

8-2017

Quantitative DWI as an Early Imaging Biomarker of the Response to Chemoradiation in Esophageal Cancer

Benjamin C. Musall

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations



Part of the [Investigative Techniques Commons](#), [Other Analytical, Diagnostic and Therapeutic Techniques and Equipment Commons](#), and the [Therapeutics Commons](#)

Recommended Citation

Musall, Benjamin C., "Quantitative DWI as an Early Imaging Biomarker of the Response to Chemoradiation in Esophageal Cancer" (2017). *The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access)*. 805.
https://digitalcommons.library.tmc.edu/utgsbs_dissertations/805

This Thesis (MS) is brought to you for free and open access by the The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences at DigitalCommons@TMC. It has been accepted for inclusion in The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digitalcommons@library.tmc.edu.

**QUANTITATIVE DWI AS AN EARLY IMAGING BIOMARKER OF THE
RESPONSE TO CHEMORADIATION IN ESOPHAGEAL CANCER**

by

Benjamin Charles Musall, BS

APPROVED:

Advisory Professor – Steven H Lin, MD Ph.D.

Jingfei Ma, Ph.D.

Ken-Pin Hwang, Ph.D.

R. Jason Stafford, Ph.D.

Brian P Hobbs, Ph.D.

APPROVED:

Dean, The University of Texas
MD Anderson Cancer Center UTHHealth Graduate School of Biomedical Sciences

QUANTITATIVE DWI AS AN EARLY IMAGING BIOMARKER OF THE RESPONSE
TO CHEMORADIATION IN ESOPHAGEAL CANCER

A

THESIS

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

By

Benjamin Charles Musall, BS.
Houston, Texas

August, 2017

Dedication

To my parents, Jack and Jackie.

Acknowledgements

I would like to thank my advisor, Steven Lin, for his support and boundless enthusiasm. I am honored to have been included in the work pursuing his ideas.

Special thanks to Jingfei Ma for taking me under his wing and for his tireless efforts as my mentor and advocate. I am excited to continue working for Dr. Ma in my doctoral study.

Thanks to Jong Bum “Dante” Son for his instruction, advice, sense of humor, and technical support on my project.

A huge thanks to Brian Hobbs and Bryan Fellman for their patience and thoughtful analysis of most of the data presented in this thesis.

Thanks to Brett Carter, Penny Fang, and Amy Moreno for their massive time contributions to the inter-reader agreement study.

I would also like to acknowledge and thank my other committee members for their guidance: Ken Hwang and Jason Stafford.

Finally, thanks to Bud Wendt, Frances Quintana and Betsy Kindred for their kindness and devotion to me and my fellow students.

Quantitative DWI as an early imaging biomarker of the response to chemoradiation in esophageal cancer

By: Benjamin Charles Musall, B.S.

Chair of Advisory Committee: Steven H. Lin, M.D. Ph.D.

For patients diagnosed with stages IIa-IIIb esophageal cancer, the current standard of care treatment is tri-modality therapy (TMT), where neoadjuvant chemoradiation (nCRT) is followed by surgical resection. Histopathology of resected tumors reveals that pathological complete response (pCR) is achieved in 20-30% of patients through nCRT alone. Because of the high mortality and morbidity associated with esophagectomy, it may be advantageous for patients exhibiting pCR from nCRT alone to be placed under observation rather than completing their TMT. Therefore, a method for predicting response at an early time-point during nCRT is highly desirable. Conventional methods such as endoscopic ultrasound, re-biopsy, and morphologic imaging are insufficient for this purpose. During nCRT, morphologic changes in tumors are often preceded by changes in the tumor biology. Diffusion Weighted Imaging (DWI) is an MRI modality which is sensitive to microscopic motion of water molecules in tissue. Quantitative DWI provides a measure of the cellular microenvironment which is impacted by cellularity, extra-cellular volume fraction, structure of the extracellular matrix, and cellular membranes. This work sought to investigate if changes in quantitative DWI may be used as an early imaging biomarker for the prediction of response to nCRT in esophageal cancer.

DWI scans were performed on a small group of esophageal cancer patients (stages IIa to IIIb) before, at interim, and after completion of their nCRT. Quantitative diffusion

parameter maps were estimated for DWI scans using the following models of diffusion: mono-exponential, intra-voxel incoherent motion (IVIM), and kurtosis. Summary measures of quantitative diffusion parameters were extracted from tumor voxels through volumetric contouring. These summary measures were retrospectively compared between histopathologically confirmed groupings of patients as pCR and non-pCR. The study found that the relative change in mean ADC could completely separate groupings of pCR and non-pCR patients (AUC=1) at a cutoff of 27.7%. Measurement by volume contouring was shown to be highly reproducible between readers. This pilot study demonstrates the promise of using DWI for organ sparing approaches after nCRT in esophageal cancer.

Table of Contents

Signature Page	i
Title Page.....	ii
Dedication	iii
Acknowledgements.....	iv
Abstract.....	v
List of Figures	xi
List of Tables	xvii
List of Abbreviations	xviii
1 Introduction and Background	1
1.1 Statement of Problem	1
1.1.1 General Problem	1
1.1.2 Specific Problem	2
1.2 Esophageal Cancer	3
1.2.1 Epidemiology	3
1.2.2 Detection and Staging	4
1.2.3 Treatment Strategy	6
1.2.4 Evaluation of Treatment Response	7
1.3 Magnetic Resonance Imaging	10

1.3.1	Basic Principles of MRI.....	10
1.4	Diffusion Weighted Imaging.....	20
1.4.1	Diffusion Behavior.....	20
1.4.2	Diffusion Sensitization and Contrast in MRI.....	24
1.4.3	Models of Diffusion in MRI	28
1.4.4	Advanced MRI Techniques for DWI.....	33
1.5	Hypothesis	37
1.6	Specific Aims	37
2	Materials and Methods	38
2.1	Patient Selection and Outcome Measures	38
2.2	MRI Scan Protocols	38
2.3	Diffusion Parameter Mapping.....	41
2.4	Contour Measurements of Diffusion Parameters	47
2.4.1	General Contour Strategies	47
2.4.2	Software Implementation.....	50
2.5	Inter-Reader Study	55
2.6	Statistical Analysis	56
2.6.1	Correlation of Diffusion Parameter Summary Measures with Histopathology.....	56
2.6.2	Inter-Reader Reproducibility Analysis	57

3	Results	58
3.1	Patient Enrollment.....	58
3.2	Image Quality	60
3.3	Correlation of DWI Parameters with Histopathology	64
3.3.1	Mono-exponential Model (ADC)	64
3.3.2	IVIM Model	68
3.4	Inter-Reader Study	74
4	Discussion.....	76
4.1	Response Prediction by DWI	76
4.2	Reproducibility of Contour Measurements between Readers.....	77
4.3	Limitations	79
4.4	Future Work	79
5	Appendix	82
5.1	Validation of ADC Map Calculation by Comparison to GE Functool	82
5.2	ITK-SNAP Settings and Guide	82
5.3	Alternative Manual ImageI Contouring Method.....	89
5.4	Diffusion Image Artefacts	93
5.4.1	Flashing Artefacts / Single Slice Signal Drop Out	93
5.4.2	T ₂ Shine-through.....	94

	5.5 Comparison of Diffusion Parameter Measurements between Alternate Groupings of <1% Viable Cell Patients	94
	5.6 Padhani et al. ADC Method Comparison.....	95
	5.7 Comparison of Largest Slice and Single Slice Contour Measurements.....	97
	5.8 Correlation of IVIM D and MONO ADC.....	99
6	References	100
7	Vita	113

List of Figures

Figure 1-1: Illustration depicting regions of the esophagus.....	4
Figure 1-2: Layers of the esophagus.....	5
Figure 1-3: Tumor Regression Grades by Mandard grouping depicted as illustrations.	8
Figure 1-4: Diagram indicating excess alignment of spins with B_0 to form the net magnetization M_0 . Conventional designation of Cartesian dimensions is also shown.	11
Figure 1-5: Slice selection through application of magnetic field gradient in combination with a selective bandwidth (W) RF pulse.	15
Figure 1-6: Image space and k-Space of a T_2 -weighted image of a human thorax.	17
Figure 1-7: Gradient encoding of 2D Fourier data as a both gradient waveform and as a k-space trajectory. Phase encoding is shown in orange, and frequency encoding is shown in green..	18
Figure 1-8: A typical SE sequence. The pulse sequence, gradient waveform, and readout signal are shown.	19
Figure 1-9: A typical GE sequence. The pulse sequence, gradient waveform, and readout signal are shown.	20
Figure 1-10: Illustrations depicting movement of water molecules within tissues of varying cellularity. The blue line represents the random movement of a water molecule while the black circles represent cells. Republished with permission (License ID = 4167471220873 from www.copyright.com) from original article: O'Flynn, E. A. M., and N. M. deSouza. 2011. Functional magnetic resonance: biomarkers of response in breast cancer. <i>Breast Cancer Research : BCR</i> 13: 204.....	23
Figure 1-11: Pulse sequence and simplified gradient waveform for a PGSE sequence. The gradient duration is labeled with δ , and the time elapsed between gradient is labeled with Δ .	

At the bottom, an illustration is shown comparing dephasing and rephasing effects on the magnetization of moving (red) and non-moving (blue) spins.	26
Figure 1-12: Graphical depiction of deviations in signal levels due to perfusion effects (red) at low b-values and from kurtosis effects (green) at high b-values.	29
Figure 1-13: Random orientation of capillaries within a hypothetical voxel of tissue. Republished with permission (License ID = 4167550994385 from www.copyright.com) from original article: Malyar, N. M., M. Gossel, P. E. Beighley, and E. L. Ritman. 2004. Relationship between arterial diameter and perfused tissue volume in myocardial microcirculation: a micro-CT-based analysis. <i>American journal of physiology. Heart and circulatory physiology</i> 286: H2386-2392.	30
Figure 1-14: Depiction of probability function distributions with differing kurtosis. The kurtosis observed in diffusion is positive, which is depicted as a change in the probability density function from the red plot to the green plot.....	32
Figure 1-15: EPI k-space trajectory where the phase encoding direction is blipped, and the frequency encoding direction is continuous.	34
Figure 2-1: Comparison of LMS fitted ADC plot with plotted measured signal values for a single voxel.	42
Figure 2-2: Comparison of LMS fitted IVIM model with measured signal values for a single voxel.....	44
Figure 2-3: Comparison of LMS fitted DKI model with measured signal values for a single voxel. A plot of the kurtosis model is seen in red, while a plot of the DKI-calculated ADC as a mono-exponential decay is seen in green. Deviation of the signal from mono-exponential behavior can be seen at high b-values.....	46

Figure 2-4: This figure demonstrates the indistinct SI boundaries in the b200 images and the distinct SI boundaries in ADC maps. It can be seen that the b200 images demonstrate a noticeable, but indistinct lowering of signal as slices are they sequentially examined in the inferior direction towards the inferior boundary of the tumor (upper row, left→right). The ADC map (bottom row) demonstrates a distinct inferior boundary when examined in this manner, which can be noticed as a large rise in ADC values between the second and third image. Red arrows in the leftmost images mark areas of low ADC corresponding to tumor, and the red arrow in the third image marks where this area of low ADC is missing.	49
Figure 2-5: Workflow for measurement of diffusion parameters through volume contouring.	53
Figure 2-6: Workflow for measurement of diffusion parameters through slice contouring...	55
Figure 3-1: Diffusion images (a, b) and ADC map (c) from MONO exam of a patient at BL. Contours made on b100 are overlain on the image in white. FDG-PET (d), T ₂ -weighted image (e), and CT (f) images of the approximate slice are included for comparison.	61
Figure 3-2: Diffusion images (a, b) and parameter maps (d-e) from IVIM exam of a patient at BL. Contours made on b100 are overlain on the image in red. A T ₂ -weighted image (c) of the approximate slice is included for comparison.	62
Figure 3-3: Diffusion images (a-c) and parameter maps (d,e) from DKI exam of a patient at BL. Contours made on b200 are overlain on the image in white. A T ₂ -weighted image (f) of the approximate slice is included for comparison.....	63
Figure 3-4: Comparison of distributions in Δ ADC mean between pCR and non-pCR groupings. P-value by MW test is shown at top. The 27.7% cutoff between pCR and non-pCR is shown	

as a dotted line. 95% confidence intervals and means are displayed with crosshairs for each group.	67
Figure 3-5: Spearman’s correlation matrix of the top 10 performing summary measures of Δ ADC.....	68
Figure 3-6: Comparison of distributions in Δ D 90 th Percentile between pCR and non-pCR groupings. P-value by MW test is shown at top. The 27.7% cutoff between pCR and non-pCR is shown as a dotted line. 95% confidence intervals and means are displayed with crosshairs for each group.	72
Figure 3-7: Spearman’s correlation matrix of the top 10 performing summary measures of IVIM diffusion parameters as relative change at IM.	73
Figure 3-8: Bland-Altman plots comparing inter-reader variability for ADC 25 th percentile (a) and ADC mean (b) from MONO scans.	76
Figure 5-1: Minimization of volume for semi-automated contouring. The new smaller volume is outlined in red lines. The display includes axial (Upper Left), sagittal (Upper Right) and coronal (Bottom Right) views.....	83
Figure 5-2: Final adjustments of thresholding for speed images of a minimized scan volume. The display includes axial (Upper Left), sagittal (Upper Right) and coronal (Bottom Right) views.	85
Figure 5-3: An example of good bubble seed placement within the tumor volume.....	86
Figure 5-4: Screenshot of active contouring step before active contouring is activated.	87
Figure 5-5: Screenshot of completed contour after iterative update of contour boundaries during active contouring has been stopped with appropriate timing.	88

Figure 5-6: Rough outlining of tumor using several sweeps of the paintbrush tool. This should take on the order of a few seconds for each slice.	90
Figure 5-7: Example of edge contouring and lumen exclusion through raising minimum included signal level using the volume histogram tool in ImageI. A red arrow shows the movement of the slider, which was manually dragged by the user.	91
Figure 5-8: Final completion of the contour by manual adjustment of edges with eraser tool. The red arrows show points where the contour from Figure 5-7 was adjusted.	92
Figure 5-9: Massive shift of tumor signal on a single slice due to motion, causing a “flashing” effect when slices are scrolled through quickly. This is caused large scale motion or by parallel imaging correction failure, which leads to wrap-around artefacts in the phase direction (middle image).	93
Figure 5-10: This is an example of T ₂ shine-through in the b200 image (a) and ADC map (b) of a MONO scan. In this scan, the cancerous area of the esophagus is labeled with a red arrow and healthy stomach is labeled with a blue arrow. Although the stomach is not cancerous and is not experiencing abnormal restriction of diffusion, its T ₂ is relatively long, causing it to be bright on the diffusion image. The ADC map accounts for T ₂ shine-through, which is why the enhanced stomach on the diffusion image does not match with an area of low ADC (darker) on the ADC map.	94
Figure 5-11: Comparison of Δ ADC mean at interim between an alternate set of viable cell groupings. The 27.7% cutoff between pCR and non-pCR is shown as a dotted line.	95
Figure 5-12: Correlation of LMS Curve Fitting Method Mean ADC across all MONO Scans with Mean ADC by Alternate Methods of ADC Estimation. R ² values from linear regression are shown at the top.	96

Figure 5-13: Correlation of ADC mean measurements by Largest Slice (y-axis) and Single Slice (x-axis) methods. Graphs are titled with their respective readers, and R^2 values are shown at the top of the graphs.....	98
Figure 5-14: Correlation of Volume Mean D by IVIM with Volume ADC Mean by MONO through two different estimation methods. “LMS” corresponds to the MONO ADC method used in this paper, while “200 800” corresponds with the method put forth by Padhani et al. R^2 values from linear regression are shown at the top.....	99

List of Tables

Table 2-1. Typical scan parameters for DWI scans.	39
Table 2-2. B-values and NEX for diffusion scans.	40
Table 3-1. Patient and Pathology Characteristics	59
Table 3-2. Comparison of Baseline ADC Summary Measure Distributions between pCR and non-pCR using Mann-Whitney Test.	65
Table 3-3. Comparison of Distributions of Δ ADC Summary Measures between pCR and non-pCR using Mann-Whitney Test.	66
Table 3-4. Comparison of Baseline IVIM Parameter Summary Measures between pCR and non-pCR using MW test.	69
Table 3-5. Comparison of Distributions of Δ IVIM Summary Measures at IM between pCR and non-pCR using Mann-Whitney Test.	70
Table 3-6. Comparison of Distributions of Δ IVIM Summary Measures at FU between pCR and non-pCR using Mann-Whitney Test.	73
Table 3-7: Inter-Reader Reproducibility for MONO Δ ADC Summary Measures by Bland-Altman and ROC Analyses.	75
Table 5-1. Comparison between extracted volume ADC mean from ADC maps generated by ImageI and GE Functool.....	82

List of Abbreviations

ADC	Apparent Diffusion Coefficient
AJCC	American Joint Committee on Cancer
AP	Antero-posterior
AUC	Area Under Curve
BL	Baseline
BMI	Body Mass Index
BMT	Bimodality Therapy
CNN	Convolutional Neural Network
CT	Computed Tomography
DCE	Dynamic Contrast Enhanced
DKI	Diffusion Kurtosis Imaging
DSC	Dice-Similarity Coefficient
DWI	Diffusion Weighted Imaging
EAC	Esophageal Adenocarcinoma
ECM	Extracellular Matrix
ECVF	Extracellular Volume Fraction
EPI	Echo Planar Imaging
ESCC	Esophageal Squamous Cell Carcinoma
EUS	Endoscopic Ultrasound
FDG-PET	Fluorodeoxyglucose - Positron Emission Tomography
FID	Free Induction Decay

FOV	Field of View
FT	Fourier Transform
FU	Follow Up
GE	Gradient Echo
GE	Gastroesophageal
GE	Gradient Echo
GI	Gastrointestinal
GRAPPA	Generalized
HE	Hematoxylin and Eosin
IM	Interim
IRB	Institutional Review Board
IVIM	Intra-voxel Incoherent Motion
KM	Kaplan-Meier
KW	Kruskal-Wallis
LMS	Least Mean Squares
MDACC	MD Anderson Cancer Center
MIE	Minimally invasive esophagectomy
MONO	Mono-exponential Model
MRI	Magnetic Resonance Imaging
MW	Mann-Whitney
nCRT	Neoadjuvant Chemoradiation Therapy
NMR	Nuclear Magnetic Resonance
PACS	Picture Archiving and Communication System

pCR	Pathologic Complete Response
PGSE	Pulsed Gradient Spin Echo
RF	Radiofrequency
RF	Radiofrequency
rFOV	Reduced Field of View
ROC	Receiver Operator Curve
SE	Spin Echo
SI	Superior-Inferior
ssEPI	Single-shot Echo Planar Imaging
SSOS	Square Root Sum of Squares
TE	Echo Time
TMT	Trimodality Therapy
TNM	Tumor Node Metastasis (Staging system by AJCC)
TR	Repetition Time
TRG	Tumor Regression Grade
VATS	Video Assisted Thoracoscopic Surgery

1 Introduction and Background

1.1 Statement of Problem

1.1.1 General Problem

For a majority of cancer, patients undergo clinical management based on their cancer staging. The staging of the patients typically relies on their clinical status, lab tests, imaging, and biopsy. The actual treatment regimen is strongly influenced by clinical staging, the site of the malignancy, and the health of the patient. During the treatment, the patient may be restaged depending on the response. Because different patients of the same staging may respond to a treatment differently and every treatment has its toxicity, it is often critical to characterize individual disease and improve treatment selection on an individual basis. Evaluating treatment response for individual patients is an essential part of the clinical management of patients. In particular, prediction of treatment response at an early time point during therapy would allow for alteration of a patients therapy regimen to minimize the toxicity and maximize the efficiency of the treatment of the patients.

