence[126] was extended to 3D. The sequence diagram of one example shot was shown in Figure 5-1. In each shot, the sequence starts with the diffusion module using the traditional Stejskal-Tanner pulsed gradient spin echo technique[107], then followed by a bSSFP

readout module. Stabilizer gradients were added before the 90˚ tip-up pulse in the diffusion module, and before and after the echo in each readout session to convert the shot-to-shot k-space magnitude inconsistency to phase inconsistency[126]. A 2D navigator was acquired during the built-in linear catalyzation pulses for each shot to estimate and correct the phase inconsistency between different shots. The catalyzation pulses were mainly used to reduce signal oscillation and to achieve a steady-state signal intensity during the image acquisition. In this study, 16 navigator echoes were acquired during this ramp-up stage. This navigator was only phase-encoded to cover the central ky lines. It assumes that the through-slab phase variation is negligible, and has been shown to be valid for slab thickness less than 30mm in the brain[127,128].

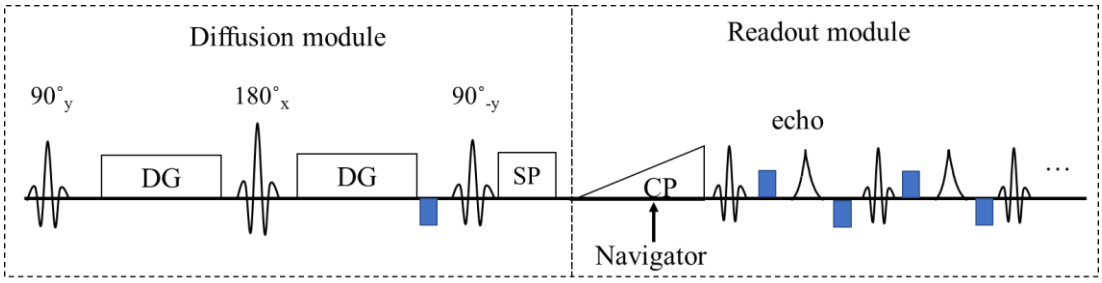

Figure 5-1 Sequence diagram of one example shot. Blue rectangles represent the magnitude stabilizers. In this study, the navigator was acquired in each shot during the 16 linear catalyzation pulses.

5.2.2 **Image reconstruction**

The 3D diffusion data was then phase corrected slice-by-slice and the phase correction was performed separately for each individual imaging shot for a given slice in the z direction[129]. A 1D Fourier transform was first applied along the Kz direction. All the subsequent image reconstruction and signal processing were performed separately for each of the slices in the Kx-Ky-z space after the Fourier transformation in the Kz direction. For each slice, the k-space data from each individual imaging shot was used to reconstruct one image. Under an ideal scenario where no motion occurs between shots, the reconstructed image series from multiple shots would be identical with each other and the rank of the Casorati matrix, in which each column consists of the vectorized image

pixel intensities from each individual reconstructed image, would be 1. However, motion exists between shots in practical situation, which introduces additional shot-to-shot spatially varying phase variations and causes the corresponding Casorati matrix to have a rank of greater than 1. We assume that by correcting for the phase variations, the low-rank property of the Casorati matrix can be exploited to recover an artifact-free diffusion image from acquired multi-shot data. Specifically, for each slice, define y_{ij} as the acquired k-space data matrix (size: $n_x * n_y$) from the ith shot (within a total of L shots, i.e. $i = 1,2, \dots, L$) and the jth coil (within a total of n_c coils, i.e. $i = 1,2, \dots, n_c$), and define x_i as the coil-combined image matrix (size: $n_x * n_y$) independently reconstructed from the ith shot, x as the matrix (size: $n_x * n_y * L$) concatenating all images from different shots, F as the Fourier transform operator, D_i as the under-sampling operator for i^{th} shot, P_{ij} as the phase correction matrix (size: $n_x * n_y$) which is the phase of the navigator image obtained using the navigator echoes acquired during the ramp-up stage, and S_j as the jth coil sensitivity map. Considering imperfections in phase compensation, a locally low-rank constrained optimization problem can be formulated as follows to solve for the underlying artifact-free image:

$$min_x \sum_{j=1}^{n_c} \sum_{i=1}^{L} || D_i F S_j P_{ij} x_i - y_{ij} ||_2^2 + \gamma \sum_{b \in \Omega} || C_b x ||_*$$

where Ω is a set of small image blocks (size: $b_x * b_y * L$) partitioned from x, C_b is the operator that takes image block b from the set Ω and forms its Casorati matrix (size: $b_x b_y * L$), $|| \cdot ||_*$ is the nuclear norm, and γ is the regularization parameter. Several studies have shown the improved performance of locally low-rank constraint over globally low-rank constraint[130,131]. The optimization is performed for all the slices, and the final diffusion image is calculated as the sum of square combination of x along the shot dimension. The reconstruction was performed using

the Berkeley Advanced Reconstruction Toolbox (BART)[132], which solves the low-rank optimization using the alternating direction method of multipliers (ADMM).

5.2.3 **Phantom study**

To evaluate and compare the geometric fidelity and ADC accuracy of the proposed DP-MS and the conventional DW-ssEPI sequence, a diffusion phantom (High Precision Device, Inc. Model 128. Boulder, USA) was scanned on a 1.5T scanner (Avanto, Siemens Healthineers, Erlangen, Germany). As diffusivity is temperature-dependent, ice temperature study and room temperature studies were conducted to analyze the ADC accuracy. In the ice temperature study, the phantom was filled in with crushed ice one day before the scan and stored in the refrigerator overnight. Additional ice was added to displace the melted water on the morning of the scan day. In the room temperature study, the phantom was put in the scanner room one day before the scan to allow the phantom reaching the same temperature as the scanner room. To meet the high-resolution requirement for SBRT and SRS treatment planning, DW-ssEPI and DP-MS were acquired at $1.5mm^3$ isotropic resolution. Scan parameters for the two sequences were: TR=6100/2000ms, TE=86/120ms, field of view=$240x240mm^2$, resolution=$1.5mm^3$, bandwidth= 1562/780 Hz/px, echo spacing=0.75/5.6ms, average= 12/1, 32 slices (25% oversampling in DP-MS), b-value= 0 and 800 s/mm^2, scan time=2min46s/9min36s. In DP-MS, each kz plane was covered with four shots. Turbo spin-echo sequence with matched resolution and coverage was acquired as the geometric reference.

Target registration error was used as a surrogate of the geometry accuracy. To calculate TRE, landmarks were selected on the TSE image, which is called the reference image, and the DW-ssEPI/DP-MS image, which is called the test image. TRE is defined as the root mean square of the Euclidean distance between corresponding landmark locations on the reference and the test images.

Images acquired at room temperature were used for the calculation. A total of 8 landmarks were selected on each slice of images, including the topmost, bottommost, leftmost and rightmost points of the phantom and the central vials. The paired t-test was used to compare the TRE from DW-ssEPI and DP-MS at significant level of $p<0.01$.

To quantify the ADC accuracy, ADC values for the thirteen diffusion vials were recorded across all slices and then averaged. For the ice temperature study, ADC values from DP-MS were compared with ADCs from DW-ssEPI as well and the reference values at 0˚C provided by the phantom vendor. For the room temperature study, ADCs from the two diffusion sequences were compared.

5.2.4 **In vivo study**

Five healthy volunteers were recruited to investigate the geometric accuracy and ADC accuracy of the DP-MS sequence for treatment planning of the brain at 1.5T. The study was approved by the Institutional Review Board approval. A written consent form was obtained prior to each study.

In the first volunteer, one shot per kz plane, two shots per kz plane, and four shots per kz plane were acquired to compared and identify the best protocol for in vivo imaging. ECG triggered acquisition was adopted to minimize the influence of cardiac pulsatile motion. The sequence parameters include: TR=2 cardiac cycles, TE=120ms, FOV=240x240mm^2, resolution=1.5mm^3, 16 slices with 25% over-sampling, bandwidth=780 Hz/px, average=1, and b-value of 0 and 800 s/mm^2 in three orthogonal directions. The scan time for the one-shot, two-shot, and four-shot protocols were approximately 2min40s, 5min20s, and 10min40s, respectively.

Based on the results from the first volunteer, the four shots per kz plane protocol was selected and applied in the remaining four volunteers. A total of 48 slices in three slabs were acquired to

evaluate and compare geometric fidelity and ADC accuracy of DP-MS and DW-ssEPI. Eight to ten landmarks were picked in each slice to calculate the TRE. These landmarks include the topmost, bottommost, leftmost and rightmost points of the brain, and some easily identifiable structures such as the eyeball, the lateral ventricle, etc. Paired t-test was carried out to evaluate whether TREs form the two diffusion techniques were significantly different at $p<0.01$. To investigate the ADC accuracy, a region of interest was drawn on the white matter for all slices. Additional ROIs were drawn on the cerebrum and CSF region for slices that contain the cerebrum and the ventricle. The Bland-Altman analysis was performed to evaluate the ADC agreement between DW-ssEPI and DP-MS for white matter, cerebrum, and CSF.

5.3 Result

5.3.1 Phantom Experiments

Phantom image comparison is shown in Figure 5-2. As indicated by the blue dotted lines, the phantom shell was strongly distorted using the DW-ssEPI approach. There was susceptibility related artifact as pointed by the blue arrow. In addition, several diffusion vials were elongated and were no longer in round shape. In comparison, no visible distortion or artifact was present in DP-MS image.

Compared with TSE images, the maximum landmark displacement was 8.34mm for DW-ssEPI, and 3.13mm for DP-MS. TREs from DP-MS (0.70±0.50 mm) were significantly lower than TREs from DW-ssEPI (2.09±1.75mm).

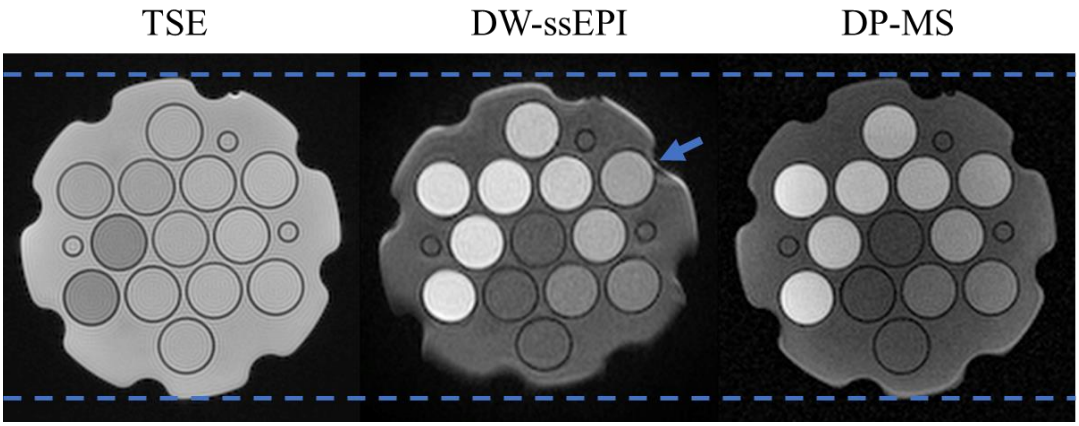

Figure 5-2 Phantom image comparison. Images from left to right are: TSE, DW-ssEPI (b=800 s/mm^2), and DP-MS (b=800 s/mm^2). The blue dotted lines and arrow point to distortion and artifact in the DW-ssEPI image.

Figure 5-3 shows the ice temperature and room temperature ADC results. In the ice temperature study, the temperature of the phantom was 1.4˚C before the imaging and remained at 1.4 ˚C after the imaging. Due to higher temperature than 0˚C, ADCs from DW-ssEPI and DP-MS were higher than the reference value. For diffusivity greater than 0.4x10^{-3}mm/s (the first nine diffusivity vials), the percent difference between DW-ssEPI and DP-MS was less than 4%. The last four diffusivity measurements were less accurate and had a high standard deviation due to the lack of high b value to characterize low diffusivity. In the room temperature study, the phantom temperature was 22.8˚C before the imaging and increased to 23.2˚C after the imaging due to RF heating. Overall, there was a good agreement of ADC measured from the two diffusion sequences. The percent difference was less than 7%.

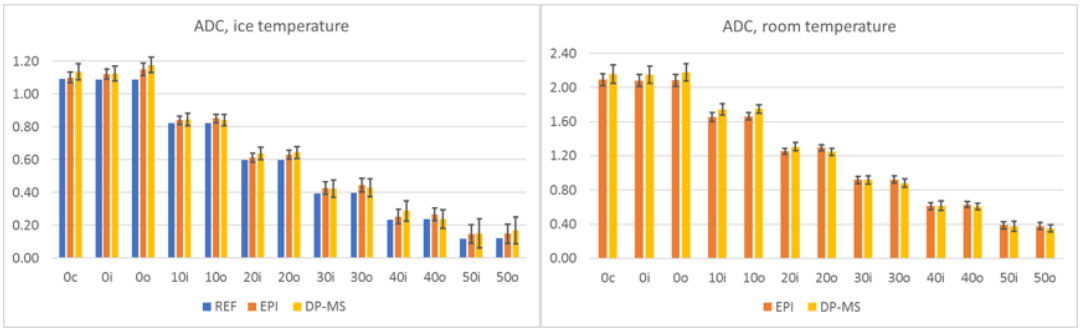

Figure 5-3 ADC accuracy at ice temperature and room temperature. The x-axis shows the labels of the thirteen vials. The number represents the percent concentration of PVP solutions (higher PVP, lower diffusivity), and c (center), i (inner), o (outer) represent the relative location of the vial. The y-axis is the reference and measured ADC values in $10^{-3}mm^2/s$. The error bar represents the standard deviation.

5.3.2 **In vivo Experiments**

The comparison between one shot, two shots and four shots per kz plane was shown in Figure 5-4. It is apparent that the one-shot protocol had relative low SNR. The image was not acceptable for clinical use. Applying two shots per kz plane improved the SNR, whereas the long echo train and associated signal decay induced blurring. As pointed by the triangles, the fine details of grey matter structure were blurred. The four shots per kz plane approach provided high-quality images with decent SNR. The grey matter structure can be clearly visualized. Applying more shots per each kz plane has the potential of further improving the image quality. However, the scan time would be too long to be adopted clinically. Therefore, the four shots protocol was selected and investigated in the following in vivo studies.

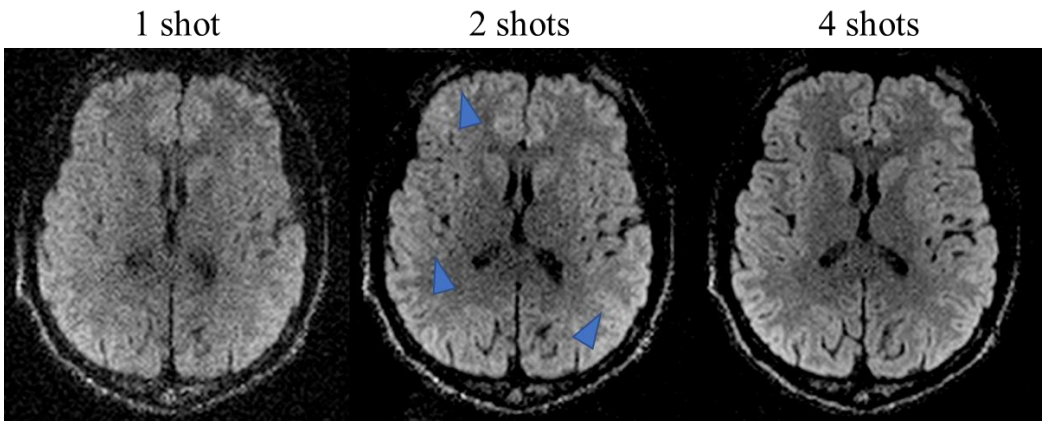

Figure 5-4 Comparison of b=800 s/mm^2 images acquired using one shot per kz, two shots per kz, and 4 shots per kz. The blue triangles points to the blurring of the grey matter structure.

Figure 5-5 shows the comparison of diffusion images between DW-ssEPI and DP-MS. As pointed by the blue triangles, strong distortion and signal pile-up artifacts around the tissue air boundary were apparent on the DW-ssEPI images. The distortion is particularly strong for the eyeballs and at the base of the brain due to increased susceptibility. This could induce erroneous contour and consequently miscalculated dose to the patient. In comparison, the proposed DP-MS approach provided high-quality diffusion images with no observable distortion.