Cancer therapies, with the obvious exception of surgery, are usually implemented over a series of weeks. For radiation therapy, a long treatment time is often needed to allow normal tissues to recover between treatments. For chemotherapy, long treatment times are needed to achieve the desired malignant cell kill while maintaining toxic chemotherapy agents at levels acceptable to the patient health. At the end of the therapy regimen, the response to treatment is often assessed by imaging and the histopathology of surgical specimens. Whenever available, histopathological metrics are considered a “gold standard”

for assessing how well an individual responded to therapy. There are also non-invasive metrics available after surgery: continued observance of changes in tumor morphology through imaging and prognostic FDG-PET imaging post-therapy. Although these excellent prognostic tools are available post-therapy, morphologic measures through conventional imaging as well as re-biopsy are insufficient for interrogation of patient response early during therapy. Changes in the tumor microenvironment, which are not visible morphologically, occur early during therapy. Quantitative Diffusion Weighted Imaging (DWI) provides a measure of these changes through its sensitivity to diffusion behavior of water molecules in tissue. Because of this, it may hold value as an early predictor of response.

1.1.2 Specific Problem

For esophageal cancers of stages IIa – IIIb, the current standardized of care is for patients to be treated with neoadjuvant chemoradiation (nCRT), followed by surgical resection (1). Histopathological metrics post-surgery reveal that 20-30% of patients exhibit pathologic complete response (pCR) from chemoradiation alone (2-4). Surgical methods for removing esophageal tumors are highly invasive, with high risk of mortality and morbidity in these patients (5).

A method of predicting a patient's response early during their therapy would allow for pCR patients to forego surgery and instead be placed under observation. Conventional imaging methods, re-biopsy, and some basic genetic mutation characterization methods have been found to be insufficient for the purpose of predicting patient response (6-10). There is an unmet clinical need for a method which can provide an early prediction of response to chemoradiation in esophageal cancer. Measured changes in intra-tumoral

diffusion behavior of water molecules during treatment with quantitative DWI may be able to fulfill this need.

1.2 Esophageal Cancer

1.2.1 Epidemiology

Esophageal cancer is the sixth leading cause of cancer-related mortality worldwide and its incidence is increasing (11). Overall 5-year survival is under 25%. Esophageal malignancies are separated into two major histologic groupings based on their tissue type of origin. Adenocarcinoma (EAC), which originates in esophageal glands, is more common in North America and Western Europe. Squamous cell carcinoma (ESCC), which originates in the squamous cells of the epithelium in the esophagus, is more common across Asia, South America, and southern Africa. A primary risk factor for EAC is high body mass index (BMI) (12). Due to the endemic of obesity, incidence of EAC is rapidly increasing. Primary risk factors for ESCC include smoking, poor diet, alcohol consumption (12). Across all esophageal cancers, incidence is four times more likely in men than in women (12).

Esophageal cancers can also be divided according to the anatomic location of the disease. The esophagus can be divided into three distinct parts. A diagram of this division can be seen in Figure 1-1.

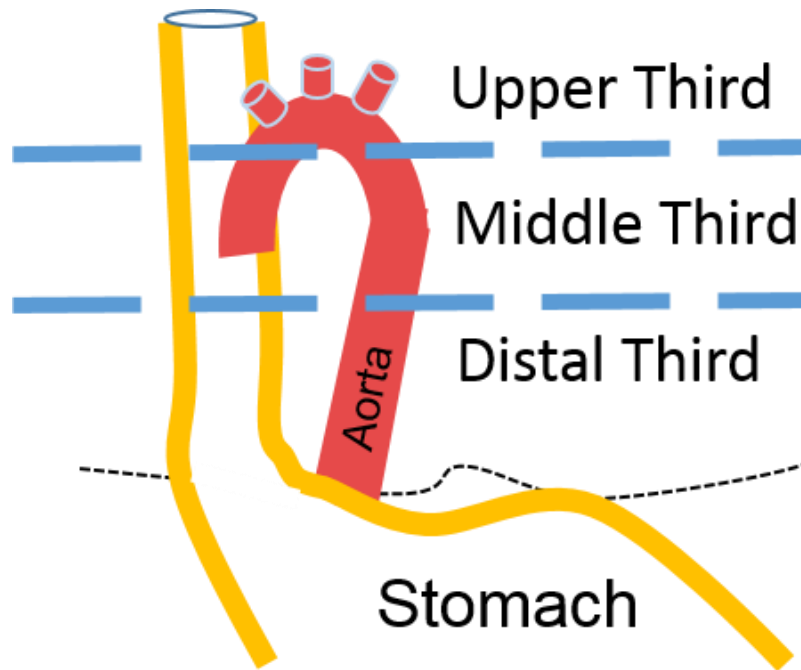


Figure 1-1: Illustration depicting regions of the esophagus.

For AEC's, the proximal third and middle third locations account for only a quarter of the cancers in esophageal cancers, while the remaining cancers are located in the distal third and gastroesophageal (GE) junction (12). ESCC is more commonly found in the middle third. Tumor location is important for radiation treatment planning and surgical planning.

1.2.2 Detection and Staging

Esophageal cancer is typically asymptomatic until sufficient tumor growth has occurred. Because of this, it is usually detected in later stages. Clinical symptoms include dysphagia (difficulty swallowing) and chest pain. After symptoms are reported, patients are examined and sent for further tests. Imaging methods used for diagnosing esophageal cancer include chest x-ray, barium swallow fluoroscopy, contrast CT, and FDG-PET. Once it is confirmed that the patient has cancer, their cancer is then staged.

Staging of esophageal cancer may include information gained from diagnostic imaging methods (13). It will also include endoscopic ultrasound (EUS) and biopsy. First, lesion characteristics are used to define the cancer by the TNM system of the American Joint Committee on Cancer (AJCC). In this system, lesions are given an initial T-stage based on the degree of their invasion from their site of origin into layers of surrounding tissue. In the esophagus, T1 implies that the cancer has spread into the epithelium; T2 implies spread into the muscularis propria; T3 implies that the cancer is growing into the adventitia, and T4 implies local spread to nearby organs. An illustration of these tissue layers is given in Figure 1-2.

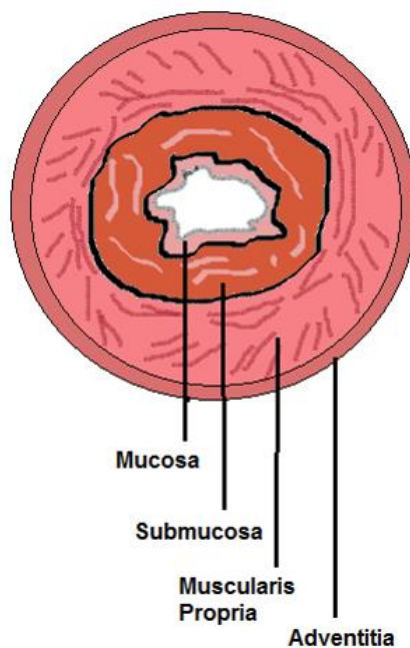


Figure 1-2: Layers of the esophagus.

N in TNM represents the relative number of local lymph nodes involved. Further, a binary M-staging is given, where M0 represents no metastatic spread and M1 represents metastatic spread. After this, groupings of the cancer are made based on location, TNM

staging, and biopsy grading of cell appearance. Stage I patients have no nodal involvement, no invasion beyond the epithelium, no metastases, and are well differentiated. Stages II-II are stratified into several levels based on variations of T and N stages as well as tumor location and cell grading. Stage IV implies metastatic disease.

1.2.3 Treatment Strategy

A treatment course is based on a comprehensive consideration of the initial staging and the health of the patient. For stage I, surgery is implemented immediately, and is usually curative (14). For stage IV, palliative treatment is given. For stages IIa through IIIb, the optimum treatment for patients is trimodality therapy (TMT) (15).

1.2.3.1 Neoadjuvant Chemoradiation Therapy

During TMT, neoadjuvant chemoradiotherapy (nCRT) is used to shrink the tumor and followed by resection in patients who are healthy enough to be considered for surgery. For radiation, tumors are irradiated to a total dose of 50.4 Gy (1.8 Gy/frac) over a period of approximately five weeks (1). Chemotherapy will accompany radiation, consisting of a combination of cisplatin and either docetaxel or fluorouracil (1). After completion of nCRT, a patient will undergo surgery. Some patients may be unable to undergo surgery because of health problems. In this case, their treatment regimen, consisting of only nCRT, is referred to as bimodality therapy (BMT).

1.2.3.2 Surgical Excision

Surgical resection has shown benefits for locoregional control over definitive chemoradiation (16). Esophagectomy requires a combined approach towards tumors from

both thoracic and abdominal cavities (17). Historically, this was performed through an open thoracotomy, where an incision is made in the pleural space of the chest (18). More recently, minimally invasive esophagectomy (MIE) through video-assisted thoracoscopic surgery (VATS) has become available. MIE does not require a formal pleural incision, and may be advantageous over open esophagectomy (19). Both techniques requires general anesthesia and insertion of breathing tubes into the lung. During esophagectomy, a number of adjacent lymph nodes are also extracted and examined for cancer spread.

Even in experienced centers, post-operative mortality and morbidity from esophagectomy are in the range of 7-8% (20). This is due to both the radical nature of the surgery as well as the weakened state of the cancer patients after the chemoradiation.

1.2.4 Evaluation of Treatment Response

1.2.4.1 Post-CRT Methods

Before surgery, a follow-up (FU) FDG-PET scan is taken for the detection of residual disease, as well as to check for metastatic spread. Qualitative evaluation of FDG-PET scans after induction chemotherapy has been shown to be highly prognostic of clinical outcome for both BMT and TMT patients (21).

After surgery, sections of resected tumor are sent to a pathology laboratory and examined for residual disease. To accomplish this, the resected tissue is formed into slides and stained with the Hematoxylin and Eosin (H&E) stain. Hematoxylin stains nucleic acids such as DNA and RNA with a deep blue-purple color, and Eosin is a non-specific stain which labels proteins with a light pink. This H&E staining method is considered a gold standard for the examination of cancer histopathology (22). In this presentation, trained

experts in histopathology can discern cellular structures within the resected tissue. For the purpose of evaluating residual disease status, histopathologists will search H&E stained slides for cancer cells and cancer structures which have survived treatment. Several methods of grouping patients by their post-treatment pathology, or restaging, have been presented in literature (23). An early grouping method was described for esophageal cancer by Mandard et al. and is based on a qualitative assessment of surviving cancer cells (24). Visual depiction of these groupings and their descriptions can be seen in Figure 1-3.

Tumor Regression	Residual Cancer
TRG1	No residual cancer
TRG2	Rare residual cancer
TRG3	Fibrosis outgrowing residual cancer
TRG4	Residual cancer outgrowing fibrosis
TRG5	Absence of regressive changes

Figure 1-3: Tumor Regression Grades by Mandard grouping depicted as illustrations.

In the Mandard method, patients are placed in tumor regression groups based on subjective qualitative measures. At MDACC, an quantitative adaptation of the Mandard method known as the four-tiered grading system is used (25). In this method, a term called viable cell percentage is derived by comparing of the relative areas occupied by cancer cells and normal tissue. In the four-tiered system, TRG1 corresponds to 0% viable cancer cells or pathologic complete response (pCR). TRG2 corresponds to measures of tumor cells between

<1% to 10% viable cells. TRG3 corresponds to >10% to 50% viable cancer cells, and TRG4 corresponds to >50% viable tumor cells.

1.2.4.2 Early Prediction of Response

Previously, several possible methods of predicting response of esophageal therapy during nCRT have been examined. Endoscopic US combined with biopsy has been shown as insufficient for predicting pCR (6, 7). A potential reason is that morphologic changes visible with US may not be present at an early time-point during therapy, and biopsy is always limited by its sampling errors. Likewise, CT and T₂-weighted MR imaging are not able to predict response (26). Genetic characterization has shown minor successes (8-10) but does not yet have the sensitivity needed to predict pCR. Quantitative FDG-PET has also shown some success (7, 27), but is difficult to measure during therapy because of non-specific uptake caused by radiation induced inflammation. FDG-PET results may improve in future studies due to improved protocol design and improved methods for acquiring FDG-PET parameter summary measurements.

Multi-parametric MRI is being investigated for predicting therapy response in cancers with increasing frequency and varying success (28). In a previous study, quantitative diffusion MRI has demonstrated excellent sensitivity to pCR (29). The protocols within this thesis were designed, in part, as a validation of this previous study by Van Rossum et al. Dynamic Contrast Enhanced (DCE) MRI has also been shown to have value for response prediction (30), but suffers from poor reproducibility.

1.3 Magnetic Resonance Imaging

This purpose of this section is two-fold. The first is to demonstrate an understanding of MRI from a fundamental level. The second is to later extend these fundamental explanations to the more specific and advanced MRI methods used in the MRI research protocols of this thesis.

1.3.1 Basic Principles of MRI

1.3.1.1 Behavior of Protons in a Magnetic Field

For atomic isotopes with an odd number of nucleons, the net angular momentum of the system is non-zero. In the case of protium (^1H), the contribution of a single proton creates a net angular momentum or “spin” of $\pm \frac{1}{2}$. When a proton is placed in an external magnetic field, this non-zero spin will align with the direction of the external magnetic field and precess about this direction. The frequency of precession is referred to as the Larmor frequency, and it is dependent on both the strength of the external magnetic field (B_0) as well as the intrinsic gyromagnetic ratio of the proton (42.58 MHz/T). This relationship is described in the Larmor equation:

$$\text{Equation 1: } \omega_0 = \gamma B_0$$

where ω_0 is the Larmor frequency, γ is the gyromagnetic ratio, and B_0 is the external magnetic field strength.

In a group of protons, not all of the protons will align with the magnetic field (parallel). A slightly smaller amount of them will align in the direction opposite the external magnetic field (anti-parallel). At 25°C, thermal energy is sufficient to induce changes

between these two quantum states in individual proton spins. Despite this, a sufficiently large group of protons will always have slightly more spins in the parallel state, creating a bulk magnetization. This magnetization is proportional to B_0 and inversely proportional to temperature. This relationship is described through combination of a Maxwell's Laws and a Boltzmann distribution in Equation 2:

$$\text{Equation 2: } M_0 = \frac{\rho_0 \gamma^2 \hbar^2}{4kT} B_0$$

where M_0 is the net magnetization per unit volume, ρ_0 is the density of protons, γ is the gyromagnetic ratio, \hbar is the Planck's constant, k is the Boltzmann constant, T is temperature, and B_0 is the external magnetic field. In MRI, B_0 refers to the static magnetic field. This net magnetization is depicted in Figure 1-4.

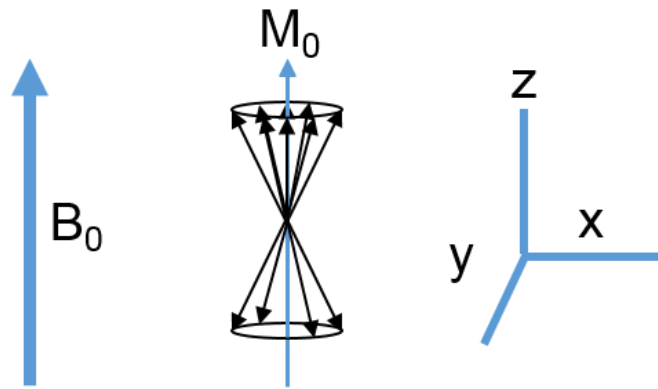


Figure 1-4: Diagram indicating excess alignment of spins with B_0 to form the net magnetization M_0 . Conventional designation of Cartesian dimensions is also shown.

The high density of water protons in vivo combined with high magnetic field strengths in clinical MRI scanners create a relatively large bulk magnetization in tissue. This bulk magnetization and its properties are used for imaging with MRI.