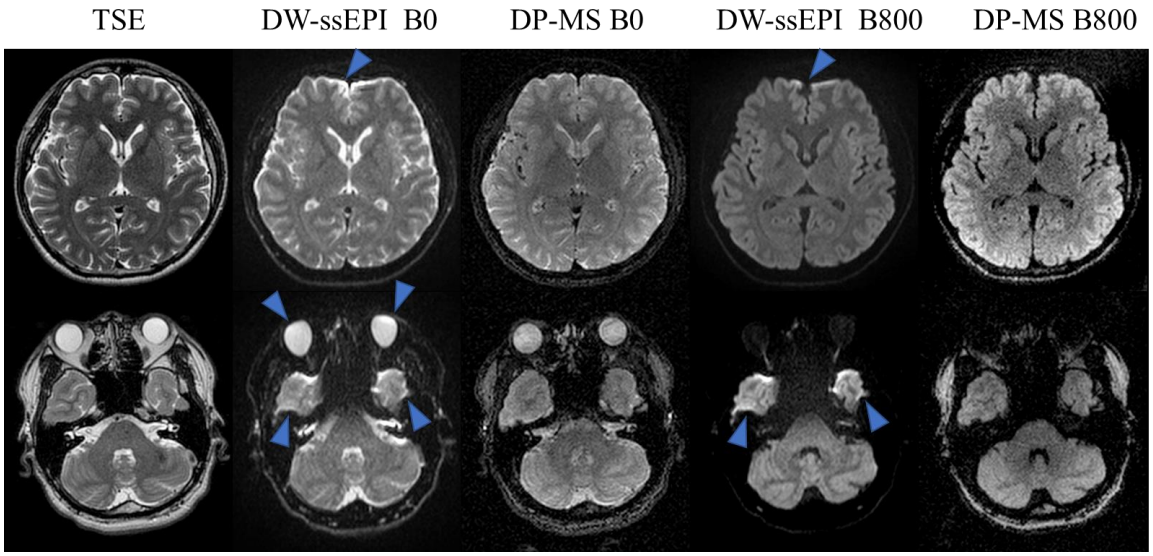

Figure 5-5 Representative transversal plane images from two volunteers. The anatomical reference image, images at b-value of 0 s/mm^2 and 800 s/mm^2 from DW-ssEPI and DP-MS were shown. The blue triangles point to distortion and signal pile-up artifacts in DW-ssEPI.

Quantitively TRE measurements from the four volunteers were shown in Table 5-1. The proposed DP-MS approach significantly improved geometric fidelity. The mean TRE of DW-ssEPI was over 2mm for all four volunteers, and the maximum displacement was almost 10mm. Overall, the DP-MS sequence provided submillimeter geometric accuracy.

Table 5-1 TRE measurement of DW-ssEPI and DP-MS on four volunteers.

TRE in mm	DW-ssEPI	DP-MS
vol1	2.40±2.04 (9.40)	0.76±0.50 (1.97)
Vol2	2.16±1.87 (8.42)	0.74±0.49 (1.88)
Vol3	2.25±1.89 (8.10)	0.73±0.38 (1.78)
Vol4	2.03±1.82 (7.41)	0.70±0.41 (1.71)

The 3-slab ADC maps from one volunteer were shown in Figure 5-6, and the Bland-Altman plots of the mean diffusivity for selected ROIs on the white matter, cerebellum, and CSF were shown in Figure 5-7. Overall, there were good agreements of ADC between the two diffusion approaches, and the systematic biases were low ($0.01\text{x}10^{-3}$mm^2/s, $0.02\text{x}10^{-3}$mm^2/s, and $-0.02\text{x}10^{-3}$mm^2/s for white matter, cerebellum, and CSF, respectively).

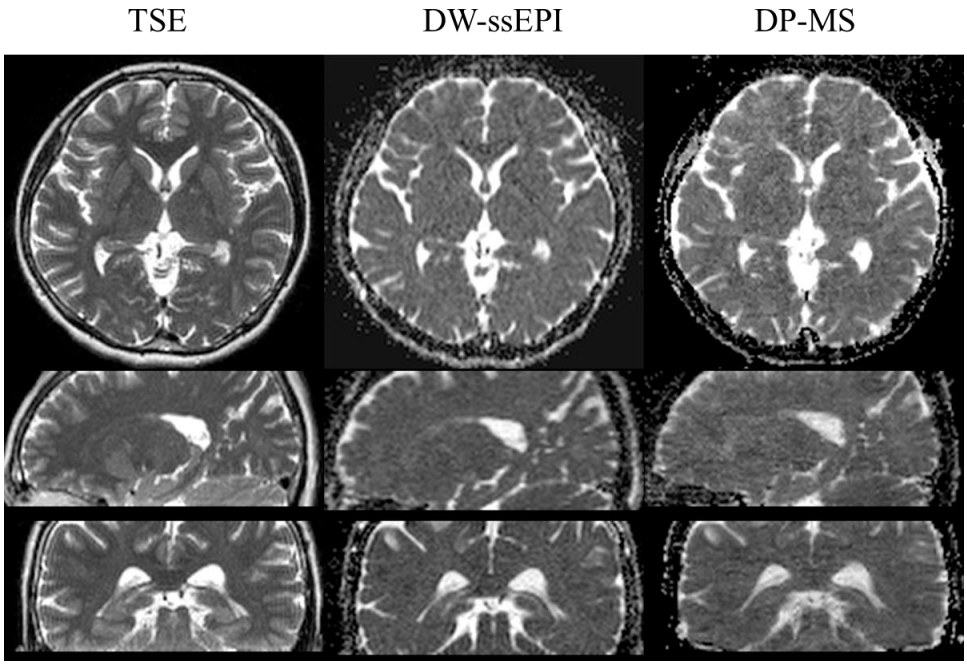

Figure 5-6 Three-plane view of the ADC map from one volunteer.

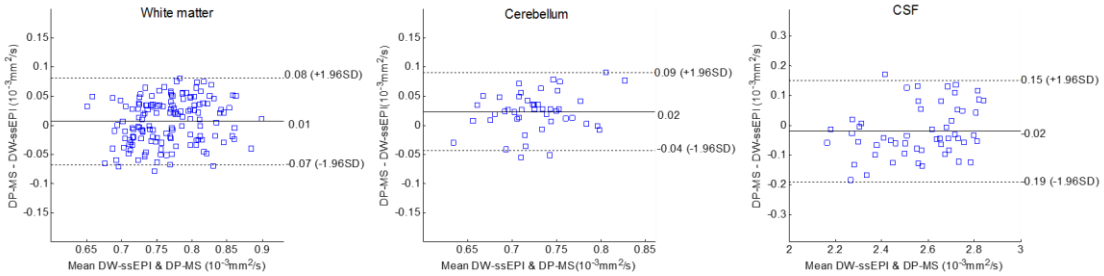

Figure 5-7 Bland-Altman plot of the white matter, cerebellum and CSF ADC between DW-ssEPI and DP-MS. Good agreements were observed for the ADC values from the two sequences.

5.4 **Discussion**

Diffusion MRI provides unique contrast reflecting tissue cellularity information for tumor characterization and response assessment. However, the strong spatial distortion prevents its utilization in the radiation oncology field, as geometric accuracy is crucial to the accurate delineation of the tumor and organs. In this work, a 3D diffusion-prepared magnitude-stabilized

bSSFP sequence was developed and validated for high-quality distortion-free diffusion imaging. Phase navigator signal was acquired during the built-in ramp up stage of the bSSFP readout to estimate the shot to shot phase variation, and a locally low-rank constrained reconstruction was developed to exploiting the local data redundancy and resolve the phase variation. The proposed sequence was validated on a diffusion phantom and healthy volunteers. It provided submillimeter geometric fidelity and desirable ADC accuracy on phantom and volunteer brain scan, which makes it a promising candidate for treatment planning and adaptive radiotherapy.

In this work, a phase inconsistency compensated locally low-rank constrained reconstruction framework was proposed. Compared with the self-navigated MUSE reconstruction[43], our strategy is less prone to g-factor effect that is relatively severe when the under-sampling factor is increased. In addition, the three-step reconstruction in MUSE requires careful tuning of multiple hyper-parameters while in the proposed one-step approach only one parameter that balances the data fidelity and locally low-rank constraint needs to be tuned. Compared with the globally low-rank reconstruction[133], which used a spiral trajectory to provide inherent navigator signal, our approach used a more robust Cartesian trajectory for data acquisition. At the same time, the locally low-rank constraint used in our method can exploit the data redundancy in a more efficient way through localizing the model space to small images blocks instead of the entire image.

Many groups have been investigating 3D diffusion sequences for the brain to achieve high-resolution high-quality diffusion imaging[88,127–129,134–136]. One common strategy is via the segmented EPI readout. In those studies, 2D or even 3D navigator was acquired after each EPI readout segment to estimate the phase and correct the shot-to-shot inconsistency. However, as mentioned in the introduction, the segmented EPI strategy can only reduce the EPI-related distortion instead of eliminating it. In this work, we selected the diffusion-prepared strategy, which

provides the flexibility to select the appropriate readout module to further reduce the distortion. The bSSFP readout was chosen due to its high SNR, high scan efficiency, and minimal distortion. As shown in the phantom and volunteer studies, the DP-MS-bSSFP sequence had submillimeter geometric fidelity.

Different from many previous works which mainly focused on qualitative and/or quantitative assessment of the image quality itself, our study also included quantitative geometric fidelity and ADC accuracy, which are important aspects for treatment planning and response assessment. High geometric accuracy is crucial for accurate tumor and organ delineation, and good ADC accuracy is important to detect tumor ADC change for adaptive radiotherapy. It is possible that one reconstructed image has high visual similarity with the gold-standard reference, but still lacks good geometric and ADC accuracy. Therefore, it is necessary to quantify the accuracy. Our study showed that the proposed method was not only able to provide images with well-delineated structures, but also allows accurate geometry and ADC measurements when compared with reference EPI acquisition. However, one has to notice that because the proposed locally low-rank reconstruction is an iterative nonlinear process, artificial ADC values can be created by manipulating regularization parameters. This is a common drawback for many iterative reconstructions. Therefore, care should be taken for the interpretation and evaluation of the ADC measured from the proposed approach.

The majority of previous 3D diffusion studies were conducted at 3T or even higher field strength, whereas many commercial MR simulators are at 1.5T. Reduced SNR caused by reduced field strength would have an adverse effect on the image reconstruction and quantification. Figure 5-8 shows the same DP-MS acquired at a 3T scanner (Prisma, Siemens Healthineers, Erlangen, Germany) using a similar protocol. Compared with Figure 5-4, images at 3T had drastically

increased SNR, and the image acquired using the 2 shot per kz protocol might already be good enough for many radiotherapy applications. Therefore, it is non-trivial to optimize and validate the sequence according to their specific system performance.

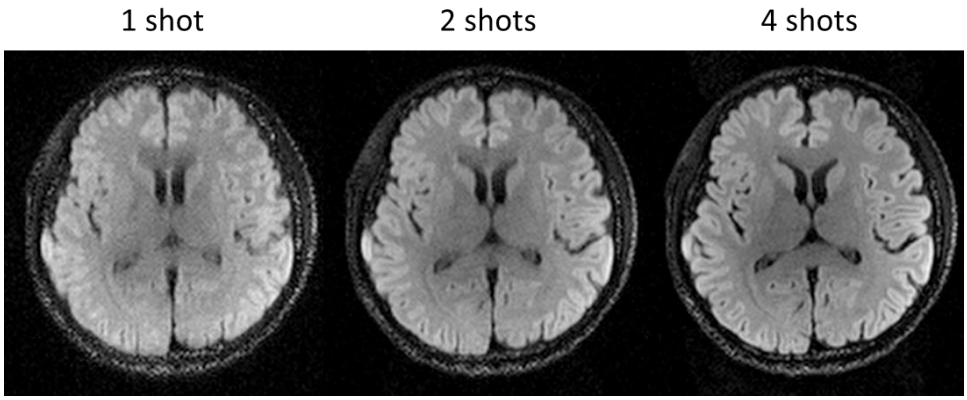

Figure 5-8 Comparison of b=800s/mm^2 images acquired at a 3T scanner using the 1 shot, 2 shots, and 4 shots per kz plane protocols.

Our work has limitations. The first limitation is that the shot to shot phase variation was estimated using a 2D navigator, leaving the through-plane motion and phase variation uncorrected. Studies have shown that the assumption of no through-plane phase variation is valid for a slab thickness of up to 30mm[127,128]. In this study, the slab thickness was 24mm, and we did not observe contamination from through-plane phase variation. However, true 3D acquisition has increased SNR for high-resolution imaging and does not have slab boundary artifacts. Therefore, it is warranted to design a 3D navigator to achieve true 3D diffusion acquisition. Secondly, the proposed sequence was validated on the diffusion phantom and healthy volunteers on the brain. We would like to optimize the sequence and apply it to other organs such as the rectum and the prostate, and evaluated it on cancer patient to study its capability for accurate tumor delineate and response assessment.

5.5 Conclusion

In this chapter, a 3D DP-MS sequence was programmed and validated on a diffusion phantom and five healthy volunteers for brain imaging. The proposed sequence had submillimeter geometry accuracy and desirable ADC accuracy, and is a promising sequence candidate for radiotherapy treatment planning and adaptive radiotherapy.

Chapter 6 Treatment response prediction: a radiomics approach

In the second half of this dissertation, we change our focus to early treatment response assessment using longitudinal DWI. Soft tissue sarcoma is heterogeneous regarding the histology. As such, its response to radiotherapy could vary a lot depending on the biology. Therefore, it is crucial to predict the early response to achieve personalized treatment and improved outcome. In this chapter, we explored radiomics approach for response prediction on localized soft tissue sarcoma undergoing hypofractionated preoperative radiotherapy. Predictive models were built to predict tumor effect score, a pathological endpoint, using selected features from longitudinal diffusion MRI. The benefits of using longitudinal data were also studied.

A version of this chapter is ready to submit: Gao Y, Kalbasi A, Shao J, Fu J, Ruan D et al. Treatment effect prediction for sarcoma patients treated with preoperative radiotherapy using radiomics features from longitudinal diffusion MRI: a preliminary study

6.1 Introduction

Soft tissue sarcoma is a rare group of tumors representing less than 1% of all malignant tumors[137]. According to the American Cancer Society, approximately 12,750 new soft tissue sarcoma cases will be diagnosed in the United States in 2019 and approximately 5,270 Americans are expected to die from it[15]. Depending on the tumor grade, tumor size and tumor type, five-year distant metastases rate ranges from less than 30% to greater than 50%, with the lung being the main site of metastases[16–19]. Among all the factors affecting the development of metastases and overall survival, the histologic grade has been recognized as the most important one[138,139]. For both

the National Cancer Institute (NCI) and the French Federation of Cancer Centers Sarcoma Group (FNCLCC) grading systems, tumor necrosis is a key factor in determining the final grade[20,140,141]. Although the correlation between tumor necrosis itself and patient overall survival remains controversial[142,143], several studies have demonstrated the prognostic value of tumor necrosis score[144–147]. For patient receiving preoperative treatment such as radiotherapy and chemotherapy, pathologic tumor necrosis is one of the most objective and reliable measurements of patient sensitivity to the treatments[143], and this treatment-induced pathologic necrosis may reflect the treatment efficacy[147,148]. In addition, hyalinization and fibrosis after treatment is another pathologic endpoint associated with favorable outcome, especially after preoperative radiotherapy[149]. Together, necrosis and hyalinization/fibrosis comprise the overall treatment effect after neoadjuvant therapy.

As soft tissue sarcomas are very diverse in terms of the biology, their response to radiation could vary significantly. Therefore, estimating and predicting treatment effect during or immediately after the treatment would be valuable in monitoring patient's response to treatment, hence provide a window for personalized treatment adaptation which enables improved treatment efficacy or reduced normal tissue complications. The superior soft-tissue contrast of MRI made it a vital tool for evaluating and predicting treatment response noninvasively for soft tissue sarcoma patients[150,151]. Diffusion MRI, in particular, has shown exceptional power in assessing tumor microstructure variations due to its ability to quantify the movement of water molecules within the extracellular region and reflecting tissue cellular density information[152–154]. The physiological information provided by diffusion MRI is potentially more sensitive and predictive than macroscopic features from the anatomical T1-weighted and T2-weighted imaging, and dynamic contrast-enhanced T1-weighted imaging (DCE) [21,151,155]. Negative correlations have been found

between tumor volume and the mean apparent diffusion coefficient (ADC) value, as well as between tumor cellularity and the minimum ADC[21,22] in patient with STS; while a positive correlation was confirmed between tumor necrosis area and the mean ADC for osteosarcoma patients[23]. Oka et al. showed that the minimum ADC was more predictive than the average ADC in evaluating the chemotherapeutic response in osteosarcoma patients[24].

However, the majority of previous studies lean heavily on simple correlations between mean ADC or minimum ADC and patient outcome using one or two time points, without assessing other features or models. In addition, because only pre-treatment and/or post-treatment imaging was acquired, the data fails to capture the dynamic changes during treatment. Extracting more features from the diffusion MRI and exploring longitudinal changes of features hold the promise of improved prediction performance. Therefore, in this study, we sought to extract features from longitudinal diffusion MRI to build a quantitative model to predict pathologic treatment effect after hypofractionated preoperative radiotherapy.

6.2 Methods

6.2.1 Patient cohort and DWI acquisition

Patients enrolled in a phase II clinical trial of hypofractionated radiation therapy for soft tissue sarcomas were recruited to a separate prospective longitudinal MRI study. All patients were diagnosed with soft tissue sarcoma of the extremity or trunk and referred to the radiation oncology department to receive neoadjuvant radiotherapy prior to the surgical removal of their resectable primary lesion. All patients had available baseline tissue specimens for pathologic assessment. For the preoperative RT, enrolled patients received a total of 30Gy delivered daily in 5 fractions with no concurrent chemotherapy. Surgery was performed usually 2-5 weeks after the completion of RT. A treatment effect score (TES) ranging from 0 to 100% was obtained from the surgical

pathology report as a surrogate of patient treatment response to the hypofractionated RT. Treatment effect refers to the degree of necrosis, hyalinization/fibrosis, and cellularity relative to the baseline biopsy. Based on the treatment effect, patients were divided into N_0 group (TES<50%) and N_1 group (TES≥50%). 50% was used as the threshold to keep consistent with the grading criteria by the French grading system[140].

A total of 36 patients were consented for the imaging study under the Institutional Review Board approval from May 2016 to June 2018. A written consent form was obtained from each patient prior to the study. Diffusion images were acquired three times throughout the RT course for each patient using a 0.35T MRI-guided radiotherapy machine (ViewRay, MRIdian, Mountain View, CA): before the first fraction of treatment (T1), immediately before or after the third fraction of treatment (T2), and immediately after the final treatment (T3). The three imaging were usually acquired every other day during the 5 daily treatment. The duration between the first and the second imaging was 3.2±1.6 days, and the duration between the second and the third imaging was 2.6±1.3 days. During each imaging session, patients were positioned at the treatment position using their corresponding immobilization device. The conventional diffusion-weighted single-shot echo-planar-imaging sequence was used to acquire the images with the following parameters: TR/TE=3400/115ms, FOV=350x350mm^2, resolution=2.7x2.7x6mm^2, slice number=10-20, bandwidth=1346Hz/px, b-value= 0, 200, 300, 400, 500 mm^2/s, averages=5. ADC maps were generated using the mono-exponential fitting.

Among the 36 patients, 6 were excluded from the analysis due to 1) no surgery performed due to patient health issue (N=2), 2) tumor is too small or tumor is difficult to be delineated on the DWI images (N=3), and 3) no residual tumor identified so the pathologic treatment effect could

not be assessed (N=1). Table 6-1 is the characteristics of the remaining 30 patients included in the analysis.

Table 6-1 Patient characteristics. (N=30)

Characteristic	Type	No. of patient (%)
Sex	Male	20 (67%)
	Female	10 (33%)
Age	Mean+std (range)	59±15 (28-82)
Primary site	Lower extremity	23 (77%)
	Upper extremity	7 (23%)
Histology	Undifferentiated pleomorphic sarcoma	11 (37%)
	Myxoid liposarcoma	7 (23%)
	Myxofibrosarcoma	4 (13%)
	Spindle cell sarcoma	3 (10%)
	Sarcoma not otherwise specified	1 (3%)
	Others	8 (13%)
Grade	1	2 (7%)
	2	12 (40%)
	3	16 (53%)
Treatment Effect	N_0 (TES<50%)	12 (40%)
	N_1 (TES≥50%)	18 (60%)

6.2.2 Feature extraction

For each DWI dataset, the gross tumor region was manually contoured on the diffusion-weighted images by a radiation oncologist with reference of gross tumor volume (GTV) contour from treatment planning images. A second radiation oncologist was involved to review and edit the contour. This contour was then transformed to the ADC map for feature extraction. Figure 6-1 shows an example of the contouring on the DWI image and the ADC map.

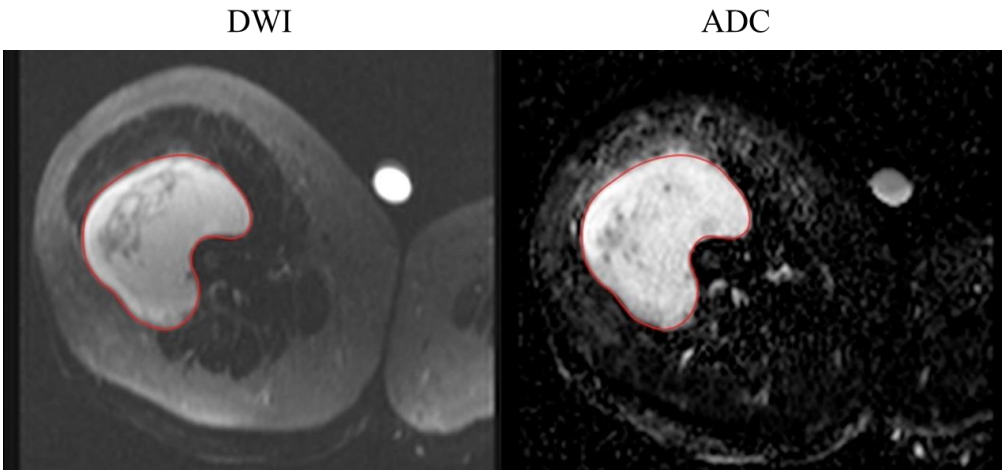

Figure 6-1 Example of DWI image and the ADC map. Tumor is contoured in red.

A total of 106 features were extracted from tumor ADC maps at each imaging time point using the Pyradiomics[46]. As shown in Table 6-2, these features included 18 first-order features, 14 shape features, 23 Gray Level Co-occurrence Matrix Features, 16 Gray Level Run Length Matrix Features, 16 Gray Level Size Zone Matrix Features, 5 Neighboring Gray Tone Difference Matrix features, and 14 Gray Level Dependence Matrix features. Delta radiomics features between different time points were calculated. These include: D31=F3-F1, D32=F3-F2, D21=F2-F1, where Fi represents features from time point i. Overall, a total of 636 features were collected for each patient. Features were then normalized by subtracting the mean and dividing by the standard deviation to improve prediction and convergence rates.

Table 6-2 Features used in this study.

First Order Feature (N=18)
Mean, Median, Minimum, Maximum, Range, 10th percentile, 90th percentile, Interquartile Range, Variance, Energy, Entropy, Kurtosis, Mean Absolute Deviation, Robust Mean Absolute Deviation, Root Mean Squared, Skewness, Total Energy, Uniformity
Shape Feature (N=14)
Elongation, Flatness, Least Axis Length, Major Axis Length, Maximum 2D Diameter Column, Maximum 2D Diameter Row, Maximum 2D Diameter Slice, Maximum 3D Diameter, Mesh Volume, Minor Axis Length, Sphericity, Surface Area, Surface Volume Ratio, Voxel Volume
GLCM Feature (N=23)
Autocorrelation, Joint Average, Cluster Prominence, Cluster Shade, Cluster Tendency, Contrast, Correlation, Difference Average, Difference Entropy, Difference Variance, Joint Energy, Joint Entropy, Informational Measure of Correlation 1, Informational Measure of Correlation 2, Inverse Difference Moment, Inverse Difference Moment Normalized, Inverse Difference, Inverse Difference Normalized, Inverse Variance, Maximum Probability, Sum Entropy, Sum Squares, Maximal Correlation Coefficient
GLRLM Feature (N=16)
Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Gray Level Variance, High Gray Level Run Emphasis, Long Run Emphasis, Long Run High Gray Level Emphasis, Long Run Low Gray Level Emphasis, Low Gray Level Run Emphasis, Run Entropy, Run Length Non-Uniformity, Run Length Non-Uniformity Normalized, Run Percentage, Run Variance, Short Run Emphasis, Short Run High Gray Level Emphasis, Short Run Low Gray Level Emphasis
GLSZM Feature (N=16)
Gray Level Non-Uniformity, Gray Level Non-Uniformity Normalized, Gray Level Variance, High Gray Level Zone Emphasis, Large Area Emphasis, Large Area High Gray Level Emphasis, Large Area Low Gray Level Emphasis, Low Gray Level Zone Emphasis, Size Zone Non-Uniformity, Size Zone Non-Uniformity Normalized, Small Area Emphasis, Small Area High Gray Level Emphasis, Small Area Low Gray Level Emphasis, Zone Entropy, Zone Percentage, Zone Variance
GLDM Feature (N=14)
Dependence Entropy, Dependence Non-Uniformity, Dependence Non-Uniformity Normalized, Dependence Variance, Gray Level Non-Uniformity, Gray Level Variance, High Gray Level Emphasis, Large Dependence Emphasis, Large Dependence High Gray Level Emphasis, Large Dependence Low Gray Level Emphasis, Low Gray Level Emphasis, Small Dependence Emphasis, Small Dependence High Gray Level Emphasis, Small Dependence Low Gray Level Emphasis
NGTDM Feature (N=5)
Busyness, Coarseness, Complexity, Contrast, Strength
Delta Radiomics
D31=F3-F1, D32=F3-F2, D21=F2-F1

* Fi: feature from time point i.

6.2.3 **Prediction using mean ADC alone**

Prediction using only the mean ADC as the feature was performed to serve as the benchmark. SVM models were built using the mean ADC from any single time points, ADC1, ADC2, ADC3, and using the delta ADC from any two time points, ADC_D21 (ADC2-ADC1), ADC_D31(ADC3-ADC1), ADC_D32 (ADC3-ADC2). Fivefold cross-validation with 50 repetitions was conducted to estimate the stability of the prediction.

6.2.4 **Feature reduction and selection**

Feature set reduction and feature selection were performed to identify the most predictive features that will be used to construct the classification model. During the feature set reduction process, features were ranked based on their predictive power quantified using AUC. Features that had a correlation coefficient greater than 0.8 but lower predictive power were removed to reduce the redundancy between features. After feature number was reduced to 25, the sequential forward selection was performed to identify the most predictive features among this reduced feature set for logistic regress and support vector machine respectively. Prediction AUC using up to 10 features was calculated, and the best feature number k was determined based on the AUC curve with respect to feature number. Fivefold cross-validation with 20 repetitions was performed when evaluating model AUC to improve the model robustness.

6.2.5 **Treatment outcome prediction**

The top k features identified from the sequential forward selection process were used to classify N_0 and N_1 groups using logistic regression and support vector machine. Five-fold cross-validation with 50 repetitions was conducted to evaluate the stability of the classification model. The AUC, sensitivity, specificity, and accuracy for each repetition were recorded. Significant

difference between logistic regression and SVM was evaluated using paired t-test at the level of $p<0.05$.

To evaluate the benefits of adding additional imaging time points and delta radiomics features, the prediction performance of SVM using features from a single time point (F1, F2, F3), using features from two time points without delta radiomics features (F1+F2, F1+F3, F2+F3), using features from all three time points without delta radiomics features (F1+F2+F3), using features from two time points with delta radiomics features (F1+F2+D21, F1+F3+D31, F2+F3+D32), and using features from all three time points and all delta radiomics features (F_{all}) were evaluated with the same workflow. Paired t-test was used to evaluate whether the performance between F_{all} and the others was significant at level $p<0.05$.

6.3 Result

Prediction AUC using mean ADC or delta ADC alone was shown in Figure 6-2. Overall, performance using only ADC as the feature was poor. The highest AUC was achieved using ADC_D31 (AUC=0.69±0.06).

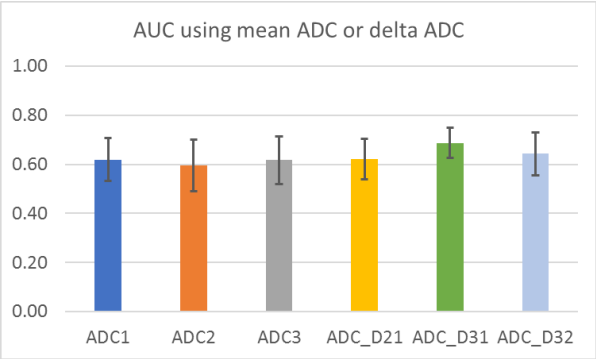

Figure 6-2. Prediction performance of SVM using only mean ADC or delta ADC

Figure 6-3 is the plot of model AUC with respect to increasing number of features during the sequential forward selection process for T_{all}. For logistic regression, the AUC reached a maximum when 3 features were used in the model and decreased with 4 or more features, possibly due to

overfitting. For SVM, the AUC reached a maximum near 4 features and then stabilized. Therefore, by inspecting the model AUC curves, 3 features were selected in logistic regression and 4 features were chosen for SVM. The 3 features selected by logistic regression were: GLSZM Zone Entropy D31, GLCM Maximal Correlation Coefficient D31, and GLDM Large Dependence High Gray Level Emphasis D31. The top 3 features identified by SVM were the same as that of logistic regression. The 4th feature of SVM was GLDM Large Dependence Low Gray Level Emphasis D31. All selected features belong to the D31 group, indicating the delta radiomics of pre-RT and post-RT is predictive of the treatment effect score.

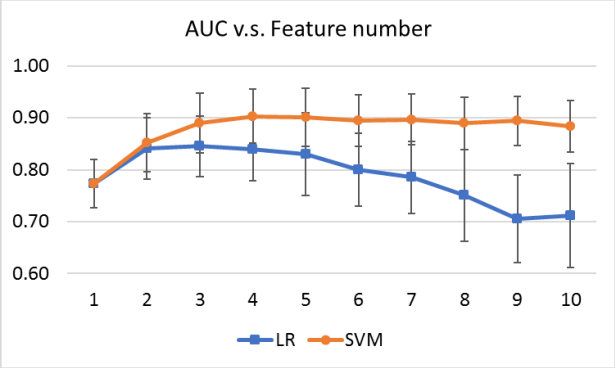

Figure 6-3 Model AUC with respect to different feature numbers from 1 to 10 for logistic regression and SVM. The error bar indicates the standard deviation. The best feature number k was determined by inspecting the AUC curve. In this case, three features were identified for LR and four features were selected for SVM

Prediction performance of logistic regression using the three selected features and SVM with the four selected features were shown in Table 6-3. SVM significantly outperformed logistic regression for all statistics. An AUC of 0.91 and an accuracy of 0.92 were achieved using SVM while the AUC and accuracy were 0.85 and 0.87 for logistic regression.

Table 6-3 Prediction performance using features from all three time points and corresponding delta radiomics.

	LR	SVM
AUC	0.85 ± 0.06	0.91 ± 0.05
Sensitivity	0.85 ± 0.07	0.90 ± 0.08
Specificity	0.94 ± 0.05	0.97 ± 0.04
Accuracy	0.87 ± 0.04	0.92 ± 0.04

Comparison of prediction performance using different time points and with/without delta radiomics features was shown in Figure 6-4. Prediction performance using only single time point was poor. AUC for F1, F2, and F3 were 0.70±0.06, 0.67±0.08, and 0.70±0.08 respectively. Including a second or third time point without delta radiomics does not necessarily improve the prediction. AUC for F1+F2, F1+F3, F2+F3, and F1+F2+F3 were 0.70±0.06, 0.73±0.07, 0.70±0.08, and 0.73±0.07 respectively. Including delta radiomics D21 or D31 drastically boosted the prediction (AUC=0.90±0.06 for F1+F2+D21, and 0.91±0.04 for F1+F3+D31), whereas D32 did not have obvious contribution in improving the AUC (AUC=0.72±0.09). For all prediction statistics, there is no significant difference between F_{all} with F1+F2+D21 or F1+F3+D31.

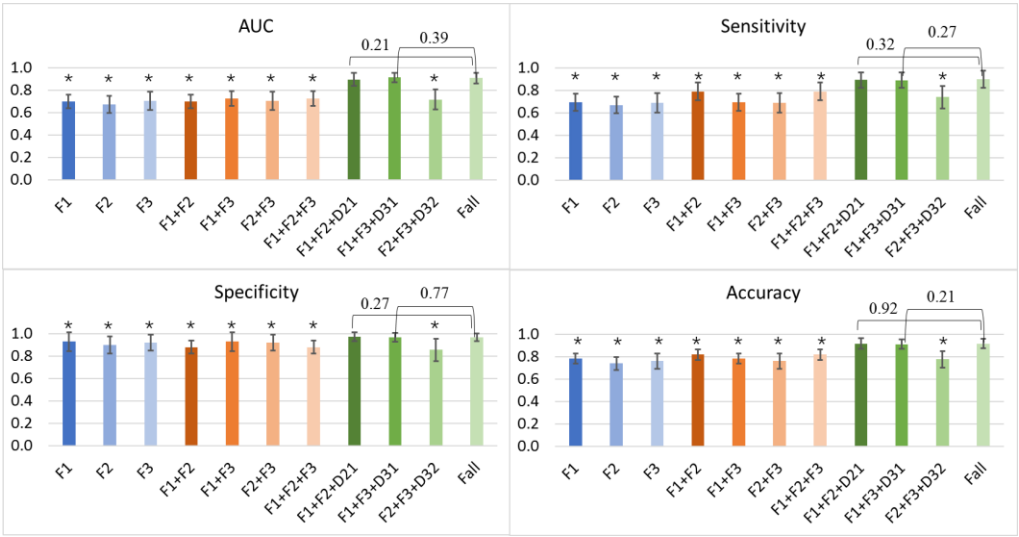

Figure 6-4 Prediction performance using different time points. The error bar indicates the standard deviation. The star (*) implies a significant difference with the performance of F_{all}.