1.3.1.2 Bloch Equations

The group behavior of spins or bulk magnetization can be derived from Maxwell's laws, and were first described mathematically by Felix Bloch in 1946 (31). Behavior of bulk magnetization in the presence of external magnetic fields can be described through the Bloch equation:

$$\text{Equation 3: } \frac{d\vec{M}(t)}{dt} = \gamma \vec{M}(t) \times \vec{B}_{ext} - \frac{M_{transverse}}{T_2} - \frac{M_{parallel} - M_{0,parallel}}{T_1}$$

where $M(t)$ is the time-variant magnetization of the sample, γ is the gyromagnetic ratio, B_{ext} is the vector of combined external magnetizations, $M_{transverse}$ is the magnetization of the sample transverse to B_0 , T_2 is the spin-spin relaxation constant, $M_{parallel}$ is the magnetization of the sample parallel to B_0 , and T_1 is the spin-lattice relaxation constant. Spin-spin relaxation, or T_2 , refers to dephasing of spins in the transverse plan by exchange of angular momentum between spins through dipole-dipole interactions, variations in the local magnetic fields, and other interactions (32). As implied in the combined Bloch equation, T_2 relaxation through gradual misalignment of individual spins results in an exponential decay of $M_{transverse}$. T_2 is also affected by B_0 inhomogeneities, which result in a shorter apparent T_2 relaxation time referred to as T_2^* . Separately, T_1 relaxation refers to the exponential recovery of $M_{parallel}$ as magnetization returns from an excited state to its steady state alignment with B_0 . Both T_2 and T_1 relaxations create exponential decay of the transverse net magnetization. T_2 refers to the time when 63% of transverse magnetization has been lost to spin-spin relaxation and T_1 refers to a time where 63% of the magnetization has realigned with B_0 . In tissue, these relation times are variable between tissue types. This variation in relaxation times between tissues is used as a source of contrast in anatomic imaging.

By selecting a time of measurement which maximizes the difference in transverse magnetization between tissues of differing T_1 or T_2 , a weighted contrast can be optimized for the image. For individual tissues, T_1 is always longer than T_2 . Thus, relatively short measurement times (echo time, TE) with little recovery between measurements (repetition time, TR) produce greater emphasis on T_1 differences and are considered T_1 -weighted. Conversely, relatively long TE and long TR emphasize T_2 differences and considered T_2 -weighted. Short TE and long TR minimize effects of both T_1 and T_2 relaxations on measurement, and give way to a proton density weighting of measurements.

1.3.1.3 Signal Generation

To detect the bulk magnetization of protons in a magnetized sample, the magnetization must be re-oriented into the transverse plane. This may be accomplished through the application of an oscillating secondary magnetic field, B_1 . In MRI, this B_1 field must oscillate with a frequency equal to Larmor frequency of protons within the B_0 field. At the B_0 strength of a clinical MRI scanner, an oscillating magnetic field of this frequency falls within the range of the electromagnetic spectrum corresponding to radio waves. Because of this, B_1 fields are also referred to as a radiofrequency (RF) pulse. The force imparted from a B_1 field transverse to B_0 can be used to nutate the net magnetization of the sample into the transverse plane. The angle of this nutation or flip is dependent on the magnitude and duration of the RF pulse, as shown in Equation 4:

$$\text{Equation 4: } \alpha = \gamma \int B_1(t) dt$$

where α is the flip angle, γ is the gyromagnetic ratio, and B_1 is the magnitude of the RF pulse. A pulse which tips magnetization from its steady state alignment with B_0 into the

transverse plane is referred to as an excitation pulse. If the magnetization is left undisturbed after its excitation, an exponential decay of its transverse component called Free Induction Decay (FID) will occur by T_2^* relaxation.

Through an interaction described by Faraday's Law, the transverse magnetization induced emf or an electric current in orthogonally placed conducting loops. In MRI, these conducting loops are known as receiver coils. Although receiver coils are more sensitive to nearby changing magnetization in accordance with the inverse square law, they are unable to fully spatially localize spins. Thus, a technique for spatially encoding spins into received signal must be utilized.

1.3.1.4 Location Encoding

The addition of spatially variant magnetic fields is used to localize magnetization in MR imaging. In Cartesian imaging, localization must be accomplished in three dimensions. While the direction of these dimensions is arbitrary and may be altered to better portray certain anatomies, a majority of MRI acquisitions are taken as axial images. For localization in axial imaging, the three spatial dimensions may be separated into two types: the dimension which is parallel to the B_0 field (the z-axis, by convention) and the two dimensions of the transverse plane (x-y plane, by convention).

Localization along the z-axis is also known as slice selection. For visualization of axial images, variations in signal are displayed on the transverse plane or "slice", while the z-axis is implicit and unseen dimension corresponding to the thickness of the visualized anatomy. A technique for selection of a slice of certain location and thickness was first described by Peter Mansfield in the 1970's (33). The desired thickness and location of a slice is selected during excitation with the addition of a gradient field G_z , which creates a linear

variation in the Larmor frequency along the z-axis. A slice of spins may be selectively excited using an RF pulse whose central frequency and bandwidth are limited to the Larmor frequency range present within spins of the desired slice. This relation is detailed in Equation 5:

$$\text{Equation 5: } BW_{RF} = \gamma G_z \Delta z$$

where BW_{RF} is the frequency bandwidth of the RF pulse, γ is the gyromagnetic ratio, G_z is the slice selection gradient, and Δz is the thickness of the sample intended for excitation. In this equation, it is implied that the central frequency of the RF pulse corresponds to the Larmor frequency of spins centered in the z-dimension of the slice. This process of slice selective excitation may be reapplied if multiple RF pulses are used sequentially. A visual illustration of slice selective excitation can be seen in Figure 1-5.

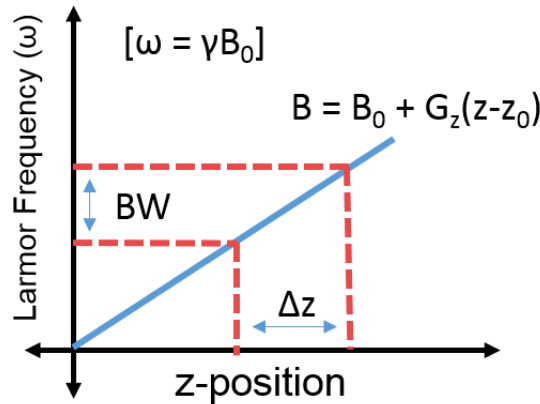


Figure 1-5: Slice selection through application of magnetic field gradient in combination with a selective bandwidth (W) RF pulse.

Although encoding in the transverse dimensions is also accomplished with spatially varying magnetic fields, their implementation is different. Unlike slice selection, the two transverse dimensions are encoded directly into the signal which is read-out by the receiver

coils. The two linear magnetic fields, G_x and G_y , do not isolate signal from a specific point within the imaging plane. Rather, these gradients produce spatial variations in the contribution of all excited spins. Transverse gradients are used to create spatial variation of magnetic field across the transverse dimension, which produces spatial variation in the Larmor frequency of spins. Over the duration of the gradient, the variable spin frequencies will result in spatially varying directionality of magnetization. This may also be referred to as an accumulation of phase. The amount of phase accumulated during the application of the gradient is dependent on the position of the spin as well as the magnitude and duration of the gradient. This is described mathematically by Equation 6:

$$\textbf{Equation 6: } \theta = \gamma \int G_x(x, t) x dt$$

where θ is the accumulated phase, γ is the gyromagnetic ratio, $G_x(x)$ is the strength of the gradient at location x , and x is the location of the spin.

Because phase is periodic, its variation will occur cyclically along the direction of the gradient. This phenomenon is better known as spatial frequency, with units of inverse unit length. The spatial frequency in spins produced through application of a gradient field is described in Equation 7:

$$\textbf{Equation 7: } k = \gamma \int G(t) dt$$

where k is the spatial frequency, γ is the gyromagnetic ratio, and G is the magnitude of the gradient. During an acquisition, information will be sampled using spatial frequencies of varying magnitude and direction across the transverse plane. The signal collected from this acquisition can be plotted in k -space, which is the conjugate of image space. A comparison of these spaces can be seen in Figure 1-6.

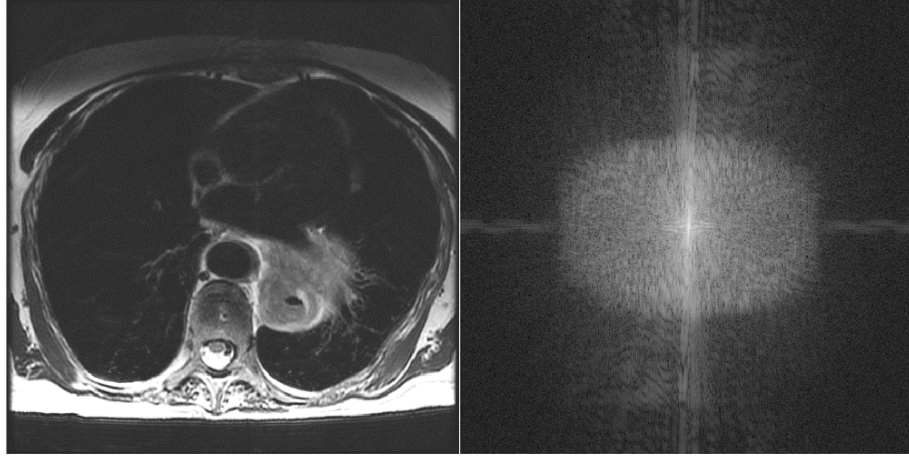


Figure 1-6: Image space and k-Space of a T₂-weighted image of a human thorax.

By using the inverse Fourier transform to change the dimensionality of the data, the k-space acquisition may be converted into image space. The Fourier relationship between k-space and image space is described by Equation 8:

$$\text{Equation 8: } S(k_x, k_y) = \iint \rho(r) e^{-i2\pi(k_x x + k_y y)}$$

where $S(k_x, k_y)$ is the distribution of signal in k-space, $\rho(r)$ is the spatial distribution of spins or proton density, k_x/k_y are the Cartesian coordinates of different spatial frequencies in k-space, and x/y is the location of spins in Cartesian image space. The central area of k-space (near the origin) contains information on low spatial frequencies of the system. It serves as the primary source of image contrast and has a relatively high signal. The outer section of k-space contains information on higher spatial frequencies. It has relatively small signal, and its extent determines the spatial resolution of an image.

In conventional MRI, filling k-space can be more efficiently accomplished by continuous readout along one dimension (referred to as frequency encoding) and stepped along another (phase encoding). By convention, the x-direction is typically considered as the

frequency encoding direction, while the y-dimension is typically considered the phase encoding dimension. A gradient pulse diagram and k-space trajectory of this conventional acquisition method can be seen in Figure 1-7.

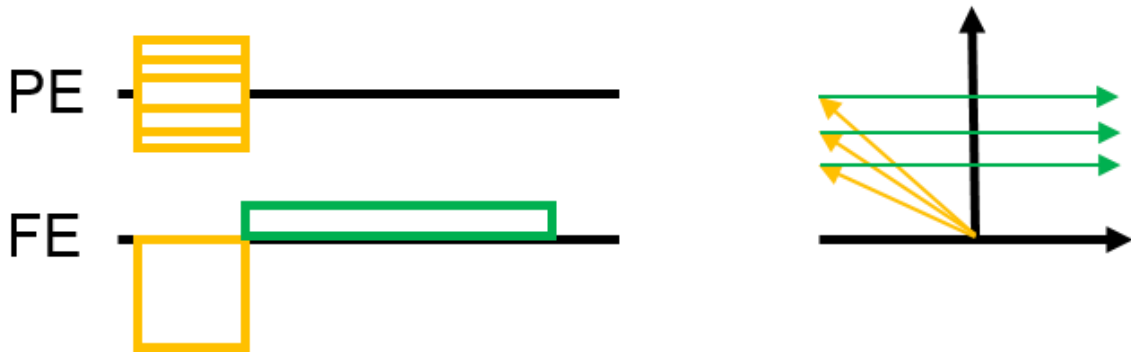


Figure 1-7: Gradient encoding of 2D Fourier data as a both gradient waveform and as a k-space trajectory. Phase encoding is shown in orange, and frequency encoding is shown in green.

1.3.1.5 Basic Pulse Sequences

In MRI, two basic pulse sequences underlie most of clinical imaging: spin-echo (SE) and gradient-echo (GE). The GE pulse sequence uses the FID signal for spatial encoding. In contrast, the SE pulse sequence uses the spin echo signal from the application of an excitation/refocusing RF pulse pair for spatial encoding.

In spin-echo sequences, a second RF pulse (B_1) is applied to the spins with a flip angle of 180° . This 180° pulse is typically applied orthogonally to both the B_0 field and the excitation pulse. The 180° pulse flips the spins in the transverse plane to their conjugate direction so that dephasing of different spins due to B_0 inhomogeneities is reversed. This rewinds the phase evolution of the spins to form coherent signal in the form of an echo. Because the dephasing effects of B_0 inhomogeneities are reversed, the observed T_2 contrast at the echo is attributable to T_2 alone rather than T_2^* . It is advantageous to select a flip angle of

90° for SE because it maximizes the signal. A pulse sequence diagram for a typical SE sequence with conventional location encoding can be seen in Figure 1-8.

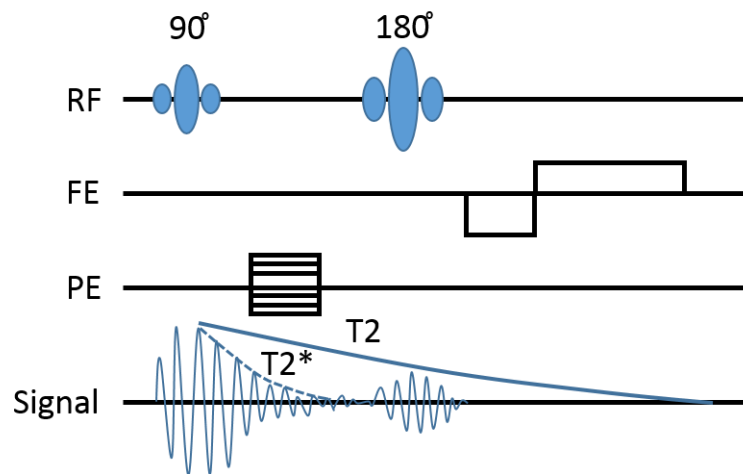


Figure 1-8: A typical SE sequence. The pulse sequence, gradient waveform, and readout signal are shown.

For GE, a true echo is not formed. Instead, a bipolar gradient is used to form a fast dephasing and rephasing of spins, during which transverse magnetization behaves similarly to an echo. The first lobe of the bipolar gradient dephases the spins by a known amount and the second opposing lobe of the gradient rephases the spins to create the “echo”. Because the effects of B_0 inhomogeneities are not reversed, the signal amplitude GE sequence decays with T_2^* . GE imaging does not require a 90° flip angle for the initial excitation pulse, allowing for selection of a flip angle which is optimized by application and anatomy of interest. A pulse sequence diagram for a typical GE sequence with conventional location encoding can be seen in Figure 1-9.

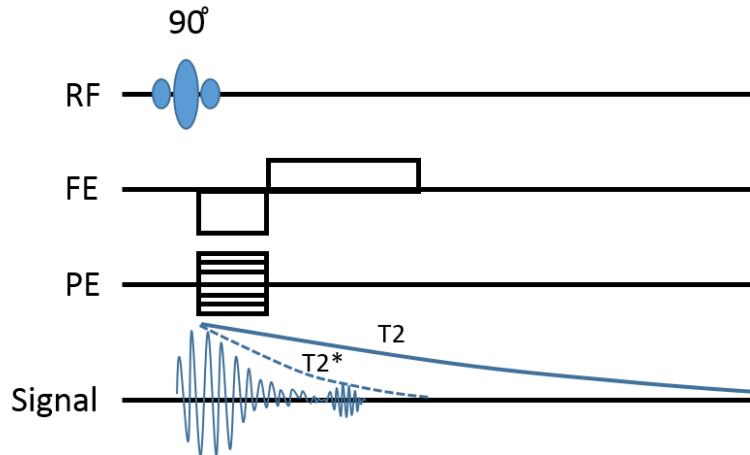


Figure 1-9: A typical GE sequence. The pulse sequence, gradient waveform, and readout signal are shown.

Pulse sequence selection is application dependent. For T_2 -weighted imaging, SE-based sequences are favorable because they can remove T_2^* effects from contrast. GE imaging is useful for fast T_1 -weighted imaging because of short TR and low flip angle. These are generalizations for the basic forms of these pulse sequences. Advanced methods of applying SE and GE have been developed which defy the limits of their general forms and extend their usefulness. For diffusion imaging, both may be used, but SE is favored for reasons described in the following section.

1.4 Diffusion Weighted Imaging

1.4.1 Diffusion Behavior

1.4.1.1 Idealized Diffusion by Random Brownian Motion

Diffusion refers to the random Brownian motion of particles driven by entropy and thermal energy. It was first described mathematically by Adolf Fick (34), who presented

diffusion as random movement of a species along a concentration gradient. This formulation is shown in Equation 9:

$$\text{Equation 9: } \frac{dc_i}{dt} = D_i \nabla^2 c_i$$

where c_i is the concentration of a molecular species within a medium and D is the diffusion coefficient with units of length squared over time. With the assumption of Gaussian diffusion behavior, the diffusion coefficient can be more simply explained with Albert Einstein's derivation in Equation 10:

$$\text{Equation 10: } D = \frac{\langle x^2 \rangle}{2t}$$

where D is the diffusion coefficient from Fick's Law, $\langle x^2 \rangle$ is the 1D mean-squared displacement of the particle from its original position, and t is the length of time over which the particle diffused. For a particle of water in a homogenous water medium at a typical human body temperature (37°C), the diffusion coefficient would be approximately $1 \times 10^{-3} \text{ mm}^2/\text{s}$ (35). Solution of Equation 10 shows that if a water molecule was allowed to diffuse in these conditions for 100 ms, it would travel an average of 25 μm from its original position. It can be seen in this derivation, that for a particle with a given diffusion coefficient, diffusion over a certain time scale corresponds with a certain "length-scale".