6.4 Discussion

Magnetic resonance imaging is a standard modality in the diagnosis and treatment planning of soft tissue sarcoma[156]. In addition to standard MRI sequences, diffusion MRI may provide additional information to characterize soft tissue sarcoma at diagnosis[22,157–159]. In other tumor types, diffusion MRI has shown promise as a tool to assess treatment response[160–163]. However, there is limited data on the role of diffusion MRI in assessment of treatment response in soft tissue sarcoma; the existing data in this field lean heavily on simple correlations between mean ADC or minimum ADC and patient outcome using one or two time points, without assessing other more complex predictive features or predictive models. In this study, we showed that for our patient cohort, mean ADC or delta ADC alone did not reach a high prediction performance for treatment effect prediction. Radiomics features were needed to better capture tumor texture changes to improve the prediction.

Establishing the optimal imaging acquisition timing during a radiotherapy treatment course is a critical step toward an efficient large-scale treatment response study trial. In this study, diffusion imaging was acquired three times throughout the 5-fraction treatment. By comparing the prediction performance using different time point, it can be seen that information from single or multiple time points was not sufficient to reflect the tumor's response to the treatment. It is crucial to include the change of feature to capture tumor response to radiation. We also found that baseline information is needed for a good prediction. It needs to be combined with either mid-treatment or post-treatment data to calculate the delta radiomics. Our results indicate that acquiring two imaging time points, either pre- and mid-, or pre- and post-, is sufficient for response prediction on our patient cohort.

All images of this study were acquired using the onboard MRI scanner of the ViewRay system. This MR-guided radiotherapy system made the longitudinal imaging acquisition throughout the treatment logistically practical. In addition, the large bore and flat table design of the system allow the use of most of the immobilization devices, so that the imaging could be set up based on patient's treatment position and reproduced for all following imaging sessions. However, there are several limitations of the MRgRT system. The field strength of the MR component is 0.35 Tesla designed specifically for the integration of MRI and RT. However, the low magnetic field leads to low image SNR, and the signal level of DWI approaches noise floor level with increasing b value[164]. Therefore, the highest b-value used in this study was 500 mm^2/s. Excluding high b-value could have a negative impact on the detection of tissue with low diffusivity.

At the time of compiling this study, other commonly used outcome evaluation criteria such as the development of metastases or long-term survival data were not available. Treatment effect score was chosen in this proof-of-concept study for two reasons. First, as one direct impact of RT is cell-death, this radiation-induced pathologic treatment effect (which accounts for necrosis) intuitively reflects treatment efficacy. Second, although still controversial, several studies have shown the prognostic value of necrosis score/pathologic treatment effect[144–148]. Therefore, accurate prediction of treatment effect during/after the RT but before the surgery could provide valuable information to patients' care team to plan or adjust treatment strategies.

The main limitation of this work is the small patient cohort. 52 patients were recruited to participate in the single-institution preoperative trial of hypofractionated radiotherapy for from May 2016 to June 2018. Among them, 36 patients were enrolled in this longitudinal MRI study, and 30 patients had a full data set available for analysis. Repetitive cross-validation was applied to estimate the robustness. However, the results were potentially biased because we lack an

independent test set. We are now enrolling more patients on an expansion cohort and setting up an individual test set to build a more robust and accurate model. In addition, manual contour is required for feature extraction, and this contouring process has variability. Features, particularly shape features, could be sensitive to contour uncertainty. Other classification models that do not require tumor segmentation, such as conventional neural network-based classifier, could be implemented for treatment effect prediction. Lastly, although patients enrolled in this study received the same treatment scheme, their histology subtypes are different. Sarcoma is known for its histologic and biologic diversity, and these differences are often reflected in diverging imaging characteristics. The lack of homogeneity in imaging impacts the power of our study. Ultimately, we aim to study patients based on the histology category and analyze their response to treatment separately once a larger patient cohort is available.

6.5 Conclusion

Radiomics features from longitudinal diffusion MRI were explored to predict post-surgery tumor necrosis score after RT for sarcoma patients. The SVM model built with predictive radiomics features provided high prediction performance. Based on the observations of this study, we found that mean ADC, or delta ADC, or radiomics features alone was not sufficient for response prediction. Including delta radiomics features of mid- or post- treatment relative to the baseline can optimize the prediction of treatment effect score, a pathologic and clinical endpoint.

Chapter 7 Treatment response prediction: a deep learning approach

In chapter 6, a radiomics-based approach was investigated for early treatment response prediction. However, there was no independent test data in the model due to the small patient size, leaving the model vulnerable to overfitting. In this chapter, an ACGAN-based model was built to augment the patient data. The VGG-19 network was trained to predict treatment effect score on independent test patients.

A version of this chapter is ready to submit: Gao Y, Ghodrati V, Kalbasi A, Fu J, Ruan D, et al. Preliminary study of deep learning-based treatment response prediction using longitudinal diffusion MRI on sarcoma patients

7.1 Introduction

Soft tissue sarcomas are a rare group of tumors and are heterogeneous with regard to location, histologic type, and prognosis[14]. Surgery removal of the tumor is the main treatment scheme for soft-tissue sarcomas, and radiotherapy was usually given pre- or post- the surgery to facilitate the tumor resection. Therefore, assessment of pre-operative radiotherapy treatment could be crucial in achieving personalized treatment, as follow-up treatment and timing of surgery may vary depending on the treatment response. MRI has become the methodology of choice for early treatment response assessment as it provides exquisite soft-tissue contrast[150,151]. Diffusion MRI, in particular, is gaining popularity since the functional information it provides is potentially more sensitive than anatomical MRI [21,151,155]. Studies have shown that baseline apparent diffusion coefficient value and change of ADC during the treatment are predictive of the patient response[21–

[24]. However, most of the existing studies lean heavily on simple correlations between ADC and patient outcome, without exploring quantitative models to predict the response. A predictive model would enable early detection of complete responders and non-responders, which opens the window for treatment adaptation.

Deep learning is growing rapidly in medical imaging analysis. It becomes the methodology of choice for image segmentation, detection, classification, and registration for various disease sites[53–55]. Several studies have applied neural networks to predict treatment response using CT images on lung, rectal, and bladder patient[55–57]. Promising results were obtained which outperformed the conventional linear regression or support vector machine. However, to the best of our knowledge, deep learning has not been applied to predict treatment response on sarcoma patients.

Another problem encountered by many medical applications is the small data size, which consequently limits the stability and generalizability of the model. Data augmentation is a well-established technique to artificially increase the dataset size[165,166]. Simple data augmentation includes image translation, rotation, flip, and scale. However, information gained from those simple modifications is limited. Another emerging technique to enrich the dataset is Generative Adversarial Networks (GANs), in which two networks were trained in an adversarial process: one discriminator network was trained to discriminate between real and fake images, whereas the other network was trained to generate fake images to fool the discriminator[58]. High-quality realistic images were synthesized via this competing process. GAN-based data augmentation framework has been applied to many medical applications such as image segmentation and lesion classification[59–62]. Frid-Adar et al. showed that improved liver lesion classification was achieved by adding the synthetic augmentation than the classic augmentation[61]. Madani et al. found that

GAN data augmentation provided higher accuracy than traditional data augmentation on cardiovascular abnormality classification using X-ray image[62]. Whether GAN-based data augmentation could facilitate treatment response prediction remains unknown.

Therefore, the goal of this study is to employ a deep learning-based workflow to predict treatment response on soft tissue sarcoma patients. A GAN-based model was trained to augment the longitudinal diffusion images for improved prediction.

7.2 Methods

7.2.1 Problem formulation

Treatment response prediction consists of feature extraction and classification. In the classical approach, the first step has been achieved by extracting explicit human-crafted features of the data and the latter one by building a linear or nonlinear classifier to separate the features. The mentioned steps in the modern deep feedforward neural network has been considered together and through the training episodes, most effective implicit features would be extracted. Following is a simple mathematical illustration for the general classification network.

The goal of the deep feedforward neural network in classification is learning to approximate the most optimal function F^* that maps the input data to their corresponding labels. In general, a deep classifier network in image classification consists of a sequential chain of convolutional layers. A deep feedforward neural network with n-layer can be represented as a composite of n nonlinear function as:

$$F(x) = f^n(f^{n-1}(f^{n-2}(\dots . f^1(x) \dots)))$$

where, f^1 represents the first convolution layer that applied for the image input x, and output of first layer $f^1(x)$ is feeding to the second layer. Typically, each layer in a convolutional neural network composed of convolutional kernels followed by non-linear activation i.e. ReLU.

In multi-category classification, assigned non-linearity to the last convolution layer is a soft-max function that produces the probability for a pre-activation z_i:

$$Softmax(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_i)}$$

In the training process, trainable parameters of $F(x)$ i.e. convolutional kernels and biases are updated until a specific distance between the $F(x)$ and optimal function F^* reduced enough. Log-likelihood or cross-entropy is the most widely used distance in classification tasks:

$$\log softmax(z)_i = z_i - log \sum_j \exp(z_j)$$

which consists of two parts: the first part z_i represents the true probability and second term negative log of sum of the false probabilities. Intuitively, maximizing cross-entropy or negative log-likelihood means increasing true prediction and pushing down the false predictions. In the training phase of a deep classifier network, all the weights and bias values are adjusted through backpropagation chain rules to maximize the probability of the ground truth labels for a given image.

7.2.2 **Patient cohort**

The patient cohort is the same as Chapter 6. All patients were recruited prospectively from a phase II clinical trial of hypofractionated radiation therapy for soft tissue sarcomas. They were diagnosed with soft tissue sarcoma of the extremity or trunk and received a total of 30Gy delivered daily in 5 fractions. No concurrent chemotherapy was performed. Post-therapy removal of

respectable primary lesion was performed 2-5 weeks after the completion of RT. A pathological treatment effect score was assessed which reflects the degree of necrosis, hyalinization/fibrosis, and cellularity of the tumor relative to the baseline biopsy. This score ranges from 0 to 100% where a high score represents a good response to the RT treatment. Patients were divided into three classes based on the treatment effect score: poor response class C1 (TES≤20%), moderate response class C2 (20%<TES<90%), and good response class C3 (TES≥90%)[24,167,168]. A total of 30 patients were recruited and had a complete imaging dataset. Patient characteristics were summarized in Table 7-1.

Table 7-1 Patient characteristics

Characteristic	Type	No. of patient (%)
Sex	Male	20 (67%)
	Female	10 (33%)
Age	Mean+std (range)	59±15 (28-82)
Primary site	Lower extremity	23 (77%)
	Upper extremity	7 (23%)
Histology	Undifferentiated pleomorphic sarcoma	11 (37%)
	Myxoid liposarcoma	7 (23%)
	Myxofibrosarcoma	4 (13%)
	Spindle cell sarcoma	3 (10%)
	Sarcoma not otherwise specified	1 (3%)
	Others	8 (13%)
Grade	1	2 (7%)
	2	12 (40%)
	3	16 (53%)
Treatment Effect	C1 (TES≤20%)	6 (20%)
	C2 (20%< TES <90%)	19 (63.3%)
	C3 (TES ≥90%)	5 (16.7%)

7.2.3 **Imaging protocol**

The diffusion imaging study was performed three times throughout the treatment course on a 0.35T MRI-guided radiotherapy machine (ViewRay, MRIdian, Mountain View, CA) using the DW-ssEPI sequence. The first imaging was performed before the start of RT treatment, the second

imaging was acquired before/after the third fraction of RT, and the last imaging was acquired after the completion of RT. Diffusion images acquired at the second and third timepoint were registered to the image acquired at the first timepoint. The gross tumor volume was manually segmented by a radiation oncologist on diffusion images and transferred to the fitted ADC maps.

7.2.4 **Data augmentation network**

To cope with the small datasets and the limited number of annotated samples as well as satisfying the requirements of a large data size for a deep learning model, an ACGAN model was used to augment the data[64] (Figure 7-1 (a)). ACGAN is a variant of conditional GANs in which an additional auxiliary decoder was added on the discriminator to enable the model to output the class label as well as the real and fake decision. A class label ([1,0,0], [0,1,0], or [0,0,1]) and a random noise image were fed into the generator G. The generator tries to synthesize the data to be indistinguishable with the real data, while the discriminator tries to increase the sum of the log-likelihood of the real image and true class. The generator used in this study was the same as the generator in the original ACGAN. As shown in Figure 7-1 (b), the discriminator has six layers. Each of the first four layers starts with a convolutional layer with kernel size 3 and stride 2, followed by the batch normalization and leaky ReLU. The last two are dense layers with 100 and 1 neurons.

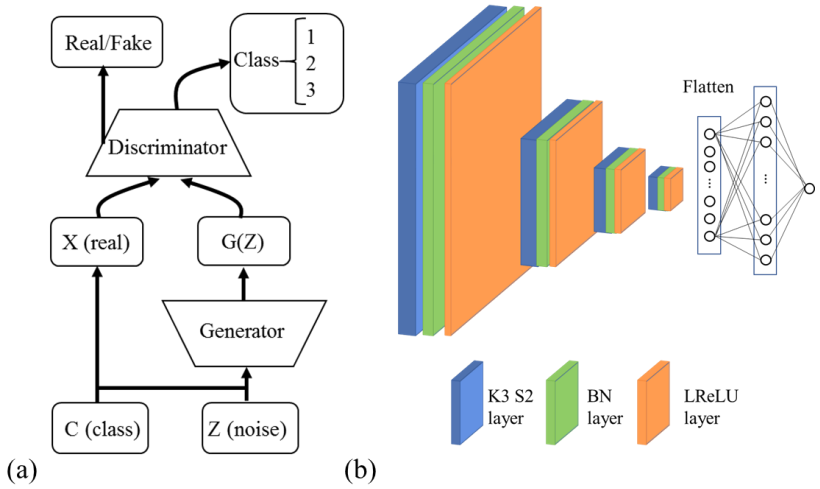

Figure 7-1 (a) Architecture of the ACGAN network. (b) Details of the discriminator.

7.2.5 **Response prediction network**

Treatment response prediction network was constructed based on a deep convolutional network VGG-19 proposed by the Visual Geometry Group[65]. As illustrated in Figure 7-2, two dense layers with 10 neurons and 3 neurons were added on top of the pre-trained VGG-19 to predict the treatment response class. ADC maps from the three time points were fed into the network as three channels. Categorical cross-entropy was used as the objective function.

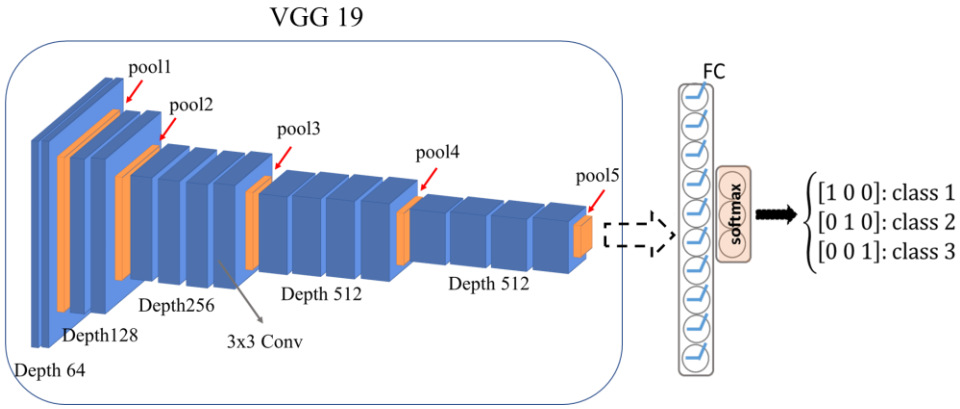

Figure 7-2 Architecture of the VGG-19 based prediction network structure

7.2.6 **Workflow**

The overall workflow of the ACGAN-VGG network was shown in Figure 7-3. It consists of two networks: data augmentation network and response prediction network. In the augmentation step, four patients (one from C1, two from C2 and one from C3) were randomly selected to serve as the final test set. The ACGAN augmentation network was trained on the remaining 26 patients (around 300 slices). 15,000 synthetic slices were generated for each class using the trained ACGAN model, resulting in a total of 45,000 synthesized slices.

The synthesized slices were then used to train the VGG-19 prediction network, while the original data from the 26 patients were used as the validation set. Separating the generated and real patient images as training and validation also severs as an evaluation of the performance of ACGAN: a close training and validation accuracy would imply that the generated samples are a good representation of the original patient data. This trained model was then tested on the four patients that were unseen from the data augmentation and prediction network training process.

To demonstrate the stability and reliability of the ACGAN-VGG workflow, the entire data augmentation and response prediction process was repeated five rounds. Four different patients were removed as the test set during each round, and the data augmentation and prediction model training were based on the remaining 26 patients (296-313 slices).

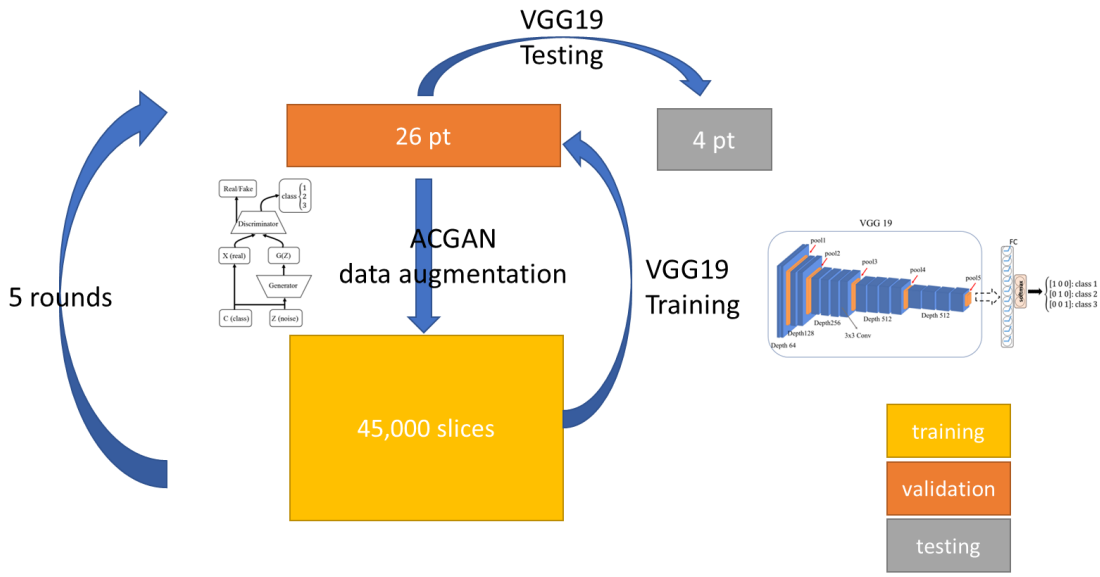

Figure 7-3 Overall workflow of ACGAN-VGG network.

7.2.7 **Evaluation**

To show the similarity between the generated data and real patient data, we trained the prediction network based on the synthesized data and validated it based on real patient data. Training and validation accuracies were reported. As the gap between training and validation accuracy lower, the distribution of synthesized data is closer to real data.

The prediction performance was evaluated in two folds: the accuracy of slice-based prediction and the accuracy of patient-based prediction. In the patient-based prediction, the majority predicted class was assigned for each patient. Accuracy, sensitivity, and specificity metrics were reported. Those metrics were defined as followed:

$$\text{Accuracy} = \frac{\sum TP}{Total\ tumor\ slices}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

where P represents the slices from the interested category, and N represents slices from the other two categories.

7.3 Result

7.3.1 Data augmentation

The trained ACGAN model produced 15,000 synthetic slices for each class. Figure 7-4 shows longitudinal ADC maps of one patient example and two synthetic examples for the three classes. The synthesized ADC images had similar patterns as the real patient ADC maps, and had wide variations in terms of shape and intensity. This intuitively demonstrates that the trained ACGAN model could produce high-quality ADC maps that mimic the real patient data.

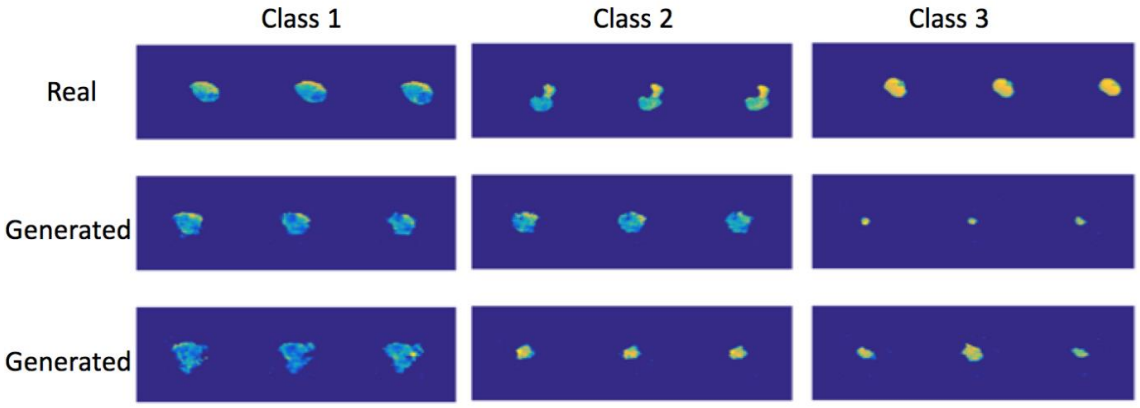

Figure 7-4 Examples of longitudinal patient data and ACGAN synthesized data for each class. Two generated examples were shown. For each case, ADC maps from the three time points were displayed.

7.3.2 **Response prediction**

The prediction accuracy on training, validation, and test sets over the five rounds were shown in Figure 7-5. The averaged accuracy was: 94.3%±2.7%, 90.1%±4.7%, and 87.7%±4.9% for training, validation, and testing, respectively. It is to be emphasized that the prediction network was trained on the ACGAN generated images and evaluated on the real patient ADC images. Therefore, the matching training and validation accuracies imply that the synthesized data resemble the real data. In other words, this confirmed the validity of the ACGAN data augmentation network. The prediction accuracy on the unseen test set was close to the validation accuracy. This indicated the trained VGG-19 prediction network had stable performance and no overfitting issue. Overall, round 2 had the worst performance (accuracy=79.6%) on the test set and round 4 had the best prediction (accuracy=92.9%).

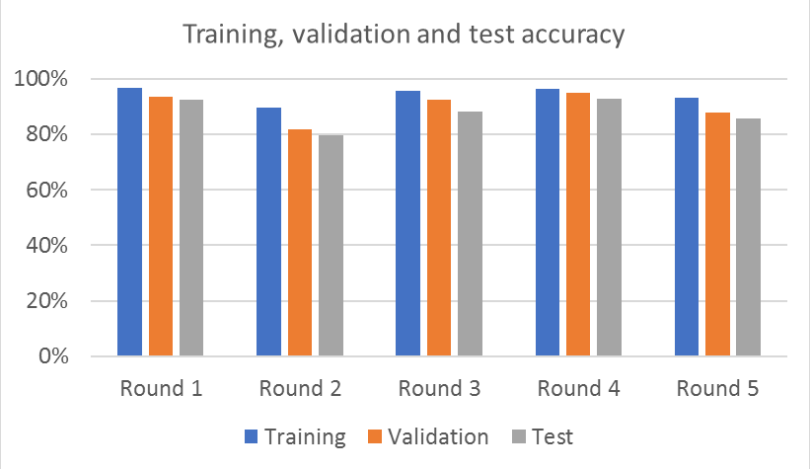

Figure 7-5 Training, validation, and test accuracy over the five rounds

Prediction results on the test sets over the five rounds were summarized as a confusion matrix in Table 7-2. The VGG-19 network had the best performance in identifying class 3, which provided a sensitivity of 91% and specificity of 97%. Similar sensitivity and specificity were achieved for class 2 (89.8% and 89.2%). Specificity was higher than sensitivity in class 1 (93.8% and 81.6%).

Table 7-2 Confusion matrix of the slice-based prediction results on the test sets.

	Pred C1	Pred C2	Pred C3	Sensitivity
Actual C1	62	10	4	81.6%
Actual C2	9	97	2	89.8%
Actual C3	1	4	48	90.6%
Specificity	93.8%	89.2%	96.7%	

The confusion matrix for the patient-based prediction was shown in Table 7-3. The model had 100% sensitivity and specificity on predicting class 3. One patient that belongs to C1 was predicted as C2.

Table 7-3 Confusion matrix of the patient-based prediction results on the test sets.

	Pred C1	Pred C2	Pred C3	Sensitivity
Actual C1	4	1	0	80%
Actual C2	0	10	0	100%
Actual C3	0	0	5	100%
Specificity	100%	90%	100%	

Table 7-4 shows the detailed prediction results on the test sets. The “slice” column shows the correctly predicted slice number over the total slice number for each patient. In the “patient” column, “F” represents incorrect classification and “T” represents accurate classification. For patient 1 in round 1, 2 out of 4 slices were classified incorrectly. Therefore, the model was not able to classify it correctly. For all the remaining patients, more than 75% of slices were predicted correctly. Overall, the patient-based accuracy was 95%.

Table 7-4 Prediction results on the test sets for each patient. Slice column: correctly predicted slice number over the total number of slices. Patient column: F - incorrect classification; T - accurate classification

	Round 1		Round 2		Round 3		Round 4		Round 5	
	Slice (Cor./tot.)	patient	Slice (Cor./tot.)	patient	Slice (Cor./tot.)	patient	Slice (Cor./tot.)	patient	Slice (Cor./tot.)	patient
Patient 1	2/4	F	12/14	T	9/10	T	8/8	T	6/8	T
Patient 2	8/8	T	4/5	T	17/19	T	13/13	T	10/10	T
Patient 3	18/18	T	15/20	T	9/10	T	5/6	T	17/20	T
Patient 4	8/9	T	8/10	T	10/12	T	13/15	T	15/18	T

7.4 Discussion

In this work, we sought to predict the post-RT treatment effect score of sarcoma patients using longitudinal DWI. An ACGAN data augmentation network was trained to augment the data for improved stability and generalizability of the prediction. The synthesized ADC maps along with patient images were employed to train and validate the VGG-19 predication network, which was then tested on unseen patient data. The whole process was repeated five times to evaluate the stability of the proposed workflow. High accuracies were achieved on the training, validation and test datasets. Overall the accuracy was 88% on the independent test sets for slice-based prediction, and 95% for patient-based prediction. To the best of our acknowledge, this is the first work that employs longitudinal diffusion data and deep learning to predict treatment effect on sarcoma patients. Although the result is still preliminary, it demonstrated the feasibility of the proposed workflow on predicting post-RT treatment response.

Early treatment response prediction holds the promises of achieving personalized treatment. Patients in this study received neoadjuvant radiation therapy before surgical removal of their respective primary lesions. The capability of assessing and predicting their response to RT provides a window for personalized treatment adaptation: surgery removal may be avoided in

patients with complete response to RT, whereas radiation boosting might be beneficial to further reduce tumor margin in the non-responding region. All these potentially enable improved treatment efficacy and/or reduced normal tissue complications.

Empirical results have been shown that the generalization power of the deep networks is more than the shallow networks[169]. Also, obviously using large datasets in terms of variety could result in effective training and consequently boost test performance of the deep neural networks. However, it is challenging to access and acquire a large size of longitudinal data in the medical field. In this study, we focused on sarcoma patient, which is a relatively rare group of tumor representing less than 1% of all malignant tumors[137]. A total of 52 patients were recruited in our institute to participate in this preoperative clinical trial from May 2016 to June 2018. 36 of them were consented to the longitudinal DWI study, with only 30 patients had a complete dataset. To overcome the limited dataset size, we adopted data augmentation. In this work, we employed the ACGAN network to augment the data[64]. A total of 45,000 slices were synthesized based on images from 26 patients (around 300 slices). It is shown that trained ACGAN network could generate synthetic images that learned the density of the real patient data. Those generated images had a wide variation on intensity and shape.

Although GAN has become a popular tool to generate images, evaluating the performance of GAN remains challenging. Conventional evaluation metrics include Inception score and Frechet Inception distance[170–172]. However, all these rely on pre-trained networks developed based on natural image database, hence are not applicable in our application. Therefore, we modified a newly proposed GAN-train GAN-test criteria to assess the performance of synthesized images[173]. In the original paper, two classification networks were trained: one network was trained using GAN generated images and tested on the real data to compute the GAN-train score, and another

network was trained using the real data and tested on GAN generated images to obtain a GAN-test score. These two networks evaluated the distribution difference between the learned images and the real images, where high GAN-train and GAN-test scores demonstrate that the generated samples are diverse enough and are realistic approximation to the original patient images. However, in our application, the GAN-test would not be stable due to small training sample size (around 300 slices vs. over ten thousands trainable parameters). Instead, we trained our prediction network on the generated images, validated it on the real images, and tested it on unseen patient images. High and similar training, validation, and test accuracies implied that the generated samples learned the real image distribution and were diverse enough to predict the real case. Further investigated in needed to fully evaluate the diversity and quality of ACGAN generated images.

In this work, VGG-19 proposed by the Visual Geometry Group was used to predict the treatment outcome[65]. VGG-19 was designed to classify selected samples of ImageNet to 1000 categories with less than 10 percent error rate and can be used as a feature extractor instead of a classifier. The last layer of the VGG-19 was removed and an image can be passed through the rest of the network to obtain its feature vector. We adapted the VGG-19 network by adding two additional trainable layers. Promising results of 88% and 95% accuracy were achieved for slice-based and patient-based prediction on independent test sets.

There are limitations of this work. The first limitation is that the prediction model is built to predict treatment response for each slice whereas only one single score is available for each patient. The treatment effect score is determined by pathologists as an overall assessment of treatment-induced pathologic necrosis, hyalinization, and fibrosis to that patient. In this work, it is assumed that all tumor imaging slices had the same score, so the data size is sufficient to allow the use of deep learning-based prediction. Although the treatment effect score reflects the radiation treatment

efficacy for the majority slices, it is possible that some tumor regions have different response pattern. Therefore, it is warranted to extend the trial, so we could have enough data to formulate patient-based prediction. Secondly, manually contoured data was used in building the model. There are two main reasons of performing tumor segmentation in this preliminary study: 1) Patients in this study includes sarcoma of the limb, arm, and hip, and manual segmentation could reduce the impact caused by different disease sites; 2) the size of the tumor varies a lot (the longest diameter ranges from as small as 2cm to as large as 15cm), so removing non-tumor part could force the network to focus on the tumor region in the prediction. However, to avoid contour uncertainty and ease the overall workflow, we are extending this model to accommodate inputting original un-contoured images. Auto-segmentation of the tumor region is an alternative plan we would like to explore. Lastly, at the time of compiling this study, other commonly adopted evaluation matrices such as the development of metastases and long-term survival were not available. We would like to test our proposed workflow on predicting those matrices once the data is available.

7.5 **Conclusion**

An ACGAN model was trained to augment longitudinal ADC maps. The trained model provided high-quality realistic images. A VGG-19 network was modified to predict sarcoma patient treatment outcome using data augmented from the ACGAN model. Accuracies of 88% and 95% were achieved on independent test sets for slice-based and patient-based prediction, respectively.

Chapter 8 Conclusion

In summary, in order to achieve the goal of early treatment response prediction using distortion-free diffusion MRI, we focused on the development of distortion-free diffusion sequences and investigation of treatment response prediction using conventional diffusion sequence in parallel. In Chapter 3, a multi-shot DP-TSE sequence was programed for a low-field MRI-guided radiotherapy system for distortion-free diffusion imaging. To resolve the k-space magnitude inconsistency problem in the DP approach, a magnitude stabilizer technique was developed and validated using 2D DP-MS-bSSFP sequence in Chapter 4. The proposed sequence was then extended to 3D in Chapter 5 to meet the high-resolution requirement for treatment planning and treatment adaptation. We then showed promising preliminary results of using longitudinal DWI for treatment response on sarcoma patients. The radiomics approach and deep learning approach were explored in Chapter 6 and 7. In this chapter, technical contributions achieved in this dissertation are summarized. Potential areas that need future studies are discussed.

8.1 Summary of contribution

8.1.1 Low-field diffusion sequence development

MRI is gaining momentum in the radiation oncology field in recent years. Up to date, there are two MRgRT systems commercially available, and several MRgRT systems are under development[9,10,174,175]. One MRgRT system adopted low-field (0.35T) MR design (ViewRay, MRIdian, Mountain View, CA) to minimize the electron return effect. As diffusion MRI is an SNR demanding imaging technique, it becomes particularly challenging to achieve high-quality DWI at low field strength. As shown in Chapter 3, the conventional DW-ssEPI had increased distortion due to poor gradient performance, and the ADC accuracy was compromised because of low SNR.