Einstein further described the diffusion coefficient in his Ph.D. thesis through the Stokes-Einstein equation for 3D diffusion of spherical particles in a medium. This is shown in Equation 11:

$$\text{Equation 11: } D = \frac{RT}{N_a} \frac{1}{6\pi\eta_a r}$$

where D is the diffusion coefficient, R is the gas constant, T is the absolute temperature, N_a is Avogadro's number, η_a is the viscosity of the solvent, and r is the radius of the particle. This equation implies that the diffusion coefficient of a spherical particle is dependent on three variables: temperature, particle size, and the viscosity of the solvent.

1.4.1.2 Diffusion in Tissue

In tissue, temperature is constrained to its range within the human body, which is not highly variable in term of absolute temperature. Particle size and shape are constricted to that of water. Although MRI is also sensitive to magnetization of fat, efforts are taken to remove fat contribution to signal from the contribution of water in DWI. This leaves the viscosity as the major impactor of the diffusion coefficients in MRI measurements of diffusion. However, the Stokes-Einstein equation is limited to systems of quiescent and homogenous fluid. Biological tissue satisfies neither of these assumptions.

The assumption of quiescence fails due to perfusion and bulk motion of water molecules in tissue. These perfusion effects occur on larger length-scales than pure diffusion, but are important in MRI measurement of diffusion. The assumption of homogeneity fails because of the varied makeup of biological tissue. Water molecules exists in a biological solvent of heterogeneous composition. While moving, these water molecules also may encounter barriers such as cellular membranes, organelle membranes, large macromolecules, and components of the extra-cellular matrix (ECM). Because of this, the diffusion coefficient of water in tissue is a product of complex set of interactions between water molecules and various organic structures. Previously, MRI measurements of diffusion in tissue has been shown to relate to cellularity (36), ECM composition (37), extracellular volume fraction (ECVF) (38), and membrane permeability (39). Historically, measurement of diffusion

coefficient by MRI rose in popularity because the diffusion coefficient is able to detect stroke and classify it by type after the stroke occurs (40).

1.4.1.3 Diffusion in Cancer

In cancerous tissues, rampant uncontrolled cell growth leads to high cellularity within tumors. This, among other possible factors, leads to lower mobility of water in malignant tissue and a lower diffusion coefficient in malignant tissue. An illustration showing restricted diffusion can be seen in Figure 1-10.

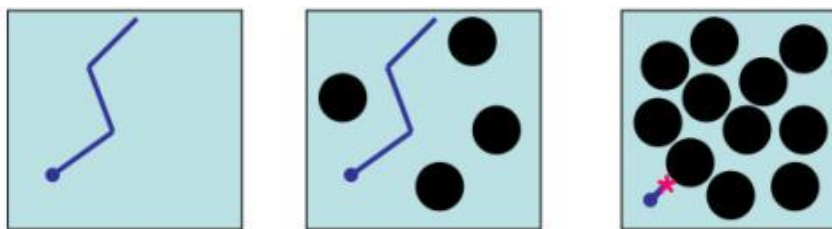


Figure 1-10: Illustrations depicting movement of water molecules within tissues of varying cellularity. The blue line represents the random movement of a water molecule while the black circles represent cells. Republished with permission (License ID = 4167471220873 from www.copyright.com) from original article: O'Flynn, E. A. M., and N. M. deSouza. 2011. Functional magnetic resonance: biomarkers of response in breast cancer. *Breast Cancer Research : BCR* 13: 204.

Because of this phenomena, images with diffusion-based contrast as well as estimated maps of the diffusion coefficient are able to delineate cancer and normal tissue. This has been used for detection (41-43), T-staging (44), and N-staging (45) for a variety of cancers.

This project is focused on changes in diffusion throughout therapy. In tumors treated with conventional therapies such as chemotherapy and radiation (and some other therapies), water mobility increases after treatment (26). It has been shown in several anatomic sites that changes in diffusion at an early time-point during treatment are different between patients who exhibit good response to treatment and patients who exhibit poor response to treatment

(29, 46-48). Patients exhibiting good response to therapy show a larger relative increase in water mobility in comparison to patients with partial or poor response to therapy.

1.4.2 Diffusion Sensitization and Contrast in MRI

The effects of diffusion are ever present in MRI images. This is because T_2 relaxation is partially mediated by diffusion. Diffusion weighted imaging (DWI) describes MRI pulse sequences and acquisition methods which attempt to isolate diffusion as a contrast mechanism. The Bloch equation (Equation 3) has previously been extended by H.C. Torrey using Fick's Law (Equation 9) to include the effects of diffusion (49). The Bloch-Torrey formulation is shown below in Equation 12:

$$\text{Equation 12: } \frac{d\vec{M}(t)}{dt} = \gamma \vec{M}(t) \times \vec{B}_{ext} - \frac{M_{transverse}}{T_2} - \frac{M_{parallel} - M_{0,parallel}}{T_1} + D \nabla^2 \vec{M}(t)$$

where $M(t)$ is the time-variant magnetization of the sample, γ is the gyromagnetic ratio, B_{ext} is the vector of combined external magnetizations, $M_{transverse}$ is the magnetization of the sample transverse to B_0 , T_2 is the spin-spin relaxation constant, $M_{parallel}$ is the magnetization of the sample parallel to B_0 , T_1 is the spin-lattice relaxation constant, and D is the diffusion coefficient. If relaxation terms are neglected, only transverse magnetization is considered, and only gradient fields are accounted for, Equation 13 will result:

$$\text{Equation 13: } \frac{d\vec{M}(\vec{r}, t)}{dt} = -i\gamma(\vec{r} \cdot \vec{G}(x))\vec{M}(\vec{r}, t) + D \nabla^2 \vec{M}(\vec{r}, t)$$

where $M(r,t)$ is the time and space dependent magnetization of the sample, γ is the gyromagnetic ratio, the vector r is the spin location, $G(x)$ is the external gradient as a

function of location, and D is the diffusion coefficient. For the case in which $D=0$, integration of Equation 13 will result in Equation 6.

The method of diffusion sensitization which is used in MRI was initially demonstrated by E.O. Stejskal and J.E. Tanner in an NMR experiment (50). Their method takes the form of a pair of pulsed gradients (known as Stejskal –Tanner gradients) the first of which encodes spin location with dephasing throughout the sample and the second of which provides the opposite location-based rephasing, forcing realignment of non-moving spins. Spins that experience these gradients at different positions will not experience complete rephasing. This causes dephasing of the spins and a loss of signal in areas where spins are mobile. The pulsed gradient SE (PGSE) implementation of Stejskal-Tanner gradients as well as an illustration of its effects on moving and non-moving spins can be seen in Figure 1-11.

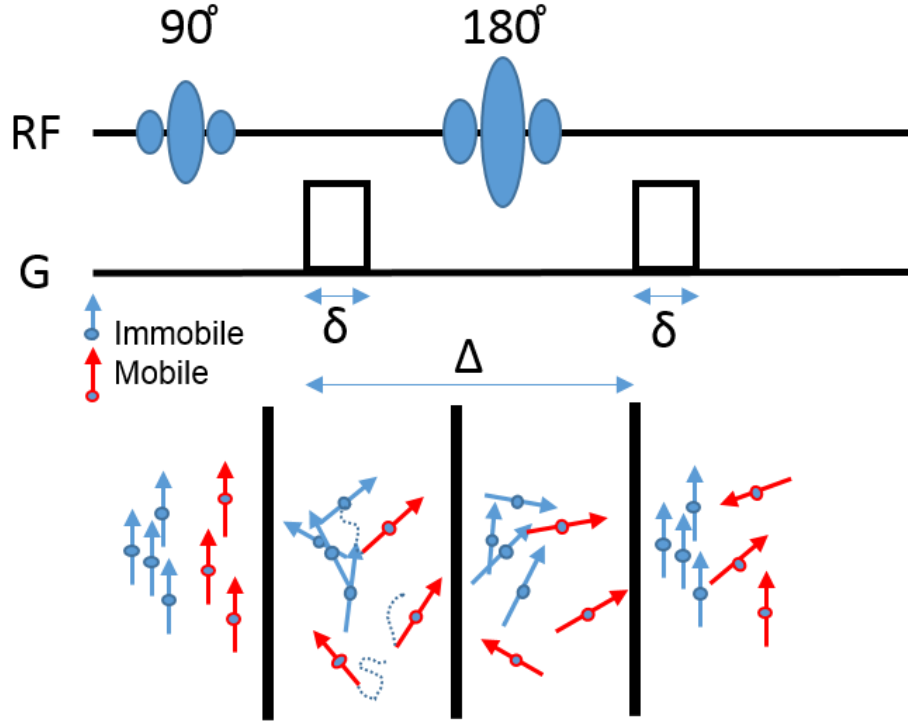


Figure 1-11: Pulse sequence and simplified gradient waveform for a PGSE sequence. The gradient duration is labeled with δ , and the time elapsed between gradient is labeled with Δ . At the bottom, an illustration is shown comparing dephasing and rephasing effects on the magnetization of moving (red) and non-moving (blue) spins.

The effects of these diffusion sensitization gradients on MR signal can be described through further derivation of Equation 13. Stejskal and Tanner used bounds imposed by their pulsed gradient experiment to derive a mono-exponential model of diffusion from the Bloch-Torrey equation, which is shown in Equation 14.

$$\text{Equation 14: } \frac{S}{S_0} = e^{-D \gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3})}$$

where S is the signal level of the diffusion images, γ is the gyromagnetic ratio, G is the magnitude of a single gradient, δ is the duration of a single gradient pulse, and Δ is the time lapse between the pulsed gradients. This derivation shows that the degree of signal loss is

dependent upon several variables. The variable of interest is the diffusion coefficient of the tissue. The remainder of the variables affecting the degree of signal loss include the strength of the location encoding gradients (magnitude), the duration of the location encoding pulse, and the time between the pulses (during which spins are allowed to move). These variables are grouped into a single parameter known as the b-value (alternatively as the b-factor). The formulation for the b-value is shown in Equation 15:

$$\text{Equation 15: } b = \gamma^2 G^2 \delta^2 (\Delta - \frac{\delta}{3})$$

where γ is the gyromagnetic ratio, G is the magnitude of a single gradient, δ is the duration of a single gradient pulse, and Δ is the time lapse between the pulsed gradients. Diffusion images acquired with a certain b-value are referred to by bVALUE (Ex: Images acquired at $b = 200 \text{ s/mm}^2$ are referred to as b200 images). The maximum b-value an MRI scanner can produce with constant Δ is limited by its gradient hardware. It should be noted that identical b-values might be achieved with different combinations of gradient pulse areas, gradient pulse durations, and diffusion times. For the purposes of the experiments in this thesis, b-value is assumed to be the only technique factor impacting signal, although this is not necessarily true (51).

Other sources of contrast are minimized in diffusion scans by maintaining TE and TR across diffusion images with different b-values. This helps to minimize deviation from the assumptions made in Equation 13.

1.4.3 Models of Diffusion in MRI

1.4.3.1 Mono-Exponential Model

The simplest model of diffusion, the mono-exponential model (Equation 14), is the product of direct derivation of the Bloch-Torrey equation. Even if anatomic motion is ignored, the diffusion coefficient from this equation is the result of a complex combination of spin motion phenomena which are interrogated with Stejskal-Tanner gradients. In acknowledgement of these complex phenomena underlying this measurement, the mono-exponential diffusion coefficient is referred to as the apparent diffusion coefficient (ADC).

Violations of the assumptions used to derive Equation 14 may be included in more complex diffusion models. This is done both for the purpose of better estimating the true diffusion coefficient as well as identifying other factors affecting diffusion contrast separate from initial assumptions. In order to separate these other factors from the diffusion coefficient, b-values must be chosen which characterize spin motion on two separate ranges length-scales. These two ranges of b-values must be chosen so that in one range, the diffusion coefficient is the dominant contrast mechanism and in the other range, signal levels are impacted by both the diffusion coefficient and the other factor. Mathematical comparisons of signals over these two range of b-values will then allow for isolation of the diffusion model parameters. At low b-values, perfusion of microvasculature causes a deviation from mono-exponential diffusion behavior shown as signal loss beyond that expected of diffusion alone. At high b-values, deviations from the assumption of Gaussian diffusion, or kurtosis, cause an opposite deviation in signal where measured signal is higher

than what is expected from diffusion alone. A plot detailing these effects can be found in Figure 1-12.

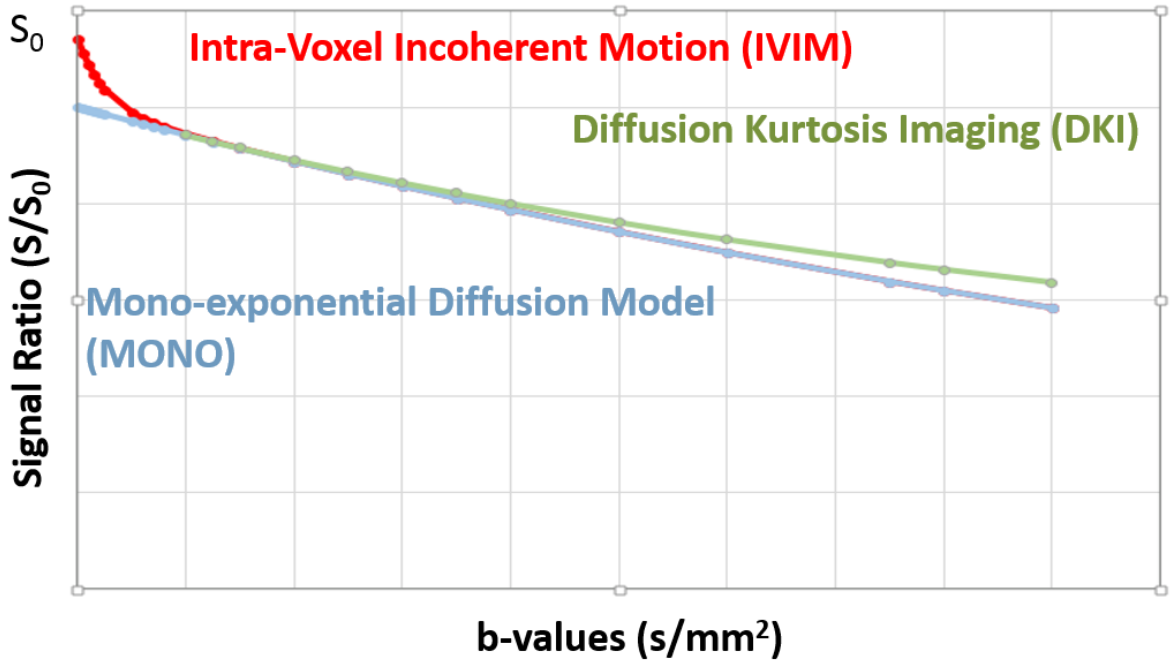


Figure 1-12: Graphical depiction of deviations in signal levels due to perfusion effects (red) at low b-values and from kurtosis effects (green) at high b-values.

1.4.3.2 Intra-voxel Incoherent Motion Model

In the case of lower b-values (tissue dependent, typically in the approximate range of 0 to 300 s/mm²), larger length-scales of motion are interrogated and signal loss is due to a combination of diffusion and perfusion effects. Perfusion stems from microcirculation in capillaries within a given voxel, and is the dominant mechanism of signal loss in this range of b-values. Signal loss from perfusion will reach a limit, and at higher ranges of b-values will not further impact signal loss. A modeling method for perfusion effects, originally described by Denis Le Bihan, is the intra-voxel incoherent motion (IVIM) model (52). In this model, it is assumed that the orientation of capillaries within a given voxel is random, and

movement of fluids through these voxels mimics the random motion of diffusion on a larger length-scale. A visual depiction of this random capillary orientation can be seen in Figure 1-13.

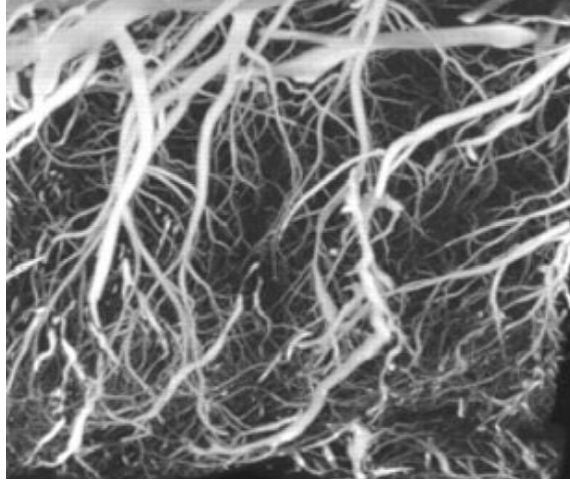


Figure 1-13: Random orientation of capillaries within a hypothetical voxel of tissue. Republished with permission (License ID = 4167550994385 from www.copyright.com) from original article: Malyar, N. M., M. Gossel, P. E. Beighley, and E. L. Ritman. 2004. Relationship between arterial diameter and perfused tissue volume in myocardial microcirculation: a micro-CT-based analysis. *American journal of physiology. Heart and circulatory physiology* 286: H2386-2392.

On this basis, contribution of perfusion motion to signal loss may be simply modeled as a second exponential, whose contribution to signal loss is limited to a maximum effect reached at low b-values. This IVIM model is stated in Equation 16.

$$\text{Equation 16: } \frac{S}{S_0} = (1 - f)e^{-b \cdot D} + fe^{-b \cdot D^*}$$

where f is the perfusion fraction, D is the true diffusion coefficient, and D* is the pseudo-diffusion coefficient representing the b-value dependent contribution of perfusion to signal loss. In general, this D* coefficient is an order of magnitude larger than D coefficient (52).

This model has been used for application in cancer with varying success (43, 53, 54). When selecting b-values for this exam, it is important to select multiple low b-values to fully characterize the perfusion effects on signal loss (55).