To enable functional imaging at low-field strength, it is urgent to develop a diffusion sequence that is free from distortion and has good ADC accuracy. In Chapter 3, a multi-shot DP-TSE sequence was programed and optimized on the 0.35T MRgRT system. K-space view-ordering was modified to suppress shot-to-shot k-space inconstancy. The sequence provided desirable geometric and ADC performance on the phantom and patient scans.

8.1.2 Magnitude stabilizer approach for k-space magnitude inconsistency

Diffusion-prepared strategy has the flexibility of optimizing the image acquisition type, contrast, and SNR by choosing different readout module. However, due to the application of the spoiler gradient after the flip back RF pulse, there exist a k-space magnitude inconsistency in the diffusion prepared strategy. As the conjugate signal has been spoiled, it is challenging to resolve the lost signal. This magnitude inconsistency problem has not been explicitly studied and solved in previous literature. In Chapter 4, we proposed a magnitude stabilizer approach. By adding crusher gradients before the flip back RF pulse, and before and after each echo, the malignant k-space magnitude inconsistency problem is converted to benign k-space phase inconsistency, which can consequently be eliminated by many existing phase correction techniques. We demonstrated that high-quality distortion-free images could be obtained from the magnitude stabilized diffusion-prepared bSSFP sequence. Results from diffusion-prepared FLASH sequence were shown to verify the generalizability of the magnitude stabilizer strategy.

8.1.3 3D distortion-free diffusion sequence for radiotherapy applications

Many radiotherapy applications such as tumor and organ delineation required high spatial resolution to accurately characterize the structures. Therefore, 3D imaging is frequently required in radiotherapy studies to achieve isotropic resolution. In Chapter 5, the diffusion-prepared magnitude-stabilized bSSFP sequence was extended to 3D for high-resolution (1.5mm isotropic)

distortion-free imaging at 1.5T. A locally low-rank-based reconstruction workflow was developed to resolve the k-space inconsistency along with the built-in navigator. Desirable quantitative phantom and in vivo geometry and ADC accuracies were obtained. This sequence is ready to be tested on MR-simulator and used for treatment planning and adaptive radiotherapy of the brain.

8.1.4 Early response prediction using longitudinal diffusion MRI

Predicting treatment response early on is a crucial step in achieving personalized medicine based on individualized patient feedback. In chapter 6, we developed a radiomics workflow for early treatment response prediction using longitudinal diffusion MRI on sarcoma patient. Prognostic features were identified, and predictive models were built. The model is ready to be tested on future patients, and could potentially provide valuable information about the efficacy of the radiotherapy treatment. Another contribution of this work is that we investigated and compared the predictive power using different imaging time points and with or without delta radiomics. Since it is impractical to acquire images after every fraction of RT, identifying the best imaging time point for response prediction is critical for a successful clinical trial. Therefore, the results from this work provide a reference for future longitudinal study of sarcoma patients.

8.1.5 Deep learning-based data augmentation and response prediction

Limited dataset is a common problem in many medical applications. Simple data augmentation techniques such as image translation and rotation have been used in many previous studies to increase the sample size. However, there is little information gain from those data augmentation methods. To overcome this, we applied a novel ACGAN network to learn the underlying pattern of the image in Chapter 7. A large number of realistic images with a wide range of shape and signal intensity were obtained to aid the prediction model. A state-of-art VGG-19 network was modified to predict the treatment response. Promising prediction results were

obtained on individual test sets. We believe the data augmentation and prediction networks workflow we implemented could be applied to many other medical applications for improved performance.

8.2 Future outlook

8.2.1 Patient validation of the DP-MS approach

In this work, the proposed DP-MS sequence was evaluated on phantom and healthy volunteers. Geometric fidelity and ADC accuracy between the conventional DW-ssEPI and DP-MS were compared. Although promising results were obtained on phantom and volunteer studies, no patient study was performed at this time due to some clinical difficulties. To evaluate the clinical benefits of the DP-MS sequence, it is necessary to validate the sequence on brain tumor patients to fully study its capability for tumor detection, tumor and organ delineation, and response assessment. In addition, we would like to compare the dosimetric difference of treatment plans developed using DW-ssEPI images and DP-MS images. Therefore, future patient study is warranted to validate and translate the sequence to the clinic.

8.2.2 Extending the DP-MS approach on other organs

Currently, the DP-MS sequence was only validated on the brain, as the brain is one of the major organs that diffusion imaging is applied on. To enable functional-based treatment planning and treatment adaptation for non-brain cancer patients, it is necessary to optimize the sequence and extend it to other body parts such as rectum, prostate, and liver. There are two main challenges associated with non-brain diffusion imaging. The first one is reduced SNR due to lower T2 relaxation time of body tissue. For example, at 1.5T, T2 relaxation time of the gray matter is about 90ms, whereas only 40ms for the liver. Therefore, it could be difficult to get enough SNR given

the half of signal loss in the DP-MS approach. One potential solution is to design a 3D navigator to achieve true 3D imaging. This could boost the SNR significantly. Another challenge for body diffusion imaging is motion, especially respiratory motion. Although k-space navigator is acquired to estimate the phase inconsistency caused by motion and system imperfection, it is difficult to correct the inconsistency when motion is strong. To solve this, navigator gated acquisition could be implemented to minimize motion artifact.

8.2.3 **Long term survival perdition**

In the two treatment response studies, the sarcoma patients were prospectively recruited from a clinical trial. Therefore, other treatment response index, such as long-term survival and distance metastasis, is not available at this time. Those indexes are the most important indicator of the successfulness of the treatment. In the two studies, we used the treatment effect score, which is a pathological endpoint mainly based on tumor necrosis, as an immediate surrogate of patient response to radiotherapy. Although there are studies showing the correlation between tumor necrosis and long-term survival, the finding is somewhat controversial. Therefore, it is necessary to modify and re-train the model once long-term survival and distance metastasis data is available.

8.2.4 **Model validation on an external cohort**

One limitation of the early treatment response prediction work is the small patient number. Due to the rareness of sarcoma patient, and difficulty of conducting the prospective longitudinal study, only 30 patients had complete data set over the two years of the study. In the radiomics approach, no patient was held out as an independent test set, leaving the results prone to overfitting. In the deep learning study, although data augmentation was performed, and 4 patients were held out as an independent test set in each round of testing, it is unknown whether the model could still maintain desirable performance on the external patient cohort. Therefore, it is important to acquire

patient data on an extended cohort to further improve and validate the performance of the radiomics and deep learning models.

8.2.5 **Treatment response prediction using distortion-free DWI**

As the title alluded, our ultimate goal is to use longitudinal distortion-free DWI for early treatment response prediction. To achieve this in an efficient manner, distortion-free sequence development and response prediction using the conventional DWI were investigated in parallel (chapter 3-5 and chapter 6-7 respectively). However, tumor shape changes could be an important indicator of patient outcome, whereas distorted DWI cannot faithfully reflect tumor shape, as spatial distortion is affected by patient geometric variation, patient daily setup, and system stability. Therefore, it is necessary to combine the results from the two paths together once the DP-MS sequence is verified on patient. It would also be interesting to compare the prediction performance using DW-ssEPI with performance using DP-MS to verify the value of distortion-free diffusion technique.

REFERENCE

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019;69(1):7-34. doi:10.3322/caac.21551

2. Kupelian P, Sonke J-J. Magnetic Resonance–Guided Adaptive Radiotherapy: A Solution to the Future. *Semin Radiat Oncol*. 2014;24(3):227-232. doi:10.1016/j.semradonc.2014.02.013

3. Mannelli L, Nougaret S, Vargas HA, Do RKG. Advances in diffusion-weighted imaging. *Radiol Clin North Am*. 2015;53(3):569-581. doi:10.1016/j.rcl.2015.01.002

4. Li SP, Padhani AR. Tumor response assessments with diffusion and perfusion MRI. *J Magn Reson Imaging JMRI*. 2012;35(4):745-763. doi:10.1002/jmri.22838

5. Bhatnagar P, Subesinghe M, Patel C, Prestwich R, Scarsbrook AF. Functional Imaging for Radiation Treatment Planning, Response Assessment, and Adaptive Therapy in Head and Neck Cancer. *RadioGraphics*. 2013;33(7):1909-1929. doi:10.1148/rg.337125163

6. Tsien C, Cao Y, Chenevert T. Clinical applications for diffusion magnetic resonance imaging in radiotherapy. *Semin Radiat Oncol*. 2014;24(3):218-226. doi:10.1016/j.semradonc.2014.02.004

7. Tomura N, Narita K, Izumi J, et al. Diffusion changes in a tumor and peritumoral tissue after stereotactic irradiation for brain tumors: possible prediction of treatment response. *J Comput Assist Tomogr*. 2006;30(3):496-500.

8. Malayeri AA, El Khouli RH, Zaheer A, et al. Principles and applications of diffusion-weighted imaging in cancer detection, staging, and treatment follow-up. *Radiogr Rev Publ Radiol Soc N Am Inc*. 2011;31(6):1773-1791. doi:10.1148/rg.316115515

9. Mutic S, Dempsey JF. The ViewRay System: Magnetic Resonance–Guided and Controlled Radiotherapy. *Semin Radiat Oncol*. 2014;24(3):196-199. doi:10.1016/j.semradonc.2014.02.008

10. Raaymakers BW, Jürgenliemk-Schulz IM, Bol GH, et al. First patients treated with a 1.5 T MRI-Linac: clinical proof of concept of a high-precision, high-field MRI guided radiotherapy treatment. *Phys Med Biol*. 2017;62(23):L41-L50. doi:10.1088/1361-6560/aa9517

11. Bernstein MA, King KF, Zhou XJ. *Handbook of MRI Pulse Sequences*. Elsevier; 2004.

12. Zaharchuk G, Saritas EU, Andre JB, et al. Reduced field-of-view diffusion imaging of the human spinal cord: comparison with conventional single-shot echo-planar imaging. *AJNR Am J Neuroradiol*. 2011;32(5):813-820. doi:10.3174/ajnr.A2418

13. Porter DA, Heidemann RM. High resolution diffusion-weighted imaging using readout-segmented echo-planar imaging, parallel imaging and a two-dimensional navigator-based reacquisition. *Magn Reson Med*. 2009;62(2):468-475. doi:10.1002/mrm.22024

14. Coindre JM, Terrier P, Guillou L, et al. Predictive value of grade for metastasis development in the main histologic types of adult soft tissue sarcomas: a study of 1240 patients from the French Federation of Cancer Centers Sarcoma Group. *Cancer*. 2001;91(10):1914-1926.

15. Key Statistics for Soft Tissue Sarcomas. https://www.cancer.org/cancer/soft-tissue-sarcoma/about/key-statistics.html. Accessed July 30, 2019.

16. Billingsley KG, Burt ME, Jara E, et al. Pulmonary metastases from soft tissue sarcoma: analysis of patterns of diseases and postmetastasis survival. *Ann Surg*. 1999;229(5):602-610; discussion 610-612.

17. Weitz J, Antonescu CR, Brennan MF. Localized extremity soft tissue sarcoma: improved knowledge with unchanged survival over time. *J Clin Oncol Off J Am Soc Clin Oncol*. 2003;21(14):2719-2725. doi:10.1200/JCO.2003.02.026

18. Zagars GK, Ballo MT, Pisters PWT, et al. Prognostic factors for patients with localized soft-tissue sarcoma treated with conservation surgery and radiation therapy: an analysis of 1225 patients. *Cancer*. 2003;97(10):2530-2543. doi:10.1002/cncr.11365

19. Chibon F, Lagarde P, Salas S, et al. Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat Med*. 2010;16(7):781-787. doi:10.1038/nm.2174

20. Coindre J-M. Grading of soft tissue sarcomas: review and update. *Arch Pathol Lab Med*. 2006;130(10):1448-1453. doi:10.1043/1543-2165(2006)130[1448:GOSTSR]2.0.CO;2

21. Dudeck O, Zeile M, Pink D, et al. Diffusion-weighted magnetic resonance imaging allows monitoring of anticancer treatment effects in patients with soft-tissue sarcomas. *J Magn Reson Imaging JMRI*. 2008;27(5):1109-1113. doi:10.1002/jmri.21358

22. Schnapauff D, Zeile M, Niederhagen MB, et al. Diffusion-weighted echo-planar magnetic resonance imaging for the assessment of tumor cellularity in patients with soft-tissue sarcomas. *J Magn Reson Imaging*. 2009;29(6):1355-1359. doi:10.1002/jmri.21755

23. Uhl M, Saueressig U, van Buiren M, et al. Osteosarcoma: preliminary results of in vivo assessment of tumor necrosis after chemotherapy with diffusion- and perfusion-weighted magnetic resonance imaging. *Invest Radiol*. 2006;41(8):618-623. doi:10.1097/01.rli.0000225398.17315.68

24. Oka K, Yakushiji T, Sato H, Hirai T, Yamashita Y, Mizuta H. The value of diffusion-weighted imaging for monitoring the chemotherapeutic response of osteosarcoma: a comparison between average apparent diffusion coefficient and minimum apparent

diffusion coefficient. *Skeletal Radiol*. 2010;39(2):141-146. doi:10.1007/s00256-009-0830-7

25. Lövblad KO, Laubach HJ, Baird AE, et al. Clinical experience with diffusion-weighted MR in patients with acute stroke. *AJNR Am J Neuroradiol*. 1998;19(6):1061-1066.

26. Le Bihan D. Apparent Diffusion Coefficient and Beyond: What Diffusion MR Imaging Can Tell Us about Tissue Structure. *Radiology*. 2013;268(2):318-322. doi:10.1148/radiol.13130420

27. Jezzard P, Balaban RS. Correction for geometric distortion in echo planar images from B0 field variations. *Magn Reson Med*. 1995;34(1):65-73.

28. Chen NK, Wyrwicz AM. Correction for EPI distortions using multi-echo gradient-echo imaging. *Magn Reson Med*. 1999;41(6):1206-1213.

29. Chen NK, Wyrwicz AM. Optimized distortion correction technique for echo planar imaging. *Magn Reson Med*. 2001;45(3):525-528.

30. Zeng H, Gatenby JC, Zhao Y, Gore JC. New approach for correcting distortions in echo planar imaging. *Magn Reson Med*. 2004;52(6):1373-1378. doi:10.1002/mrm.20297

31. Robson MD, Gore JC, Constable RT. Measurement of the point spread function in MRI using constant time imaging. *Magn Reson Med*. 1997;38(5):733-740.

32. Reber PJ, Wong EC, Buxton RB, Frank LR. Correction of off resonance-related distortion in echo-planar imaging using EPI-based field maps. *Magn Reson Med*. 1998;39(2):328-330.

33. Wan X, Gullberg GT, Parker DL, Zeng GL. Reduction of geometric and intensity distortions in echo-planar imaging using a multireference scan. *Magn Reson Med*. 1997;37(6):932-942.

34. Morgan PS, Bowtell RW, McIntyre DJO, Worthington BS. Correction of spatial distortion in EPI due to inhomogeneous static magnetic fields using the reversed gradient method. *J Magn Reson Imaging JMRI*. 2004;19(4):499-507. doi:10.1002/jmri.20032

35. Xiang Q-S, Ye FQ. Correction for geometric distortion and N/2 ghosting in EPI by phase labeling for additional coordinate encoding (PLACE). *Magn Reson Med*. 2007;57(4):731-741. doi:10.1002/mrm.21187

36. Chen N, Oshio K, Panych LP. Application of k-space energy spectrum analysis to susceptibility field mapping and distortion correction in gradient-echo EPI. *NeuroImage*. 2006;31(2):609-622. doi:10.1016/j.neuroimage.2005.12.022

37. Jeong E-K, Kim S-E, Parker DL. High-resolution diffusion-weighted 3D MRI, using diffusion-weighted driven-equilibrium (DW-DE) and multishot segmented 3D-SSFP without navigator echoes. *Magn Reson Med*. 2003;50(4):821-829. doi:10.1002/mrm.10593

38. Nguyen C, Fan Z, Sharif B, et al. In vivo three-dimensional high resolution cardiac diffusion-weighted MRI: A motion compensated diffusion-prepared balanced steady-state free precession approach. *Magn Reson Med*. 2014;72(5):1257-1267. doi:10.1002/mrm.25038

39. Miller KL, Pauly JM. Nonlinear phase correction for navigated diffusion imaging. *Magn Reson Med*. 2003;50(2):343-353. doi:10.1002/mrm.10531

40. Ordidge RJ, Helpern JA, Qing ZX, Knight RA, Nagesh V. Correction of motional artifacts in diffusion-weighted MR images using navigator echoes. *Magn Reson Imaging*. 1994;12(3):455-460.

41. Butts K, de Crespigny A, Pauly JM, Moseley M. Diffusion-weighted interleaved echo-planar imaging with a pair of orthogonal navigator echoes. *Magn Reson Med*. 1996;35(5):763-770. doi:10.1002/mrm.1910350518

42. Jeong H-K, Gore JC, Anderson AW. High-resolution human diffusion tensor imaging using 2-D navigated multishot SENSE EPI at 7 T. *Magn Reson Med*. 2013;69(3):793-802. doi:10.1002/mrm.24320

43. Chen N-K, Guidon A, Chang H-C, Song AW. A robust multi-shot scan strategy for high-resolution diffusion weighted MRI enabled by multiplexed sensitivity-encoding (MUSE). *NeuroImage*. 2013;72:41-47. doi:10.1016/j.neuroimage.2013.01.038

44. Nguyen C, Fan Z, Xie Y, et al. In vivo diffusion-tensor MRI of the human heart on a 3 tesla clinical scanner: An optimized second order (M2) motion compensated diffusion-preparation approach. *Magn Reson Med*. 2016;76(5):1354-1363. doi:10.1002/mrm.26380

45. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2015;278(2):563-577. doi:10.1148/radiol.2015151169

46. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res*. 2017;77(21):e104-e107. doi:10.1158/0008-5472.CAN-17-0339

47. Zhang L, Fried DV, Fave XJ, Hunter LA, Yang J, Court LE. IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics. *Med Phys*. 2015;42(3):1341-1353. doi:10.1118/1.4908210

48. Brown AM, Nagala S, McLean MA, et al. Multi-institutional validation of a novel textural analysis tool for preoperative stratification of suspected thyroid tumors on diffusion-weighted MRI. *Magn Reson Med*. 2016;75(4):1708-1716. doi:10.1002/mrm.25743

49. Nioche C, Orlhac F, Boughdad S, et al. LIFEx: A Freeware for Radiomic Feature Calculation in Multimodality Imaging to Accelerate Advances in the Characterization of Tumor Heterogeneity. *Cancer Res*. 2018;78(16):4786-4789. doi:10.1158/0008-5472.CAN-18-0125

50. Vallières M, Freeman CR, Skamene SR, El Naqa I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol*. 2015;60(14):5471-5496. doi:10.1088/0031-9155/60/14/5471

51. Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. *ArXiv161207003 Cs Eess*. December 2016. http://arxiv.org/abs/1612.07003. Accessed August 21, 2019.

52. Rizzo S, Botta F, Raimondi S, et al. Radiomics: the facts and the challenges of image analysis. *Eur Radiol Exp*. 2018;2. doi:10.1186/s41747-018-0068-z

53. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88. doi:10.1016/j.media.2017.07.005

54. Shen D, Wu G, Suk H-I. Deep Learning in Medical Image Analysis. *Annu Rev Biomed Eng*. 2017;19(1):221-248. doi:10.1146/annurev-bioeng-071516-044442

55. Xu Y, Hosny A, Zeleznik R, et al. Deep Learning Predicts Lung Cancer Treatment Response from Serial Medical Imaging. *Clin Cancer Res*. April 2019. doi:10.1158/1078-0432.CCR-18-2495

56. Bibault J-E, Giraud P, Housset M, et al. Deep Learning and Radiomics predict complete response after neo-adjuvant chemoradiation for locally advanced rectal cancer. *Sci Rep*. 2018;8(1):12611. doi:10.1038/s41598-018-30657-6

57. Cha KH, Hadjiiski L, Chan H-P, et al. Bladder Cancer Treatment Response Assessment in CT using Radiomics with Deep-Learning. *Sci Rep*. 2017;7(1):8738. doi:10.1038/s41598-017-09315-w

58. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Nets. In: Ghahramani Z, Welling M, Cortes C, Lawrence ND, Weinberger KQ, eds. *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc.; 2014:2672–2680. http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf. Accessed May 6, 2019.

59. Xue Y, Xu T, Zhang H, Long R, Huang X. SegAN: Adversarial Network with Multi-scale L_1 Loss for Medical Image Segmentation. *Neuroinformatics*. 2018;16(3-4):383-392. doi:10.1007/s12021-018-9377-x

60. Nie D, Trullo R, Petitjean C, Ruan S, Shen D. Medical Image Synthesis with Context-Aware Generative Adversarial Networks. *ArXiv161205362 Cs*. December 2016. http://arxiv.org/abs/1612.05362. Accessed May 6, 2019.

61. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*. 2018;321:321-331. doi:10.1016/j.neucom.2018.09.013

62. Ali Madani, Mehdi Moradi, Alexandros Karargyris, Tanveer Syeda-Mahmood. Chest x-ray generation and data augmentation for cardiovascular abnormality classification. In: Vol 10574. ; 2018. https://doi.org/10.1117/12.2293971.

63. Mirza M, Osindero S. Conditional Generative Adversarial Nets. *ArXiv14111784 Cs Stat*. November 2014. http://arxiv.org/abs/1411.1784. Accessed August 26, 2019.

64. Odena A, Olah C, Shlens J. Conditional Image Synthesis With Auxiliary Classifier GANs. *ArXiv161009585 Cs Stat*. October 2016. http://arxiv.org/abs/1610.09585. Accessed May 6, 2019.

65. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *ArXiv14091556 Cs*. September 2014. http://arxiv.org/abs/1409.1556. Accessed May 6, 2019.

66. Thoeny HC, Ross BD. Predicting and Monitoring Cancer Treatment Response with DW-MRI. *J Magn Reson Imaging JMRI*. 2010;32(1):2-16. doi:10.1002/jmri.22167

67. Sharma U, Danishad KKA, Seenu V, Jagannathan NR. Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy. *NMR Biomed*. 2009;22(1):104-113. doi:10.1002/nbm.1245

68. Kuang F, Yan Z, Wang J, Rao Z. The value of diffusion-weighted MRI to evaluate the response to radiochemotherapy for cervical cancer. *Magn Reson Imaging*. 2014;32(4):342-349. doi:10.1016/j.mri.2013.12.007

69. Choi SP, Park KN, Park HK, et al. Diffusion-weighted magnetic resonance imaging for predicting the clinical outcome of comatose survivors after cardiac arrest: a cohort study. *Crit Care Lond Engl*. 2010;14(1):R17. doi:10.1186/cc8874

70. King AD, Chow K-K, Yu K-H, et al. Head and neck squamous cell carcinoma: diagnostic performance of diffusion-weighted MR imaging for the prediction of treatment response. *Radiology*. 2013;266(2):531-538. doi:10.1148/radiol.12120167

71. Liu L, Wu N, Ouyang H, Dai J-R, Wang W-H. Diffusion-weighted MRI in early assessment of tumour response to radiotherapy in high-risk prostate cancer. *Br J Radiol*. 2014;87(1043):20140359. doi:10.1259/bjr.20140359

72. Yang Y, Cao M, Sheng K, et al. Longitudinal diffusion MRI for treatment response assessment: Preliminary experience using an MRI-guided tri-cobalt 60 radiotherapy system. *Med Phys*. 2016;43(3):1369-1373. doi:10.1118/1.4942381

73. Gudbjartsson H, Maier SE, Mulkern RV, Mórocz IÁ, Patz S, Jolesz FA. Line scan diffusion imaging. *Magn Reson Med*. 1996;36(4):509-519. doi:10.1002/mrm.1910360403

74. Chang H, Fitzpatrick JM. A technique for accurate magnetic resonance imaging in the presence of field inhomogeneities. *IEEE Trans Med Imaging*. 1992;11(3):319-329. doi:10.1109/42.158935

75. Jeong E-K, Kim S-E, Parker DL. High-resolution diffusion-weighted 3D MRI, using diffusion-weighted driven-equilibrium (DW-DE) and multishot segmented 3D-SSFP without navigator echoes. *Magn Reson Med*. 2003;50(4):821-829. doi:10.1002/mrm.10593

76. Xie Y, Yu W, Fan Z, et al. High resolution 3D diffusion cardiovascular magnetic resonance of carotid vessel wall to detect lipid core without contrast media. *J Cardiovasc Magn Reson*. 2014;16:67. doi:10.1186/s12968-014-0067-z

77. Reese TG, Heid O, Weisskoff RM, Wedeen VJ. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magn Reson Med*. 2003;49(1):177-182. doi:10.1002/mrm.10308

78. Robson MD, Anderson AW, Gore JC. Diffusion-weighted multiple shot echo planar imaging of humans without navigation. *Magn Reson Med*. 1997;38(1):82-88. doi:10.1002/mrm.1910380113

79. Bammer R, Holdsworth SJ, Veldhuis WB, Skare ST. New methods in Diffusion Weighted and Diffusion Tensor Imaging. *Magn Reson Imaging Clin N Am*. 2009;17(2):175-204. doi:10.1016/j.mric.2009.01.011

80. Dietrich O, Heiland S, Sartor K. Noise correction for the exact determination of apparent diffusion coefficients at low SNR. *Magn Reson Med*. 2001;45(3):448-453.

81. Maurer CR, Fitzpatrick JM, Wang MY, Galloway RL, Maciunas RJ, Allen GS. Registration of head volume images using implantable fiducial markers. *IEEE Trans Med Imaging*. 1997;16(4):447-462. doi:10.1109/42.611354

82. Sener RN. Diffusion MRI: apparent diffusion coefficient (ADC) values in the normal brain and a classification of brain disorders based on ADC values. *Comput Med Imaging Graph*. 2001;25(4):299-326. doi:10.1016/S0895-6111(00)00083-5

83. Raaijmakers AJE, Raaymakers BW, Lagendijk JJW. Magnetic-field-induced dose effects in MR-guided radiotherapy systems: dependence on the magnetic field strength. *Phys Med Biol*. 2008;53(4):909. doi:10.1088/0031-9155/53/4/006

84. Iima M, Yano K, Kataoka M, et al. Quantitative non-Gaussian diffusion and intravoxel incoherent motion magnetic resonance imaging: differentiation of malignant and benign breast lesions. *Invest Radiol*. 2015;50(4):205-211. doi:10.1097/RLI.0000000000000094

85. Baron CA, Lebel RM, Wilman AH, Beaulieu C. The effect of concomitant gradient fields on diffusion tensor imaging. *Magn Reson Med*. 2012;68(4):1190-1201. doi:10.1002/mrm.24120

86. Aliotta E, Wu HH, Ennis DB. Convex optimized diffusion encoding (CODE) gradient waveforms for minimum echo time and bulk motion–compensated diffusion-weighted MRI. *Magn Reson Med*. 2017;77(2):717-729. doi:10.1002/mrm.26166

87. Pipe JG, Farthing VG, Forbes KP. Multishot diffusion-weighted FSE using PROPELLER MRI. *Magn Reson Med*. 2002;47(1):42-52.

88. Frank LR, Jung Y, Inati S, Tyszka JM, Wong EC. High efficiency, low distortion 3D diffusion tensor imaging with variable density spiral fast spin echoes (3D DW VDS RARE). *NeuroImage*. 2010;49(2):1510-1523. doi:10.1016/j.neuroimage.2009.09.010

89. Zhang Z, Zhang B, Li M, et al. Multishot cartesian turbo spin-echo diffusion imaging using iterative POCSMUSE Reconstruction. *J Magn Reson Imaging*. October 2016:n/a-n/a. doi:10.1002/jmri.25522

90. Alsop DC. Phase insensitive preparation of single-shot RARE: application to diffusion imaging in humans. *Magn Reson Med*. 1997;38(4):527-533.

91. Butts K, Riederer SJ, Ehman RL, Thompson RM, Jack CR. Interleaved echo planar imaging on a standard MRI system. *Magn Reson Med*. 1994;31(1):67-72. doi:10.1002/mrm.1910310111

92. Deng J, Miller FH, Salem R, Omary RA, Larson AC. Multishot diffusion-weighted PROPELLER magnetic resonance imaging of the abdomen. *Invest Radiol*. 2006;41(10):769-775. doi:10.1097/01.rli.0000236808.84746.95

93. Holdsworth SJ, Skare S, Newbould RD, Guzmann R, Blevins NH, Bammer R. Readout-segmented EPI for rapid high resolution diffusion imaging at 3T. *Eur J Radiol*. 2008;65(1):36-46. doi:10.1016/j.ejrad.2007.09.016

94. Skare S, Andersson JL. On the effects of gating in diffusion imaging of the brain using single shot EPI. *Magn Reson Imaging*. 2001;19(8):1125-1128.

95. O'Halloran RL, Holdsworth S, Aksoy M, Bammer R. Model for the correction of motion-induced phase errors in multishot diffusion-weighted-MRI of the head: are cardiac-motion-induced phase errors reproducible from beat-to-beat? *Magn Reson Med*. 2012;68(2):430-440. doi:10.1002/mrm.23245

96. Liu C, Bammer R, Kim D-H, Moseley ME. Self-navigated interleaved spiral (SNAILS): application to high-resolution diffusion tensor imaging. *Magn Reson Med*. 2004;52(6):1388-1396. doi:10.1002/mrm.20288

97. Li T-Q, Kim D-H, Moseley ME. High-resolution diffusion-weighted imaging with interleaved variable-density spiral acquisitions. *J Magn Reson Imaging JMRI*. 2005;21(4):468-475. doi:10.1002/jmri.20287

98. Truong T-K, Guidon A. High-Resolution Multi-Shot Spiral Diffusion Tensor Imaging with Inherent Correction of Motion-Induced Phase Errors. *Magn Reson Med Off J Soc Magn Reson Med Soc Magn Reson Med*. 2014;71(2):790-796. doi:10.1002/mrm.24709

99. Chu M-L, Chang H-C, Chung H-W, Truong T-K, Bashir MR, Chen N. POCS-based reconstruction of multiplexed sensitivity encoded MRI (POCSMUSE): A general algorithm for reducing motion-related artifacts. *Magn Reson Med*. 2015;74(5):1336-1348. doi:10.1002/mrm.25527

100. Zhang Z, Huang F, Ma X, Xie S, Guo H. Self-feeding MUSE: a robust method for high resolution diffusion imaging using interleaved EPI. *NeuroImage*. 2015;105:552-560. doi:10.1016/j.neuroimage.2014.10.022

101. Mani M, Jacob M, Kelley D, Magnotta V. Multi-shot sensitivity-encoded diffusion data recovery using structured low-rank matrix completion (MUSSELS). *Magn Reson Med*. 2017;78(2):494-507. doi:10.1002/mrm.26382

102. Bammer R, Augustin M, Prokesch RW, Stollberger R, Fazekas F. Diffusion-weighted imaging of the spinal cord: interleaved echo-planar imaging is superior to fast spin-echo. *J Magn Reson Imaging JMRI*. 2002;15(4):364-373.

103. Cervantes B, Van AT, Weidlich D, et al. Isotropic resolution diffusion tensor imaging of lumbosacral and sciatic nerves using a phase-corrected diffusion-prepared 3D turbo spin echo. *Magn Reson Med*. 2018;80(2):609-618. doi:10.1002/mrm.27072

104. Zhang Q, Coolen BF, Versluis MJ, Strijkers GJ, Nederveen AJ. Diffusion-prepared stimulated-echo turbo spin echo (DPsti-TSE): An eddy current-insensitive sequence for three-dimensional high-resolution and undistorted diffusion-weighted imaging. *NMR Biomed*. 2017;30(7):n/a-n/a. doi:10.1002/nbm.3719

105. Van AT, Cervantes B, Kooijman H, Karampinos DC. Analysis of phase error effects in multishot diffusion-prepared turbo spin echo imaging. *Quant Imaging Med Surg*. 2017;7(2):238-250. doi:10.21037/qims.2017.04.01

106. Thomas DL, Pell GS, Lythgoe MF, Gadian DG, Ordidge RJ. A quantitative method for fast diffusion imaging using magnetization-prepared TurboFLASH. *Magn Reson Med*. 1998;39(6):950-960.

107. Stejskal EO, Tanner JE. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J Chem Phys*. 1965;42(1):288-292. doi:10.1063/1.1695690

108. Sinha U, Sinha S. High speed diffusion imaging in the presence of eddy currents. *J Magn Reson Imaging JMRI*. 1996;6(4):657-666.

109. Wansapura JP, Holland SK, Dunn RS, Ball WS. NMR relaxation times in the human brain at 3.0 tesla. *J Magn Reson Imaging JMRI*. 1999;9(4):531-538.