1.4.3.3 Diffusion Kurtosis Model

When diffusion is probed on a small length-scale, barriers such as cell membranes pose a greater effect on the diffusion probability distribution (Equation 10) of water molecules. This is because these barriers block a significant portion of nearby space (relative to length-scale). The presence of barriers of any shape will alter the diffusion probability function so that shorter diffusion lengths are more likely. This is a violation of the assumption of Gaussian diffusion behavior, and results in a compression of the ideal Gaussian distribution of the diffusion probability density function. This means reducing the probability of values in the tails and increasing the probability of values close to the mean, as shown in Figure 1-14.

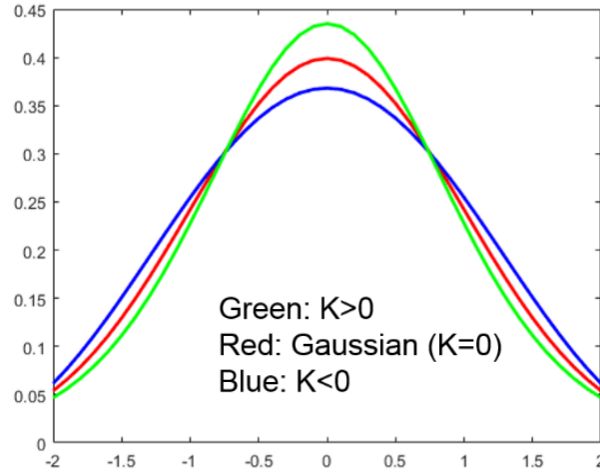


Figure 1-14: Depiction of probability function distributions with differing kurtosis. The kurtosis observed in diffusion is positive, which is depicted as a change in the probability density function from the red plot to the green plot.

This change in the distribution of the diffusion probability function can be described as a Gaussian function altered with a kurtosis parameter (56). The kurtosis of a distribution is the fourth moment mathematically, and it can be calculated for diffusion using Equation 17:

$$\text{Equation 17: } k(x) = \frac{\mu^4}{\sigma^4} = \frac{\langle (n \cdot s)^4 \rangle}{\langle (n \cdot s)^2 \rangle^2}$$

where k is the kurtosis parameter, μ is the average diffusion distance, σ is the standard deviation of the probability function, n is the unit vector and s is the diffusion distance vector. If this kurtosis term is incorporated in the derivation of Equation 14, Equation 18 will result (56):

$$\text{Equation 18: } \frac{S}{S_0} = e^{-b \cdot ADC + \frac{1}{6} b^2 ADC^2 k}$$

Examination of this equation shows an intuitive mathematical relationship between signal levels and the kurtosis effect. As length-scale is decreased (b-value increases), the

effects of kurtosis (barriers blocking diffusion) is increased (second order relationship to b-values). This reduces signal loss from diffusion, as molecules are less likely to diffuse as far. This can be seen in Figure 1-12, where kurtosis effect is evident at higher b-values (1000 s/mm² and upwards). Measurement of the kurtosis parameter is shown to be useful for various applications in cancer at several anatomic sites (57-59).

1.4.4 Advanced MRI Techniques for DWI

Many advanced MRI techniques have been developed and implemented since the birth of MRI. This section aims to explain a few important techniques which have been used in this project and are essential for clinical implementation of DWI. Several of these techniques provide acquisition acceleration. This acceleration serves to improve reproducibility and image quality of diffusion images and lower scan times. Another important technological development is the increase of gradient strength. As recently as 25 years ago, maximum b-values achievable on a clinical scanner were on the order of 200 s/mm² (51). The development of these techniques and gradient technology has been essential for clinical implementation of fast and reproducible diffusion imaging.

1.4.4.1 Echo Planar Imaging

Echo Planar Imaging (EPI) was first described by Peter Mansfield in 1977 (60). EPI is an acquisition in which all or most of k-space is collected in a single TR. Single-shot EPI (ssEPI) describes EPI in which the entirety of k-space is collected in one TR, whereas multi-shot EPI describes the acquisition of k-space within a few TRs. Within this paper, all DWI images were taken using ss-EPI techniques. The standard method of covering k-space is a raster navigation, where lines of k-space are acquired in a back-and-forth scanning trajectory.

In this method, G_x (which is referred to as the frequency encoding gradient in EPI) is used to provide continuous back-and-forth movement between the edges of k-spaces. G_y (referred to as the phase encoding gradient in EPI) is blipped to switch between phase encoding lines. The EPI acquisition is typically centered on the echo, so that the center of k-space is acquired with maximum signal. A diagram of k-space navigation by this method and the EPI gradient waveform can be found in Figure 1-15.

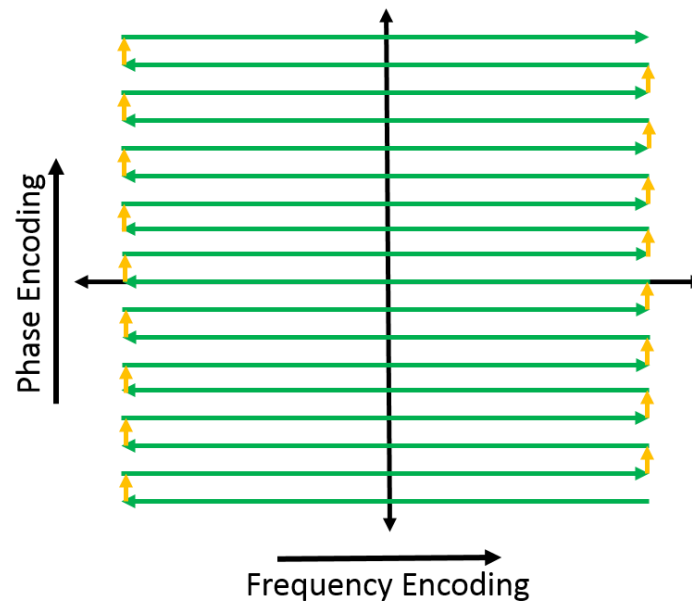


Figure 1-15: EPI k-space trajectory where the phase encoding direction is blipped, and the frequency encoding direction is continuous.

EPI offers several advantages over conventional imaging, especially for DWI. The speed of imaging offered by EPI minimizes readout time and nearly eliminates motion artifacts. EPI also is advantageous for DWI because it enforces consistency in TE/TR across diffusion images, minimizing sources of contrast besides spin mobility.

1.4.4.1 Multi-slice Acquisition

Multi-slice acquisition uses composite RF pulses to excite multiple slices simultaneously. A delay is introduced in between excitation of different slices to create a phase difference in excited signal of different slices. These signals can then be separated in post-processing of the acquired signal to create images of separate slices (32). Because multiple slices can be acquired simultaneously, the scan time of sequences can be greatly reduced.

1.4.4.2 Parallel Imaging

Parallel imaging is a technique which accelerates acquisition by taking advantage of redundancy in k-space. In this technique, k-space is symmetrically under-sampled in the phase encoding direction. This under-sampling results in a reduced field of view (FOV) in uncorrected images, which is accompanied by wrap-around aliasing artefacts. Parallel imaging uses sensitivity maps from receiver coils as a separate source of location information to resolve aliasing and geometric distortion, as well as restore FOV. This correction may be applied through different techniques in both image space and k-space. Notable methods of performing these corrections include SENSE (61) and GRAPPA (62). Parallel imaging typically provides acquisition accelerations on the order of 2-4x.

1.4.4.3 2D Selective RF Excitation

A more recently developed technique utilized in the experiments of this thesis is 2D selective RF excitation by the method of Saritas et al. (63). Several methods of 2D RF excitation have been described previously, but this particular technique is superior because of

its compatibility with adjacent slice multi-slice acquisition. In this technique, an echo-planar blipped RF is applied within “excitation k-space” to excite a 2D slab of tissue. Excitation k-space is analogous to image space in the frequency encoding or slice selection direction and cyclic in the phase encoding or slab selection direction. The phase encoding direction of excitation k-space creates a phase separation of water and fat due to chemical shift as phase evolves during the excitation period.

This technique is advantageous for several reasons. The primary reason is because it is able to reduce the number of phase encoding lines during excitation without compromising the resolution. Shortened EPI readout has the benefit of reduced image distortion and artefacts. This shortened phase encoding direction must be applied on a “short” dimension of the organ of interest. In the case of tubular organs, such as the spinal cord and esophagus, two short dimensions are available. Typically, the AP direction is chosen to be shortened because it is most impacted by breathing motion. The second advantage of this technique is that it produces an intrinsic water-fat separation during phase encoding of excitation k-space. Fat separation is important for implementation of DWI, and in DWI scans which do not incorporate 2D selective RF excitation, other methods must be utilized to remove fat contribution from signal.

1.5 Hypothesis

Two-week changes in quantitative DWI parameters are predictive of the response to neoadjuvant chemoradiation in esophageal cancer.

1.6 Specific Aims

Specific Aim I: To determine how early changes in DWI parameters in patients of esophageal cancer treated with chemoradiation correlate with measures of clinical outcome.

Specific Aim II: To evaluate reproducibility of ADC measurements between readers for both volumetric and slice contouring methods.

2 Materials and Methods

2.1 Patient Selection and Outcome Measures

Patients afflicted with esophageal cancer between stages IIa and IIIb were enrolled into this ongoing prospective study starting in 2014. These patients were given standard of care treatment in the form of nCRT followed by surgical resection. Histopathologic analysis of resected tissue provided measures of % viable cells. Based on this measure, patients were classified into TRG's according to the four-tier grading system. Disease progression after surgery was also tracked through follow-up for the purpose of eventual analyses of survival.

Patients who deviated from the study protocol were considered inevaluable, and additional patients were enrolled. This study was approved by the MDACC IRB and patient consent was obtained in all cases.

2.2 MRI Scan Protocols

MR exams for each patient were performed at three points along the course of nCRT: before nCRT at baseline (BL), two-weeks into nCRT at interim (IM), and after completion of nCRT at first follow-up (FU). All scans were performed on a single GE 3.0 T MR750W scanner using a 32-channel phased-array body coil. At the start of this study, two separate DWI protocols were included in MR exams: a standard ssEPI diffusion scan for fitting with a mono-exponential model, and an rFOV scan by FOCUS (2D RF excitation method on GE platform) for fitting with an IVIM model. For the IVIM scan FOCUS protocol, the AP direction was chosen as the reduced phase encoding dimension. This dimension was reduced by half, resulting in a rectangular FOV. Later, in January 2016, another ssEPI scan with

higher b-values was added to the study for diffusion kurtosis analyses. Henceforth, these scans will be referred to by the diffusion model they were designed for: MONO, IVIM and DKI, respectively. For MONO and DKI scans, acceleration by parallel imaging was utilized. The typical acquisition parameters for these three scans are shown in Table 2-1.

Table 2-1. Typical scan parameters for DWI scans.

Protocol	# of Slices	TE (ms)	TR (ms)	Thickness (mm)	Matrix	FOV (cm)	Approx. Scan Time (m:s)
MONO	50	69.7	5195	4.0	96 x 96	46 x 46	4:41
IVIM	16	62.4	3000	4.0	96 x 96	24 x 12	5:38
DKI	40	52.7	4500	4.5	96 x 96	42 x 42	6:21

For each DWI scan, diffusion images were taken using a distinct selection of b-values. For the MONO scan, b-values were selected based on the methodology of Padhani et al. (26). This methodology requires the selection of three b-values. First, a b₀ image is acquired for reference to T₂ images. Second, a mid-range b-value (between 200 and 500 s/mm²) is selected, which will aid ADC fitting by measuring signal loss at a lower b-value. The exact value of this mid-range b-value is specific to the anatomic site being imaged, and must be high enough to avoid variation in perfusion effects on signal loss. In the esophagus, a b-value of 200 s/mm² is sufficient for this purpose. Finally, a third high b-value must be included. This is to further characterize signal loss from diffusion in the mid b-value region where signal loss is dominated by diffusion. For the esophagus, a value of 800 s/mm² was chosen.

For IVIM scans, several low and intermediate b-values were chosen to separately characterize diffusion and perfusion effects. For DKI, several intermediate and high b-values were chosen to separately characterize Gaussian diffusion and diffusion impacted by kurtosis. Higher b-value diffusion images were acquired with multiple excitations (NEX, number of excitations) and averaged. The benefit of NEX is increased SNR. SNR increased with the square root of the number of NEX taken. B-values and their respective NEX for each of the diffusion scans are presented in Table 2-2.

Table 2-2. B-values and NEX for diffusion scans.

Protocol	b-values (s/mm²) [NEX]
MONO	0 [2] , 200 [2], 800 [6]
IVIM	0 [2], 30 [2], 60 [2], 100 [4], 600 [16]
DKI	200 [2], 600 [4], 1000 [8], 1200 [12], 1600 [16]

For each scan, diffusion sensitization gradients were applied in three directions for each non-zero b-value, and averaged to provide combined diffusion images. All diffusion scan were acquired with free-breathing of the patient. Neither co-registration nor geometric correction was applied to diffusion images. Although these corrections would provide benefit by minimizing distortions and misalignment between diffusion images due to motion, there is evidence that non-linear effects on signal level from these corrections leads to incorrect estimation of diffusion parameters. An experiment performed in-house using standard correction methods revealed a change of up to +/-5% in ADC values for MONO scans.

2.3 Diffusion Parameter Mapping

Diffusion parameter map estimation was performed using an in-house software named ImageI, which was designed and developed in our own lab. This software was implemented in the Matlab application development environment (Mathworks, Natick, MA). ImageI offers a convenient environment and a variety of useful features for the work performed in this project. Its capabilities will be expounded in the relevant methods sections of this thesis. For fitting of parameter maps, ImageI uses slightly differing methods of non-linear least mean square (LMS) curve fitting for each model. Calculating ADC is accomplished by finding the LMS solution for ADC on a voxel-by-voxel basis using all three b-values. In ImageI (Matlab), this is implemented using the *pinv* function, as shown in the pseudo-code of Equation 15.

Equation 19: Pseudo-code for Mono-exponential ADC Estimation

```
#Calculate signal ratios in log(S/S0) form  
  
Diff_image = diffusion_images(b0,b200,b800)./diffusion_images(b0_meas)  
Diff_image = log(Diff_image)  
  
#Solve inverse squares problem to estimated ADC  
  
pb = pinv(-200,-800);  
  
For All Voxels  
    ADC = pb*Diff_image(b200,b800)  
  
End
```

In the case of fitting the MONO diffusion images with a mono-exponential model, ImageI was validated against an industry standard available in the Functool application of the AW workstation by GE (Boston, MA). For the specific case of measuring ADC mean across volume contours of esophageal cancer, the GE and ImageI methods of producing ADC maps have been shown to be equivalent. Further details on this comparison can be found in the Appendix 5.1.

A second method of ADC calculation described by Padhani et al. was compared with the ImageI and industry-validated methods (26). The Padhani et al method does not use b0 images for ADC estimation and instead directly calculates ADC from the solution of b200 and b800 signals. ADC estimation in ImageI was also another standard method of ADC calculated by b0 and b800 is also included. The results of this comparison are shown in the Appendix 5.6. A graph of the curve fitting of MONO signal levels for an individual voxel can be seen in Figure 2-1.

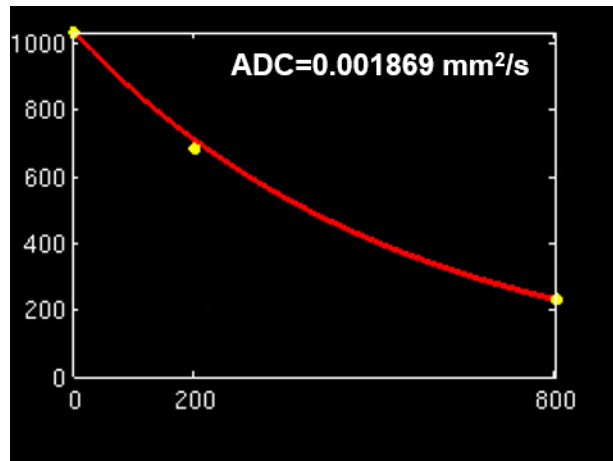


Figure 2-1: Comparison of LMS fitted ADC plot with plotted measured signal values for a single voxel.

Industry standards for fitting IVIM and DKI models are not yet available from GE AW Functool. Because of this, the ImageI methods for estimating these parameter maps has not been validated against an industry standard. In the case of IVIM, a variety of parameter fitting methods have been described in literature (64). The method used by ImageI is well-established (65). In this method, an initial guess for the true-diffusion coefficient, D , is calculated from a direct solution of signal levels from two diffusion images. From this initial ADC estimate, an estimation of S_0 is then calculated. Then, the difference between the estimated S_0 and measured S_0 is divided by the measured S_0 to create an initial guess for the perfusion fraction, f . These initial guesses were combined into the LMS algorithm for the estimation of IVIM parameter maps. During LMS estimation, S_0 signal levels were allowed to iteratively change from the measured S_0 signal. This flexibility allows for better fitting of lower b-value signals with the IVIM model. Pseudo-code for this non-linear LMS implementation is shown in Equation 20. A graph of the curve fitting of IVIM signal levels for an individual voxel can be seen in Figure 2-2.

Equation 20: Pseudo-code for IVIM Parameter Mapping

#Calculate signal ratios in $\log(S/S_0)$ form

Diff_image = diffusion_images(b0,b30,b60,b100,b600)./diffusion_images(b0_meas)

Diff_image = log(Diff_image)

#Use LMS to solve for ADC

For All Voxels

*parameter_holder=lsqcurvefit(@(f)((1-f)exp(-b.*D)+(f)exp(-b.*D*)),initial_guess,Diff_image)*

End

#Where parameter_holder is a 4D matrix containing 3D maps of perfusion fraction, pseudo-diffusion coefficient, true-diffusion coefficient, and final b0 estimation.

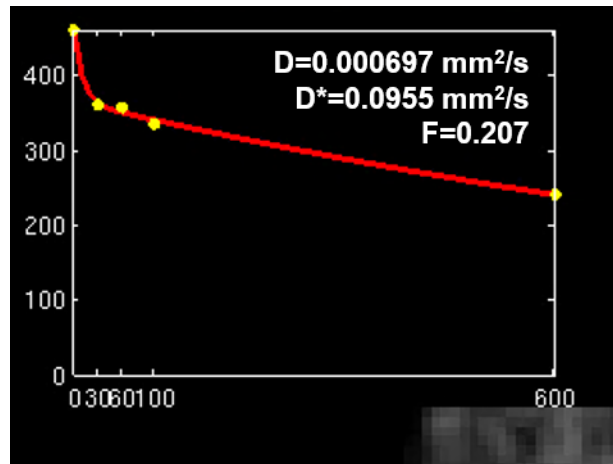


Figure 2-2: Comparison of LMS fitted IVIM model with measured signal values for a single voxel.

For DKI parameter fitting, initial estimates of ADC and S_0 were also obtained using b_{200} and b_{600} images. These estimates were used in a non-linear LMS estimation

implementation similar to IVIM, as shown in Equation 21. A graph of the curve fitting of DKI signal levels for an individual voxel can be seen in Figure 2-3.

Equation 21: Pseudo-code for DKI Parameter Mapping

#Calculate signal ratios in $\log(S/S_0)$ form

Diff_image = diffusion_images(b200,b600,b1000,b1200,b1600)./diffusion_images(b0_est)

Diff_image = log(Diff_image)

#Use LMS to solve for ADC

For All Voxels

*parameter_holder=fminsearch(@(sum(Diff_image,exp(-b.*ADC+b².*ADC².*k)),initial_guess,Diff_image)*

End

#Where parameter_holder is a 4D matrix containing 3D maps of ADC and k.

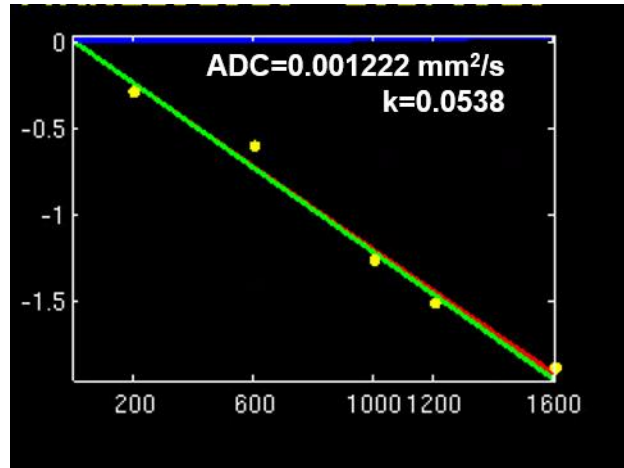


Figure 2-3: Comparison of LMS fitted DKI model with measured signal values for a single voxel. A plot of the kurtosis model is seen in red, while a plot of the DKI-calculated ADC as a mono-exponential decay is seen in green. Deviation of the signal from mono-exponential behavior can be seen at high b-values.

Across all models, some numeric bounds were used to mask out non-physical values for diffusion parameters. For the diffusion parameters ADC, D^* , D , and k , negative values were excluded. For the perfusion fraction f , negative values or values above 1.0 were excluded. No other exclusion of diffusion parameters was made based on their quantitative values.

ImageI offers fast, automated import of DICOM files followed by automatic calculation of parameter maps. On a multi-core external server, estimation of parameter maps of scans with parameters matching those in Table 2-1 took approximately 2 minutes for MONO scans, 30-40 minutes for IVIM scans, and 1.5-2 hours for DKI scans. After estimation of parameter maps, ImageI creates a single database file (.dat) which contains diffusion images, parameter maps, patient information, and a reserved space for contour masks.

2.4 Contour Measurements of Diffusion Parameters

To measure diffusion parameter values within each tumor, contours delineating the tumor from surrounding tissue were produced. Besides the technology used to make these contours, a few strategies to ensure consistency and accuracy of contours were developed.

2.4.1 General Contour Strategies

Across the contouring methods used in this paper, a few general strategies were used. First, all contours were performed on b200 images for MONO and DKI, and b100 for IVIM. This was chosen out of the possible options of contouring on any of the diffusion images or the ADC map. The decision to avoid contouring on b0 images was simple: although b0 images have the highest signal, they do not contain diffusion contrast needed to delineate tumors. Higher b-value images and ADC maps were not chosen as contouring images because although they are highly tumor specific, heterogeneity within the tumor may lead to high variations in high b-value diffusion images and ADC maps. This heterogeneity could cause readers to contour only certain sections of the tumor with more hindered diffusion. For volume contouring, it is important to include all voxels of the tumor so as to completely capture the diffusion behavior throughout. Another simple and more practical reason was included in the decision to contour on lower b-value diffusion images: at later stages of treatment, diffusion becomes less hindered. This reduces signal in the higher b-value diffusion images, making contouring more difficult by manual and semi-automatic contouring methods. By using a lower b-value, greater ease of contouring on IM and FU scans was achieved.

A second contouring strategy was introduced for the purpose of delineating the superior and inferior (SI) extents of the tumor (proximal and distal by GI anatomy). For contouring on diffusion images, the poor resolution in the slice dimension (4-4.5 mm in these scans) makes location of the superior and inferior slices of the tumor difficult. This difficulty is most severe at the inferior extent in tumors at the GE junction, where T₂ shine-through of cardiac tissue is a large issue. To standardize contouring measurements in this dimension, a method was created for finding this superior and inferior extent of the tumor. In the case of using this method on MONO scans, the b200 map is first used to locate the tumor. Then, approximate superior and inferior boundaries based on b200 images are found. After this, the reader will switch to ADC maps on a nearby slice where tumor location is certain. The dark area of the ADC map corresponding to the area of the tumor is then located. This dark region of the ADC maps is traced through slices until it disappears. This disappearance is far more distinct than SI changes in the b200 diffusion images, and marks the SI boundaries of the tumor. These boundaries can then be propagated to b200 diffusion images for contouring. Visualization of this phenomena on MONO exams is shown in Figure 2-4.

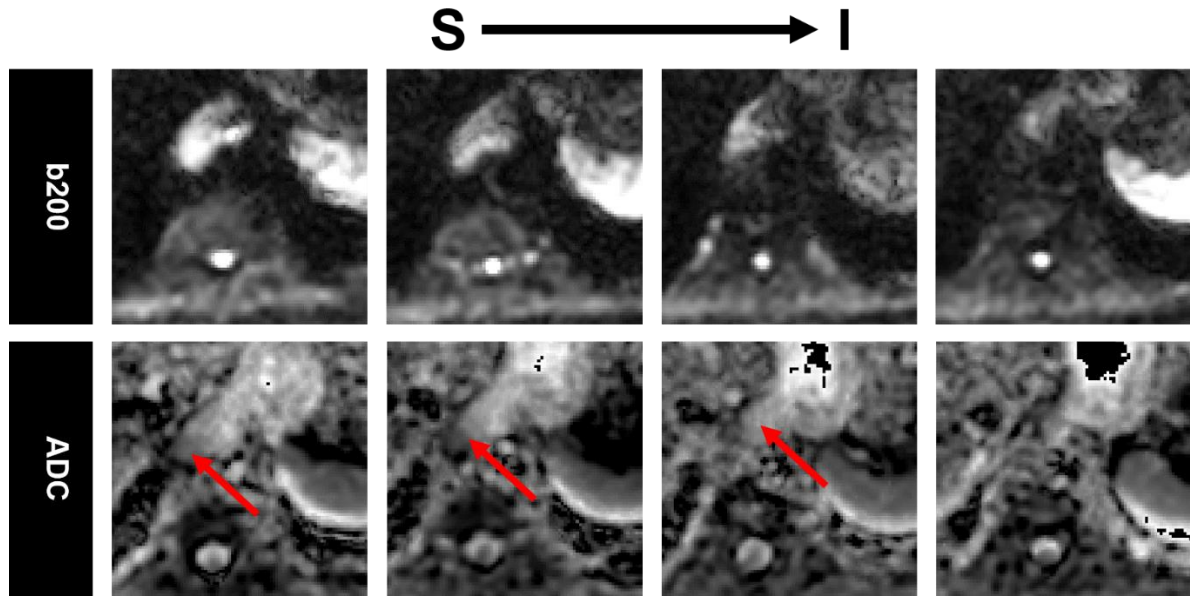


Figure 2-4: This figure demonstrates the indistinct SI boundaries in the b200 images and the distinct SI boundaries in ADC maps. It can be seen that the b200 images demonstrate a noticeable, but indistinct lowering of signal as slices are sequentially examined in the inferior direction towards the inferior boundary of the tumor (upper row, left→right). The ADC map (bottom row) demonstrates a distinct inferior boundary when examined in this manner, which can be noticed as a large rise in ADC values between the second and third image. Red arrows in the leftmost images mark areas of low ADC corresponding to tumor, and the red arrow in the third image marks where this area of low ADC is missing.

A third contouring strategy was added to account for cases in which tumor signal dropout occurs on a single slice on the diffusion images of the b-value used for contouring. This “flashing” artefact may occur because of large motion or single slice failure of parallel imaging correction. Flashing artefacts are easy to notice by the readers. When encountered, the entire slice can be excluded to prevent skewing the estimation of the tumor ADC. Images of this artefact and others can be found in the Appendix 5.4. Other b-value images were not searched for flashing artefacts so as to minimize time needed for contouring.

A fourth contouring strategy was developed to deal with misalignment of diffusion images due to motion and distortion. The method is simple: readers would slightly under-

contour the edges of tumors. This prevents the inclusion of the incorrectly estimated ADC values present at tumor borders misaligned between diffusion images.

The last notable contouring strategy used in this research was to exclude the lumen of the esophagus. ImageI has a salient feature to allow “holes” in the images to be excluded within contours. This exclusion allows further implementation of the under-contouring method around the interior edges of the tumor. Morphologic distention within the esophagus may also trap fluids within the lumen. Inclusion of voxels containing trapped liquid would not be representative of the tumor and would artificially raise ADC summary measures extracted from the contour. An example of lumen exclusion can be seen in the results section (Figure 3-2).

2.4.2 Software Implementation

Contour measurements were performed across MONO scans of all surgical patients using two distinct methods: volumetric contouring and slice contouring. Volumetric contouring is only available through research tools. In the current clinical setting of MDACC, only slice contouring is currently available through the PACS system.

2.4.2.1 Volumetric Contouring

Volumetric contouring was accomplished through the combined use of two software: ITK-SNAP (www.itksnap.org) (66) and ImageI. Although ImageI is initially used for DICOM import and parameter map estimation, primary delineation of the tumor was performed in ITK-SNAP.

ITK-SNAP is an open-source, semi-automatic contouring suite. It offers user-guided semi-automatic contouring by the level-set algorithm, among several other algorithms (66).

Contouring by the level-set algorithm is advantageous for diffusion images because the level-set algorithm is robust to high signal variation (66). This is especially important for propagation of contour boundaries in the low-resolution slice direction of diffusion images.

Within ITK-SNAP, DICOM files of the diffusion images used for contouring (EX: b200 for MONO) were imported. Contours were then produced using the Active Contour Segmentation option. For pre-segmentation, the Thresholding method was used, with a maximal selected upper threshold and a user-prescribed lower threshold. From here, several 3D “bubbles” were placed as seeds from which the initial iteration of the contour would propagate. Bubbles were selectively placed to ensure propagation of contour throughout all sections of the tumor. After this, contour evolution was executed and then terminated when the tumor was encompassed. Usually, this resulted in a useable contour. A full guide with screenshots visually detailing the steps of this method is shown in the Appendix 5.2. Once a useable contour was derived using ITK-SNAP, a binary mask was exported in DICOM format and then imported into ImageI.

Fine manual adjustments and finalizing of the contour were performed after the initial delineation by using the toolbox available within ImageI. It provides paintbrush and eraser tools, as well as polygon and freehand semi-automated contouring by the Snake algorithm. ImageI also allows for histogram slider exclusion of signal values across tumor volume or individual slices. With these tools, contours were finalized by erasing areas where the semi-automatic contour had propagated into adjacent tissues and by smoothing the edges of the contours to fit the tumor boundaries.

After contouring was completed, extraction of diffusion parameters was executed using ImageI. This included the following summary measures across volume extracts of

diffusion parameters: max, min, mean, standard deviation, median, histogram kurtosis, histogram skewness, 10th percentile, 25th percentile, 75th percentile, and 90th percentile. ImageI also extracted the length of contoured SI extent and contoured tumor volume. It should be noted that in this instance, histogram kurtosis examines the kurtosis of diffusion parameter levels across a histogram of voxels in the contour. This should not be confused with the kurtosis effect observed in high b-value diffusion imaging.

A separate extraction of all diffusion parameter summary measures was made for the largest single slice within the contour. This was done with the intent of acting as a proxy for single slice contouring. Comparison of largest slice metrics with actual single slice contouring can be found in the Appendix 5.7.

All extracted measures were combined and exported into an Excel file using a single button click within ImageI. A full workflow map of all steps between acquisition and extraction by volume contour can be found in Figure 2-5.

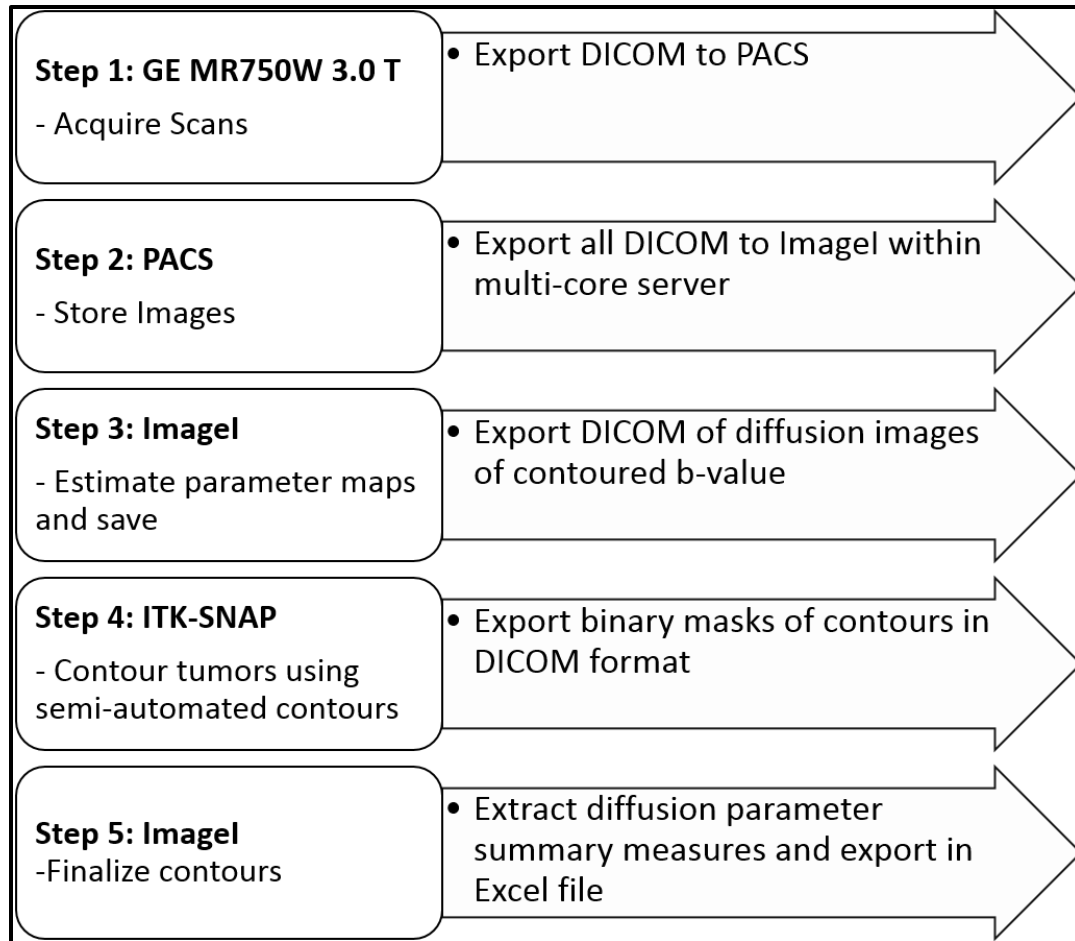


Figure 2-5: Workflow for measurement of diffusion parameters through volume contouring.

It is important to note that contouring performed by the initial reader (the author of this thesis) was not performed using ITK-SNAP. The previously described contouring method was only used by the readers in the inter-reader study. Contouring by the initial reader (R1) was performed manually using the contouring toolbox within ImageI. Further details on the strategies of their implementation is presented in the Appendix 5.3.

2.4.2.2 Slice Contouring

The implementation of slice contouring is more simplistic than that of volume contouring. First, ADC maps generated by GE Functool using diffusion images were exported to PACS. Then, using localizer mode within Phillips Intellispace, freehand contours were drawn onto the ADC map while tracing the b200 image. From here, ADC mean and standard deviation of the contour were recorded. This measurement was only performed on a single slice. Slice selection was left to each reader, whose only guidance in this selection was to select a slice which was at the midline of the tumor and judged to have the largest cross section. With this slice contouring method, readers were unable to exclude the esophageal lumen.

Within Phillip Intellispace, freehand contours were saved as presentation states and then reviewed later to insure correct recording of ADC mean and standard deviation. A full workflow map for estimation and measurement of diffusion parameters by slice contouring can be found in Figure 2-6.