110. Deshpande VS, Chung Y-C, Zhang Q, Shea SM, Li D. Reduction of transient signal oscillations in true-FISP using a linear flip angle series magnetization preparation. *Magn Reson Med*. 2003;49(1):151-157. doi:10.1002/mrm.10337

111. Stöcker T, Kaffanke J, Shah NJ. Whole-brain single-shot STEAM DTI at 4 Tesla utilizing transverse coherences for enhanced SNR. *Magn Reson Med*. 2009;61(2):372-380. doi:10.1002/mrm.21853

112. Zhang Z, Zhang B, Li M, et al. Multishot cartesian turbo spin-echo diffusion imaging using iterative POCSMUSE Reconstruction. *J Magn Reson Imaging*. 2017;46(1):167-174. doi:10.1002/jmri.25522

113. Tyagi N, Fontenla S, Zelefsky M, et al. Clinical workflow for MR-only simulation and planning in prostate. *Radiat Oncol*. 2017;12(1):119. doi:10.1186/s13014-017-0854-4

114. Benedict SH, Yenice KM, Followill D, et al. Stereotactic body radiation therapy: The report of AAPM Task Group 101. *Med Phys*. 2010;37(8):4078-4101. doi:10.1118/1.3438081

115. Rudra S, Jiang N, Rosenberg SA, et al. Using adaptive magnetic resonance image-guided radiation therapy for treatment of inoperable pancreatic cancer. *Cancer Med*. 2019;8(5):2123-2132. doi:10.1002/cam4.2100

116. Leibfarth S, Winter RM, Lyng H, Zips D, Thorwarth D. Potentials and challenges of diffusion-weighted magnetic resonance imaging in radiotherapy. *Clin Transl Radiat Oncol*. 2018;13:29-37. doi:10.1016/j.ctro.2018.09.002

117. Sugahara T, Korogi Y, Kochi M, et al. Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas. *J Magn Reson Imaging JMRI*. 1999;9(1):53-60. doi:10.1002/(sici)1522-2586(199901)9:1<53::aid-jmri7>3.0.co;2-2

118. Miao H, Fukatsu H, Ishigaki T. Prostate cancer detection with 3-T MRI: Comparison of diffusion-weighted and T2-weighted imaging. *Eur J Radiol*. 2007;61(2):297-302. doi:10.1016/j.ejrad.2006.10.002

119. Kim C, Park B, Han J, Kang T, Lee H. Diffusion-Weighted Imaging of the Prostate at 3 T for Differentiation of Malignant and Benign Tissue in Transition and Peripheral Zones: Preliminary Results. *J Comput Assist Tomogr*. 2007;31(3):449-454. doi:10.1097/01.rct.0000243456.00437.59

120. Tamada T, Sone T, Jo Y, et al. Apparent diffusion coefficient values in peripheral and transition zones of the prostate: comparison between normal and malignant prostatic tissues and correlation with histologic grade. *J Magn Reson Imaging JMRI*. 2008;28(3):720-726. doi:10.1002/jmri.21503

121. Chenevert TL, Stegman LD, Taylor JM, et al. Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors. *J Natl Cancer Inst*. 2000;92(24):2029-2036. doi:10.1093/jnci/92.24.2029

122. Galbán CJ, Mukherji SK, Chenevert TL, et al. A Feasibility Study of Parametric Response Map Analysis of Diffusion-Weighted Magnetic Resonance Imaging Scans of Head and Neck Cancer Patients for Providing Early Detection of Therapeutic Efficacy. *Transl Oncol*. 2009;2(3):184-190.

123. Koh D-M, Scurr E, Collins D, et al. Predicting response of colorectal hepatic metastasis: value of pretreatment apparent diffusion coefficients. *AJR Am J Roentgenol*. 2007;188(4):1001-1008. doi:10.2214/AJR.06.0601

124. Niwa T, Ueno M, Ohkawa S, et al. Advanced pancreatic cancer: the use of the apparent diffusion coefficient to predict response to chemotherapy. *Br J Radiol*. 2009;82(973):28-34. doi:10.1259/bjr/43911400

125. Reischauer C, Froehlich JM, Koh D-M, et al. Bone Metastases from Prostate Cancer: Assessing Treatment Response by Using Diffusion-weighted Imaging and Functional Diffusion Maps—Initial Observations. *Radiology*. 2010;257(2):523-531. doi:10.1148/radiol.10092469

126. Gao Y, Han F, Zhou Z, et al. Multishot diffusion-prepared magnitude-stabilized balanced steady-state free precession sequence for distortion-free diffusion imaging. *Magn Reson Med*. 2019;81(4):2374-2384. doi:10.1002/mrm.27565

127. Engström M, Skare S. Diffusion-weighted 3D multislab echo planar imaging for high signal-to-noise ratio efficiency and isotropic image resolution. *Magn Reson Med*. 2013;70(6):1507-1514. doi:10.1002/mrm.24594

128. Chang H-C, Hui ES, Chiu P-W, Liu X, Chen N-K. Phase correction for three-dimensional (3D) diffusion-weighted interleaved EPI using 3D multiplexed sensitivity encoding and reconstruction (3D-MUSER). *Magn Reson Med*. 2018;79(5):2702-2712. doi:10.1002/mrm.26944

129. Chang H-C, Sundman M, Petit L, et al. Human brain diffusion tensor imaging at submillimeter isotropic resolution on a 3Tesla clinical MRI scanner. *NeuroImage*. 2015;118:667-675. doi:10.1016/j.neuroimage.2015.06.016

130. Trzasko J, Manduca A, Borisch E. Local versus global low-rank promotion in dynamic MRI series reconstruction. *Proc Int Symp Magn Reson Med*. January 2011.

131. Zhang T, Pauly JM, Levesque IR. Accelerating parameter mapping with a locally low rank constraint. *Magn Reson Med*. 2015;73(2):655-661. doi:10.1002/mrm.25161

132. Uecker M, Lai P, Murphy MJ, et al. ESPIRiT-an eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA. *Magn Reson Med*. 2014;71(3):990-1001. doi:10.1002/mrm.24751

133. Liao C, Chen Y, Cao X, et al. Efficient parallel reconstruction for high resolution multishot spiral diffusion data with low rank constraint. *Magn Reson Med*. 2017;77(3):1359-1366. doi:10.1002/mrm.26199

134. Frost R, Miller KL, Tijssen RHN, Porter DA, Jezzard P. 3D multi-slab diffusion-weighted readout-segmented EPI with real-time cardiac-reordered K-space acquisition. *Magn Reson Med*. 2014;72(6):1565-1579. doi:10.1002/mrm.25062

135. Wu W, Poser BA, Douaud G, et al. High-resolution diffusion MRI at 7T using a three-dimensional multi-slab acquisition. *NeuroImage*. 2016;143:1-14. doi:10.1016/j.neuroimage.2016.08.054

136. Zhang Q, Coolen BF, Nederveen AJ, Strijkers GJ. Three-dimensional diffusion imaging with spiral encoded navigators from stimulated echoes (3D-DISPENSE). *Magn Reson Med*. 2019;81(2):1052-1065. doi:10.1002/mrm.27470

137. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018;68(1):7-30. doi:10.3322/caac.21442

138. Guillou L, Coindre JM, Bonichon F, et al. Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. *J Clin Oncol Off J Am Soc Clin Oncol*. 1997;15(1):350-362. doi:10.1200/JCO.1997.15.1.350

139. Levine EA. Prognostic factors in soft tissue sarcoma. *Semin Surg Oncol*. 1999;17(1):23-32. doi:10.1002/(SICI)1098-2388(199907/08)17:1<23::AID-SSU4>3.0.CO;2-R

140. Trojani M, Contesso G, Coindre JM, et al. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *Int J Cancer*. 1984;33(1):37-42.

141. Costa J, Wesley RA, Glatstein E, Rosenberg SA. The grading of soft tissue sarcomas results of a clinicohistopathologic correlation in a series of 163 cases. *Cancer*. 1984;53(3):530-541. doi:10.1002/1097-0142(19840201)53:3<530::AID-CNCR2820530327>3.0.CO;2-D

142. Menendez LR, Ahlmann ER, Savage K, Cluck M, Fedenko AN. Tumor necrosis has no prognostic value in neoadjuvant chemotherapy for soft tissue sarcoma. *Clin Orthop*. 2007;455:219-224. doi:10.1097/01.blo.0000238864.69486.59

143. Mullen JT, Hornicek FJ, Harmon DC, et al. Prognostic Significance of Treatment-Induced Pathologic Necrosis in Extremity and Truncal Soft Tissue Sarcoma after Neoadjuvant Chemoradiotherapy. *Cancer*. 2014;120(23):3676-3682. doi:10.1002/cncr.28945

144. el-Jabbour JN, Akhtar SS, Kerr GR, et al. Prognostic factors for survival in soft tissue sarcoma. *Br J Cancer*. 1990;62(5):857-861.

145. Hashimoto H, Daimaru Y, Takeshita S, Tsuneyoshi M, Enjoji M. Prognostic significance of histologic parameters of soft tissue sarcomas. *Cancer*. 1992;70(12):2816-2822.

146. van Unnik JA, Coindre JM, Contesso C, et al. Grading of soft tissue sarcomas: experience of the EORTC Soft Tissue and Bone Sarcoma Group. *Eur J Cancer Oxf Engl 1990*. 1993;29A(15):2089-2093.

147. Picci P, Böhling T, Bacci G, et al. Chemotherapy-induced tumor necrosis as a prognostic factor in localized Ewing's sarcoma of the extremities. *J Clin Oncol Off J Am Soc Clin Oncol*. 1997;15(4):1553-1559. doi:10.1200/JCO.1997.15.4.1553

148. Bacci G, Mercuri M, Longhi A, et al. Grade of chemotherapy-induced necrosis as a predictor of local and systemic control in 881 patients with non-metastatic osteosarcoma of the extremities treated with neoadjuvant chemotherapy in a single institution. *Eur J Cancer Oxf Engl 1990*. 2005;41(14):2079-2085. doi:10.1016/j.ejca.2005.03.036

149. Schaefer I-M, Hornick JL, Barysauskas CM, et al. Histologic Appearance After Preoperative Radiation Therapy for Soft Tissue Sarcoma: Assessment of the European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group Response Score. *Int J Radiat Oncol Biol Phys*. 2017;98(2):375-383. doi:10.1016/j.ijrobp.2017.02.087

150. Wang X, Jacobs M, Fayad L. Therapeutic Response in Musculoskeletal Soft Tissue Sarcomas: Evaluation by Magnetic Resonance Imaging. *NMR Biomed*. 2011;24(6):750-763. doi:10.1002/nbm.1731

151. Soldatos T, Ahlawat S, Montgomery E, Chalian M, Jacobs MA, Fayad LM. Multiparametric Assessment of Treatment Response in High-Grade Soft-Tissue Sarcomas with Anatomic and Functional MR Imaging Sequences. *Radiology*. 2016;278(3):831-840. doi:10.1148/radiol.2015142463

152. Costa FM, Ferreira EC, Vianna EM. Diffusion-weighted magnetic resonance imaging for the evaluation of musculoskeletal tumors. *Magn Reson Imaging Clin N Am*. 2011;19(1):159-180. doi:10.1016/j.mric.2010.10.007

153. Subhawong TK, Jacobs MA, Fayad LM. Diffusion-weighted MR imaging for characterizing musculoskeletal lesions. *Radiogr Rev Publ Radiol Soc N Am Inc*. 2014;34(5):1163-1177. doi:10.1148/rg.345140190

154. Roberge D, Skamene T, Nahal A, Turcotte RE, Powell T, Freeman C. Radiological and pathological response following pre-operative radiotherapy for soft-tissue sarcoma. *Radiother Oncol J Eur Soc Ther Radiol Oncol*. 2010;97(3):404-407. doi:10.1016/j.radonc.2010.10.007

155. Yoshida S, Koga F, Kawakami S, et al. Initial experience of diffusion-weighted magnetic resonance imaging to assess therapeutic response to induction chemoradiotherapy against

muscle-invasive bladder cancer. *Urology*. 2010;75(2):387-391. doi:10.1016/j.urology.2009.06.111

156. Noebauer-Huhmann IM, Weber M-A, Lalam RK, et al. Soft Tissue Tumors in Adults: ESSR-Approved Guidelines for Diagnostic Imaging. *Semin Musculoskelet Radiol*. 2015;19(5):475-482. doi:10.1055/s-0035-1569251

157. Yoon MA, Chee CG, Chung HW, et al. Added value of diffusion-weighted imaging to conventional MRI for predicting fascial involvement of soft tissue sarcomas. *Eur Radiol*. October 2018. doi:10.1007/s00330-018-5786-3

158. Rijswijk CSP van, Kunz P, Hogendoorn PCW, Taminiau AHM, Doornbos J, Bloem JL. Diffusion-weighted MRI in the characterization of soft-tissue tumors. *J Magn Reson Imaging*. 2002;15(3):302-307. doi:10.1002/jmri.10061

159. Lee S-Y, Jee W-H, Jung J-Y, et al. Differentiation of malignant from benign soft tissue tumours: use of additive qualitative and quantitative diffusion-weighted MR imaging to standard MR imaging at 3.0 T. *Eur Radiol*. 2016;26(3):743-754. doi:10.1007/s00330-015-3878-x

160. Byun WM, Shin SO, Chang Y, Lee SJ, Finsterbusch J, Frahm J. Diffusion-weighted MR imaging of metastatic disease of the spine: assessment of response to therapy. *AJNR Am J Neuroradiol*. 2002;23(6):906-912.

161. Moffat BA, Chenevert TL, Lawrence TS, et al. Functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. *Proc Natl Acad Sci U S A*. 2005;102(15):5524-5529. doi:10.1073/pnas.0501532102

162. Moffat BA, Hall DE, Stojanovska J, et al. Diffusion imaging for evaluation of tumor therapies in preclinical animal models. *Magn Reson Mater Phys Biol Med*. 2004;17(3):249-259. doi:10.1007/s10334-004-0079-z

163. Mardor Y, Roth Y, Ochershvilli A, et al. Pretreatment prediction of brain tumors' response to radiation therapy using high b-value diffusion-weighted MRI. *Neoplasia N Y N*. 2004;6(2):136-142. doi:10.1593/neo.03349

164. Gao Y, Han F, Zhou Z, et al. Distortion-free diffusion MRI using an MRI-guided Tri-Cobalt 60 radiotherapy system: Sequence verification and preliminary clinical experience. *Med Phys*. 2017;44(10):5357-5366. doi:10.1002/mp.12465

165. Tanner MA, Wong WH. The Calculation of Posterior Distributions by Data Augmentation. *J Am Stat Assoc*. 1987;82(398):528-540. doi:10.1080/01621459.1987.10478458

166. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. *Commun ACM*. 2017;60(6):84–90. doi:10.1145/3065386

167. Picci P, Bacci G, Campanacci M, et al. Histologic evaluation of necrosis in osteosarcoma induced by chemotherapy. Regional mapping of viable and nonviable tumor. *Cancer*. 1985;56(7):1515-1521. doi:10.1002/1097-0142(19851001)56:7<1515::aid-cncr2820560707>3.0.co;2-6

168. Wardelmann E, Haas RL, Bovée JVMG, et al. Evaluation of response after neoadjuvant treatment in soft tissue sarcomas; the European Organization for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group (EORTC-STBSG) recommendations for pathological examination and reporting. *Eur J Cancer Oxf Engl 1990*. 2016;53:84-95. doi:10.1016/j.ejca.2015.09.021

169. Zhang C, Bengio S, Hardt M, Recht B, Vinyals O. Understanding deep learning requires rethinking generalization. *ArXiv161103530 Cs*. November 2016. http://arxiv.org/abs/1611.03530. Accessed July 22, 2019.

170. Salimans T, Goodfellow I, Zaremba W, et al. Improved Techniques for Training GANs. In: Lee DD, Sugiyama M, Luxburg UV, Guyon I, Garnett R, eds. *Advances in Neural Information Processing Systems 29*. Curran Associates, Inc.; 2016:2234–2242. http://papers.nips.cc/paper/6125-improved-techniques-for-training-gans.pdf. Accessed May 6, 2019.

171. Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. In: Guyon I, Luxburg UV, Bengio S, et al., eds. *Advances in Neural Information Processing Systems 30*. Curran Associates, Inc.; 2017:6626–6637. http://papers.nips.cc/paper/7240-gans-trained-by-a-two-time-scale-update-rule-converge-to-a-local-nash-equilibrium.pdf. Accessed May 6, 2019.

172. Borji A. Pros and cons of GAN evaluation measures. *Comput Vis Image Underst*. 2019;179:41-65. doi:10.1016/j.cviu.2018.10.009

173. Shmelkov K, Schmid C, Alahari K. How good is my GAN? *ArXiv180709499 Cs*. July 2018. http://arxiv.org/abs/1807.09499. Accessed May 3, 2019.

174. Keall PJ, Barton M, Crozier S. The Australian Magnetic Resonance Imaging–Linac Program. *Semin Radiat Oncol*. 2014;24(3):203-206. doi:10.1016/j.semradonc.2014.02.015

175. Fallone BG. The Rotating Biplanar Linac–Magnetic Resonance Imaging System. *Semin Radiat Oncol*. 2014;24(3):200-202. doi:10.1016/j.semradonc.2014.02.011