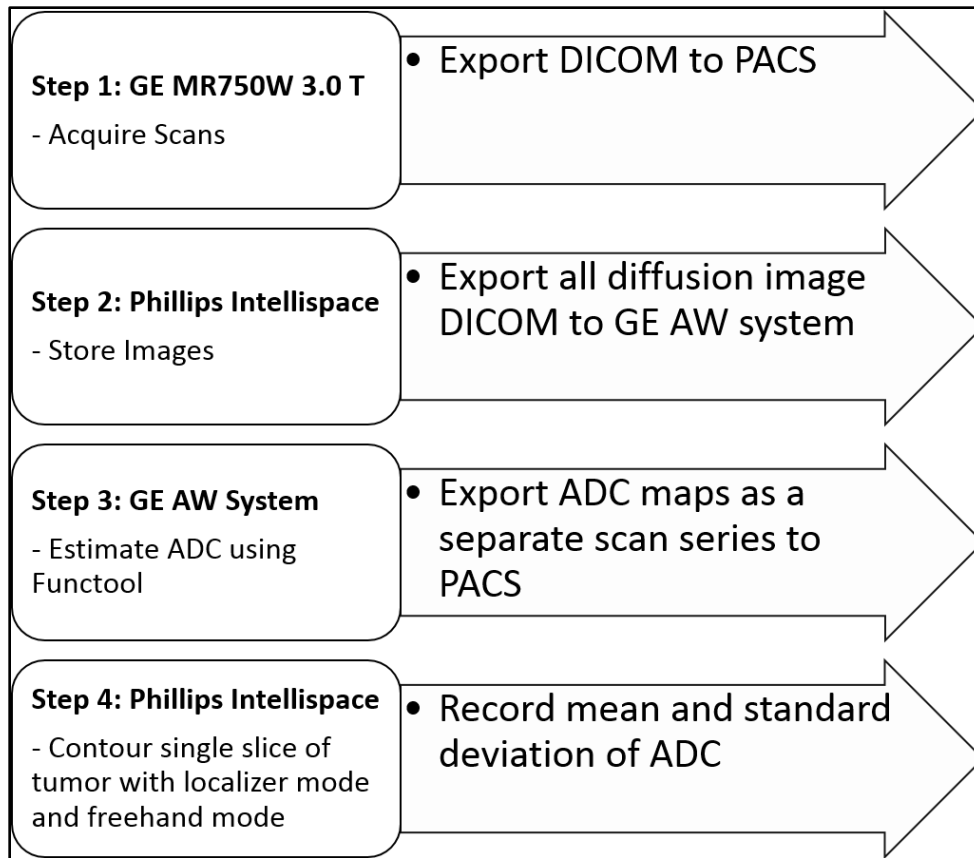


Figure 2-6: Workflow for measurement of diffusion parameters through slice contouring.

2.5 Inter-Reader Study

For the inter-reader study, four additional readers contoured MONO scans using both volumetric and slice contouring techniques. These readers included an experienced thoracic radiation oncologist, an experienced thoracic radiologist, and two radiation oncology residents. Before contouring, readers were given training across several datasets. They were then assigned randomized orders in which to contour all surgical MONO exams. They completed contours of these exams as volumetric contours using previously described volume contour methods and with slice contouring using the clinically available method.

Throughout contouring, all readers were blinded to the contours of other readers as well as the histopathological groupings of the patients.

2.6 Statistical Analysis

Statistical analysis was performed in Matlab, an open-source statistics software named R (GNU), and Prism by GraphPad Software (San Diego, California). P-values less than 0.05 were considered statistically significant.

2.6.1 Correlation of Diffusion Parameter Summary Measures with Histopathology

Diffusion parameters summary measures were compared for heterogeneity between tumor histologic groupings. Two major comparisons were made. The first compared diffusion parameter summary measures between TRG1, TRG2 and combined TRG3+ using one-way three-factor ANOVA or the Kruskal-Wallis (KW) test. A second method compared diffusion parameter summary measures between a binary grouping of patients as pCR (TRG1) and non-pCR (combined TRG2 and TRG3+) using a non-parametric t-test or Mann-Whitney (MW) u-test. The second comparison between binary groupings was extended to receiver operator curve (ROC) analysis. Optimal cutoff for classifying patients into the binary grouping were calculated using Youdens's index. To test for the possibility of building a multi-parametric model, Spearman's rank was used to assess correlation between diffusion parameter summary measures.

A third, less important comparison was also applied, where the binary definition was altered so that patients with <1% viable cell histopathology were compared with pCR

patients and a grouping of $\geq 1\%$ viable cells patients. This comparison is not relevant to a clinical endpoint. Instead, this comparison was performed to examine the sensitivity of quantitative diffusion parameters to differences between histopathologic groups which are seemingly close (0% viable cells vs $< 1\%$ viable cells). Results of this comparison can be found in the Appendix 5.5.

2.6.2 Inter-Reader Reproducibility Analysis

Inter-reader reproducibility was evaluated through three methods. First, quantitative measures were compared for reproducibility between readers using Bland-Altman analysis. Bland-Altman analysis results in a metric known as the 95% Limits of Agreement (LOA), which shows the likely distance between measurements of a system by separate readers. Second, contouring measures of individual readers were examined for performance under ROC analysis using the binary patient grouping. Both of these methods were applied to results obtained through both volume and slice contouring.

Finally, comparison of spatially contoured volumes was compared between readers using the Dice Similarity Coefficient (DSC). This was performed using Equation 22.

$$\text{Equation 22: } DSC = \frac{\# \text{ of Shared Voxels}}{(\# \text{ Voxels in Contour A}) + (\# \text{ Voxels in Contour B})}$$

Averages of DSC's across patients and between all pairings of readers were found for BL, IM and FU separately. Average DSC for each individual reader across all pairings to other readers, all patients, and all time-points were also calculated.

3 Results

3.1 Patient Enrollment

In total, 59 patients have been enrolled in the study. Several patient enrolled in the study did not undergo surgery because their disease had already metastasized, they had poor cardiopulmonary health, or they refused surgery (n = 39). These patients are not included in this analysis because currently elapsed follow-up time is inadequate for outcome analysis based on survival measures. The remaining 20 surgical patients are included within this analysis. According to the histopathological analyses of these twenty surgical patients, five exhibit pCR to nCRT (TRG1/pCR: n = 5), nine exhibit partial response to nCRT (TRG2: n = 9) and six patients exhibit limited response to nCRT (TRG3: n = 6). A more complete presentation of patient characteristics can be found in Table 3-1.

Table 3-1. Patient and Pathology Characteristics

Characteristic	N (%)
<i>Gender</i>	
Male	20 (100.0)
Female	0 (0.0)
Age (years)	62.1+/-8.0*
<i>Histological Tumor Type</i>	
Adenocarcinoma	17 (85.0)
Squamous cell carcinoma	3 (15.0)
<i>Histological Tumor grade</i>	
Well Differentiated	0 (0.0)
Moderately Differentiated	9 (55.0)
Poorly Differentiated	11 (45.0)
<i>Clinical Stage</i>	
IIa	1 (5.0)
IIb	5 (25.0)
IIIa	12 (60.0)
IIIb	2 (10.0)
<i>Tumor Location</i>	
Proximal third	0 (0.0)
Middle third	1 (5.0)
Distal third	17 (85.0)
Gastro-esophageal junction	2 (10.0)
<i>Histopathologic T-stage</i>	
ypT0	5 (25.0)
ypT1a	2 (10.0)
ypT1b	2 (10.0)
ypT2	3 (15.0)
ypT3	8 (40.0)
<i>Histopathologic N-stage</i>	
ypN0	14 (71.4)
ypN1	4 (20.0)
ypN2	2 (10.0)
ypN3	0 (0.0)
<i>Histopathologic Tumor Regression Grade</i>	
TRG 1	5 (25.0)
TRG 2	9 (19.0)
TRG 3-5	6 (30.0)

*Patient age is listed as mean+/-std

It should be noted that Table 3-1 describes the patient characteristics of all patients who underwent surgery and had useful MONO exams at BL and IM. For IVIM, one patient

(TRG1) is excluded because of a failed baseline exam (n=19 for BL and IM). Four patients did not receive FU exams because of loss to follow-up (Remaining MONO at FU: n=16; TRG1 = 4, TRG2 = 6, TRG3 = 6). In addition to this, IVIM scans were mistakenly not acquired at FU for three additional patients (remaining IVIM at FU: n=13; TRG1 = 3, TRG2 = 5, TRG3 = 5).

3.2 Image Quality

With the exception of a baseline IVIM exam with failed acquisition (n=1), all other acquired diffusion exams were of sufficient quality for quantitative volume measurements. Examples of images by exam type can be found in Figure 3-1, Figure 3-2, and Figure 3-3 for MONO, IVIM, and DKI, respectively.

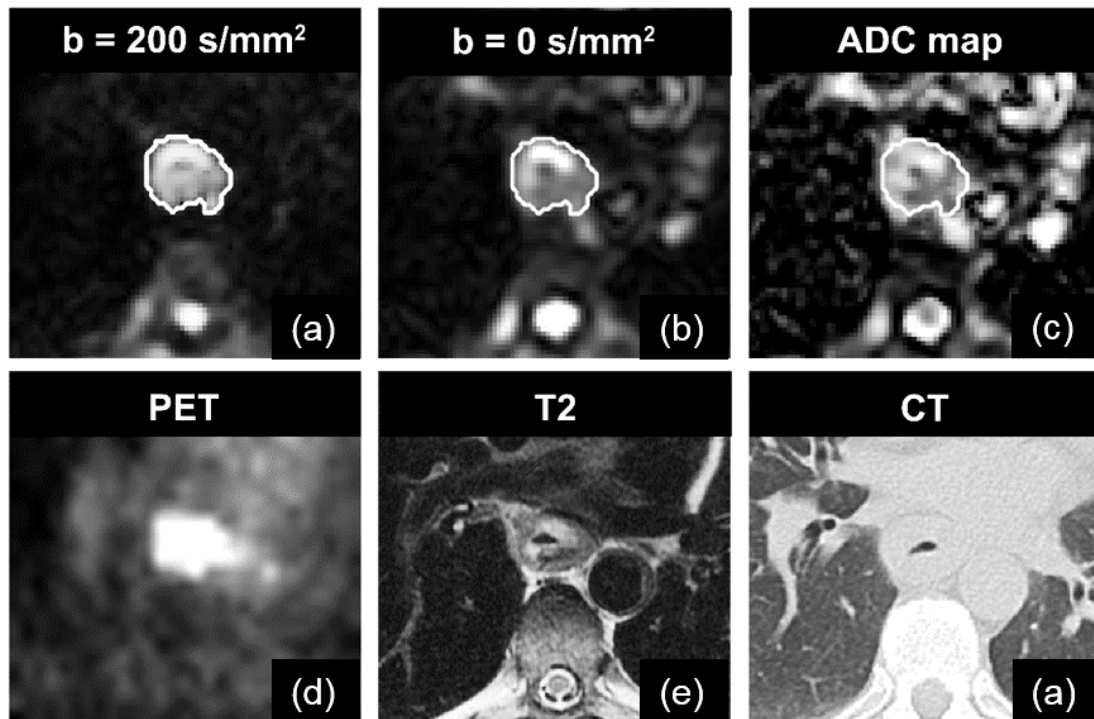


Figure 3-1: Diffusion images (a, b) and ADC map (c) from MONO exam of a patient at BL. Contours made on b100 are overlain on the image in white. FDG-PET (d), T₂-weighted image (e), and CT (f) images of the approximate slice are included for comparison.

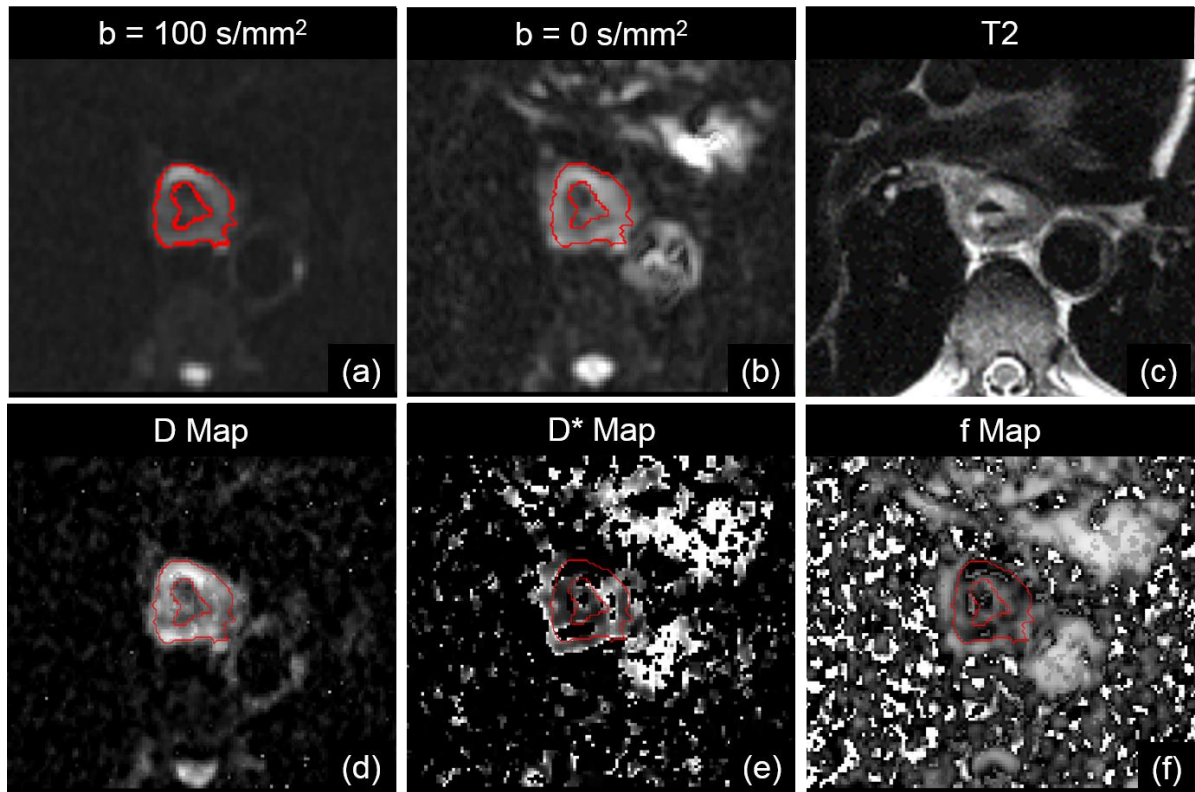


Figure 3-2: Diffusion images (a, b) and parameter maps (d-e) from IVIM exam of a patient at BL. Contours made on b100 are overlain on the image in red. A T₂-weighted image (c) of the approximate slice is included for comparison.

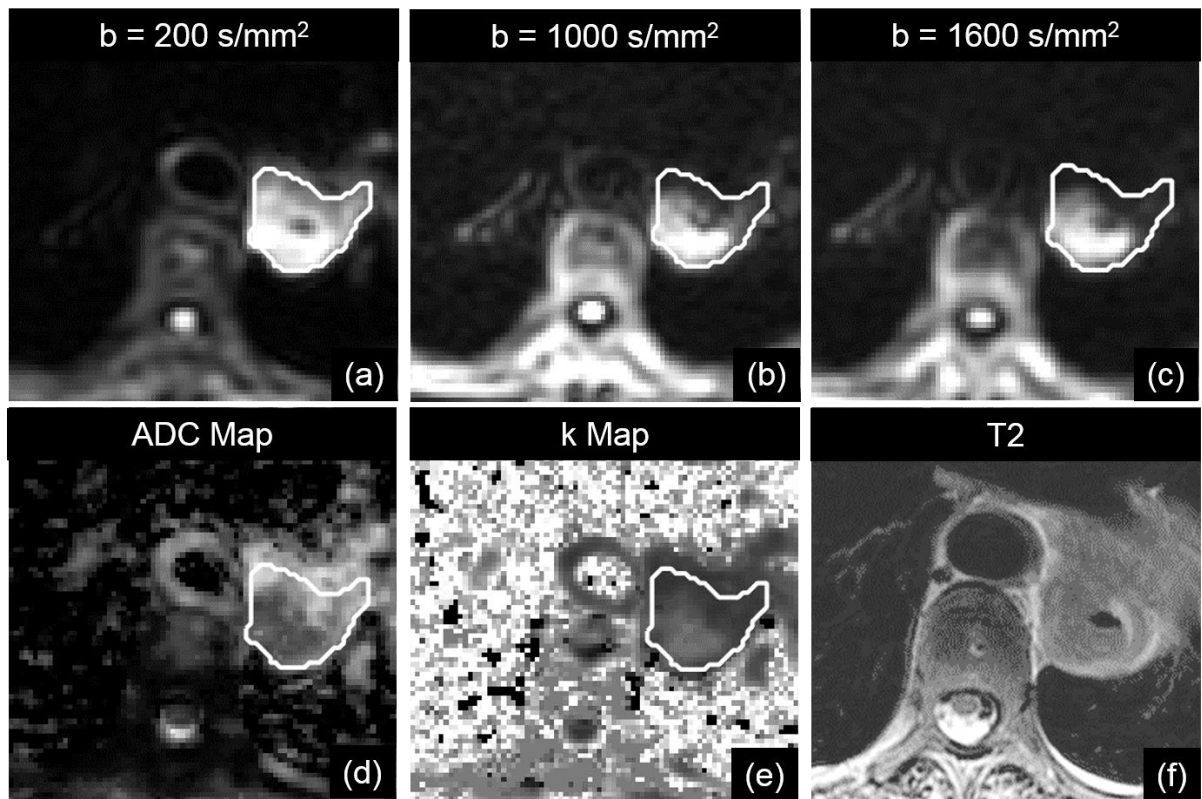


Figure 3-3: Diffusion images (a-c) and parameter maps (d,e) from DKI exam of a patient at BL. Contours made on b200 are overlain on the image in white. A T₂-weighted image (f) of the approximate slice is included for comparison.

Several artefacts were seen in the images: “flashing” artefacts (parallel imaging failure or large motion on a single slice), slight misalignment due to motion, and T₂ shine-through artefacts. Flashing artefacts were excluded through avoidance of single slices during contouring, and occurred infrequently (around 1 to 2 flashing artefacts for every 5 scans). T₂ shine-through resulted in subjectivity in contouring the inferior edge of tumor at the gastroesophageal junction. Images and further details of these artefacts can be found in the Appendix 5.4.

Two major errors in scan acquisition occurred while obtaining scans of surgical patients. The first was the failed baseline IVIM scan due to failed calibration. The second

major error was a calibration failure during the first five slices of a MONO scan. Because the tumor was not near the location of these slices, the scan was considered useable and included in this analysis.

3.3 Correlation of DWI Parameters with Histopathology

Unfortunately, DKI data cannot be reported upon presently because of insufficient patient accumulation (n = 25 patients, only 2 with surgery). The results of correlation of diffusion parameter summary measures of MONO (n=20) and IVIM (n=19) with histopathology are presented here for the volume contouring of the initial reader.

3.3.1 Mono-exponential Model (ADC)

Across ADC summary measures at all time-points and as relative changes between time-points, any differences between 3-factor TRG groupings detected with KW test were simply the extension of differences displayed with the binary grouping method (pCR vs non-pCR). No difference was found between the ADC summary measures in TRG2 and TRG3+. Because of this, KW test results are not reported here in favor of the binary comparison by MW test.

At baseline, four summary measures of ADC were found to be significantly different ($p < 0.05$) by MW test between pCR and non-pCR. ADC mean and ADC median tied for the most significant, with MW p-values of 0.0261 and AUC's of 0.840. Statistical analysis and values of these summary measures can be found in Table 3-2.

Table 3-2. Comparison of Baseline ADC Summary Measure Distributions between pCR and non-pCR using Mann-Whitney Test.

ADC Summary Measure	p-value	TRG	ADC in $\times 10^{-3} \text{ mm}^2/\text{s}$	AUC
ADC Mean	0.0261	TRG1	2.22 +/- 0.17	0.840
		TRG2	2.68 +/- 0.38	
		TRG3	2.48 +/- 0.28	
ADC Median	0.0261	TRG1	2.18 +/- 0.16	0.840
		TRG2	2.66 +/- 0.41	
		TRG3	2.45 +/- 0.30	
ADC 75 th Percentile	0.0403	TRG1	2.57 +/- 0.21	0.813
		TRG2	3.08 +/- 0.40	
		TRG3	2.86 +/- 0.33	
ADC 25 th Percentile	0.0493	TRG1	1.82 +/- 0.15	0.800
		TRG2	2.27 +/- 0.40	
		TRG3	2.10 +/- 0.28	

Baseline values of volume ADC mean were not significantly different between ESCC patients (n=2) and AEC patients (n = 18) using MW u-test (p>0.08). MW test did not demonstrate any difference in distribution of ADC summary measures between pCR and non-pCR patients at IM or FU.

Relative change in ADC (Δ ADC) summary measure at IM produced excellent separation between binary grouping of patients as pCR and non-pCR. Several ADC summary measures resulting in AUC=1 by ROC analysis. The top five Δ ADC summary measures by MW p-value are listed in Table 3-3.

Table 3-3. Comparison of Distributions of Δ ADC Summary Measures between pCR and non-pCR using Mann-Whitney Test.

Summary Measure	p-value	TRG	Relative % Change (Mean +/- STD)	AUC
ADC Mean	0.001063	TRG1	37.73 +/- 8.77	1.000
		TRG2	9.17 +/- 9.43	
		TRG3	9.77 +/- 4.01	
ADC 25 th Percentile	0.001063	TRG1	47.12 +/- 11.18	1.000
		TRG2	9.55 +/- 10.51	
		TRG3	11.60 +/- 4.40	
ADC 10 th Percentile	0.001063	TRG1	53.23 +/- 20.46	1.000
		TRG2	8.60 +/- 15.26	
		TRG3	12.51 +/- 7.31	
ADC Median	0.001443	TRG1	39.63 +/- 9.20	0.987
		TRG2	10.59 +/- 9.38	
		TRG3	10.74 +/- 3.65	
ADC 75 th Percentile	0.001944	TRG1	35.73 +/- 12.11	0.973
		TRG2	10.37 +/- 10.08	
		TRG3	8.07 +/- 4.98	

Youden's index was used to calculate an optimized cutoff for the top performing parameter, Δ ADC mean. This value was found to be a 27.7% change in ADC mean. A dot plot comparing Δ ADC mean between pCR and non-pCR can be found in Figure 3-4.

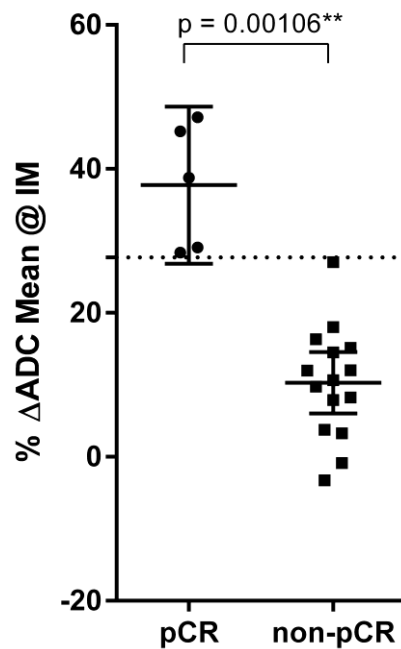


Figure 3-4: Comparison of distributions in Δ ADC mean between pCR and non-pCR groupings. P-value by MW test is shown at top. The 27.7% cutoff between pCR and non-pCR is shown as a dotted line. 95% confidence intervals and means are displayed with crosshairs for each group.

By Spearman's rank, the top Δ ADC summary measures at interim were found to be highly correlated. This result is depicted for the top 10 performing Δ ADC summary measures in Figure 3-5.

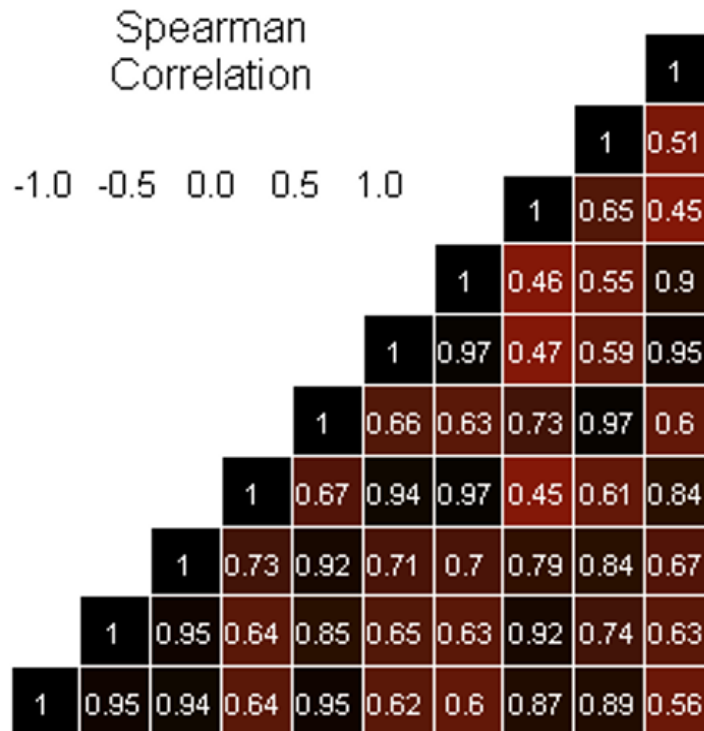


Figure 3-5: Spearman’s correlation matrix of the top 10 performing summary measures of Δ ADC.

No difference in distributions of Δ ADC summary measures at FU was found between pCR and non-pCR. Change in contoured tumor volume between BL and IM or FU were not significantly different between binary response groups.

3.3.2 IVIM Model

As with MONO, any differences between 3-factor TRG groupings detected with KW test were simply the extension of differences present in the binary grouping method (pCR vs non-pCR). Because of this, only differences based on the binary grouping are presented in this section.

Many baseline summary measures of IVIM parameters were highly different between pCR and non-pCR groupings. The IVIM parameter summary measure with the highest MW p-value at BL was D min, with an AUC = 0.950 by ROC. The top 5 performers by MW test p-value are shown in Table 3-4.

Table 3-4. Comparison of Baseline IVIM Parameter Summary Measures between pCR and non-pCR using MW test.

Summary Measure	p-value	TRG	D/D* ($\times 10^{-3}$ mm ² /s)	AUC
D Min	0.00693	TRG1	0.0613 +/- 0.1226	0.950
		TRG2	0.5045 +/- 0.2207	
		TRG3	0.3534 +/- 0.2422	
D* Mean	0.00932	TRG1	12.05 +/- 2.87	0.933
		TRG2	18.79 +/- 6.43	
		TRG3	18.44 +/- 2.80	
D Mean	0.0124	TRG1	1.04 +/- 0.53	0.917
		TRG2	2.08 +/- 0.59	
		TRG3	1.74 +/- 0.40	
D 25 th Percentile	0.0124	TRG1	0.80 +/- 0.49	0.917
		TRG2	1.70 +/- 0.53	
		TRG3	1.40 +/- 0.35	
D Median	0.0124	TRG1	1.02 +/- 0.53	0.916
		TRG2	2.03 +/- 0.58	
		TRG3	1.70 +/- 0.38	

When IVIM parameters summary measures were compared between pCR and non-pCR as relative changes at IM, 10 were found to be significantly different between groups. Of these, only two were not summary measures of the true-diffusion coefficient (D): pseudo-diffusion D* 90th percentile and perfusion fraction f 90th percentile. A list of these summary measures, their MW p-values, and their AUC's can be found in Table 3-5.

Table 3-5. Comparison of Distributions of Δ IVIM Summary Measures at IM between pCR and non-pCR using Mann-Whitney Test.

Summary Measure	p-value	TRG	Relative % Change (Mean +/- STD)	AUC
D	0.00270	TRG1	40.35 +/- 17.89	1.000
90th Percentile		TRG2	9.75 +/- 12.75	
		TRG3	4.31 +/- 5.90	
D	0.00270	TRG1	43.48 +/- 16.69	1.000
mean		TRG2	9.50 +/- 11.76	
		TRG3	6.99 +/- 10.18	
D	0.00270	TRG1	33.75 +/- 13.91	1.000
Max		TRG2	1.29 +/- 7.93	
		TRG3	0.51 +/- 9.45	
D	0.00270	TRG1	47.58 +/- 21.80	1.000
25 th Percentile		TRG2	8.77 +/- 10.88	
		TRG3	8.17 +/- 14.61	
D	0.00373	TRG1	42.70 +/- 17.78	0.983
75 th Percentile		TRG2	10.38 +/- 13.05	

		TRG3 6.43 +/- 8.55	
D	0.00373	TRG1 44.81 +/- 16.01	0.983
Median		TRG2 10.26 +/- 12.23	
		TRG3 8.44 +/- 11.23	
D*	0.00767	TRG1 662.26 +/- 650.29	0.933
Min		TRG2 62.98 +/- 212.26	
		TRG3 0.73 +/- 55.03	
D	0.01242	TRG1 39.81 +/- 25.64	0.917
Standard Deviation		TRG2 9.90 +/- 13.65	
		TRG3 0.24 +/- 7.63	
f	0.02781	TRG1 31.68 +/- 18.63	0.867
90 th Percentile		TRG2 10.63 +/- 11.15	
		TRG3 8.29 +/- 12.31	
D	0.02781	TRG1 53.57 +/- 43.60	0.867
10 th Percentile		TRG2 7.25 +/- 12.27	
		TRG3 9.28 +/- 15.24	

Youden's index was used to calculate an optimized cutoff for the top performing parameter, ΔD 90th percentile. This optimized cutoff was also found to be a 27.7% change in D 90th percentile. A dot plot comparing ΔD 90th percentile between pCR and non-pCR can be found in Figure 3-6.

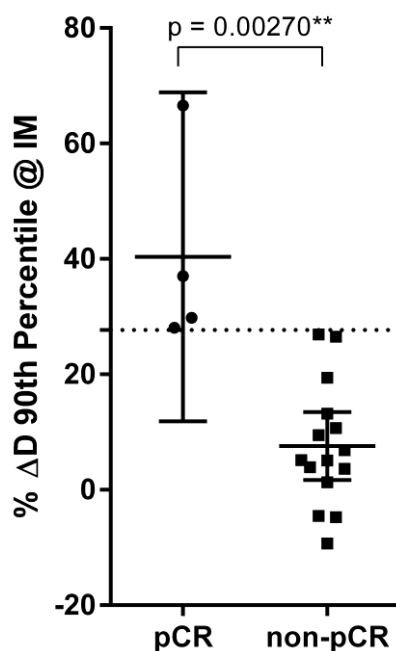


Figure 3-6: Comparison of distributions in ΔD 90th Percentile between pCR and non-pCR groupings. P-value by MW test is shown at top. The 27.7% cutoff between pCR and non-pCR is shown as a dotted line. 95% confidence intervals and means are displayed with crosshairs for each group.

By Spearman's rank, the top 10 $\Delta IVIM$ summary measures at interim were found to be highly correlated. This result is depicted in Figure 3-7.

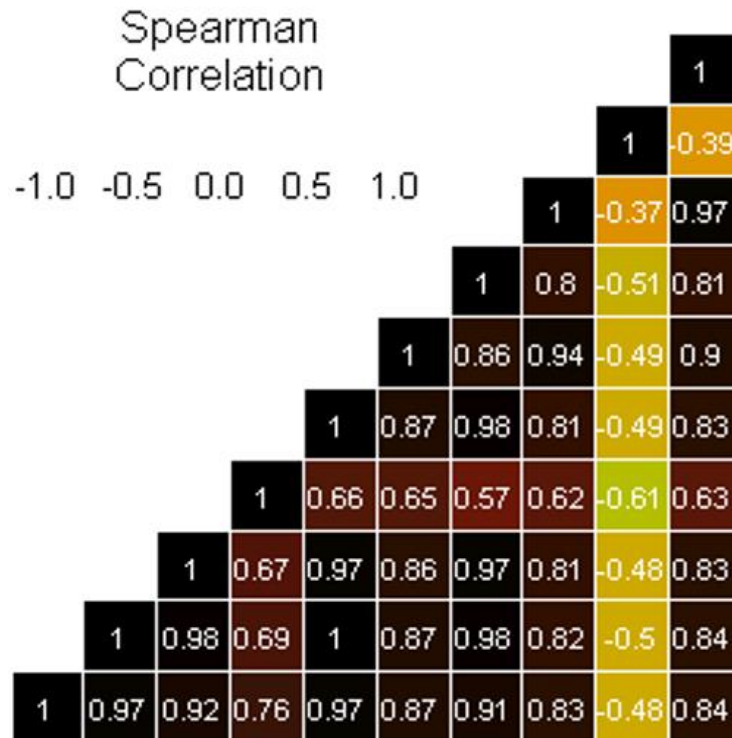


Figure 3-7: Spearman's correlation matrix of the top 10 performing summary measures of IVIM diffusion parameters as relative change at IM.

For IVIM, several significant differences in distributions of ΔD and ΔD^* summary measures as relative changes at FU were found between pCR and non-pCR. These results are not clinically useful because they do not predict response at an early time-point. All of these measures are listed in Table 3-6.

Table 3-6. Comparison of Distributions of Δ IVIM Summary Measures at FU between pCR and non-pCR using Mann-Whitney Test.

Summary Measure	p-value	TRG	Relative % Change (Mean +/- STD)
D* Standard Deviation	0.0102	TRG1	66.42 +/- 35.51
		TRG2	11.06 +/- 17.42
		TRG3	-61.92 +/- 52.36
D min	0.0158	TRG1	7.53E10 +/- 1.30E11
		TRG2	30.28 +/- 100.21
		TRG3	-57.00 +/- 58.90

D Max	0.0240	TRG1 52.42 +/- 46.42 TRG2 -9.15 +/- 18.89 TRG3 -56.60 +/- 60.41
D* 90 th Percentile	0.0240	TRG1 111.49 +/- 60.33 TRG2 17.13 +/- 34.78 TRG3 -63.18 +/- 50.82
D* Mean	0.0356	TRG1 110.36 +/- 106.42 TRG2 22.80 +/- 45.18 TRG3 -63.29 +/- 50.56
D* 75 th Percentile	0.0356	TRG1 163.58 +/- 193.84 TRG2 33.39 +/- 63.21 TRG3 -62.21 +/- 52.78

It is interesting to note that D summary measures were better classifiers of the binary grouping as relative changes at IM, while D* summary measures were better classifiers of the binary grouping as relative changes at FU. Change in contoured tumor volume between BL and IM or FU were not significantly different between binary response groups.

Because D from IVIM and ADC from MONO are both estimations of the diffusion coefficient, it was thought that a comparison of these estimations may be interesting. This comparison can be found in the Appendix 5.8.

3.4 Inter-Reader Study

Results of the inter-reader study as 95% LOA and AUC analysis can be found for top 5 performing ADC summary measures by volume contouring and all slice contouring ADC measures in Table 3-7. Bland-Altman plots are presented for the top performing summary measures of ADC as decided by maximum AUC_{min} (Δ ADC 25th Percentile and mean) in Figure 3-8.

Table 3-7: Inter-Reader Reproducibility for MONO Δ ADC Summary Measures by Bland-Altman and ROC Analyses.

Δ ADC	95% LOA	AUC				
Summary Measure	(per sd)	R1	R2	R3	R4	R5
Volume 10 th Percentile	0.940	1.000	1.000	0.987	1.000	0.960
Volume 25th Percentile	0.944	1.000	1.000	0.987	1.000	1.000
Volume Mean	1.001	1.000	1.000	0.973	0.973	1.000
Volume 75 th Percentile	1.010	0.973	1.000	0.947	0.880	0.933
Volume Median	1.021	0.987	1.000	0.933	0.960	1.000
Slice ADC Mean	1.517	0.893	0.800	0.880	0.773	0.880
Slice ADC STD DEV	1.738	0.507	0.533	0.613	0.600	0.600

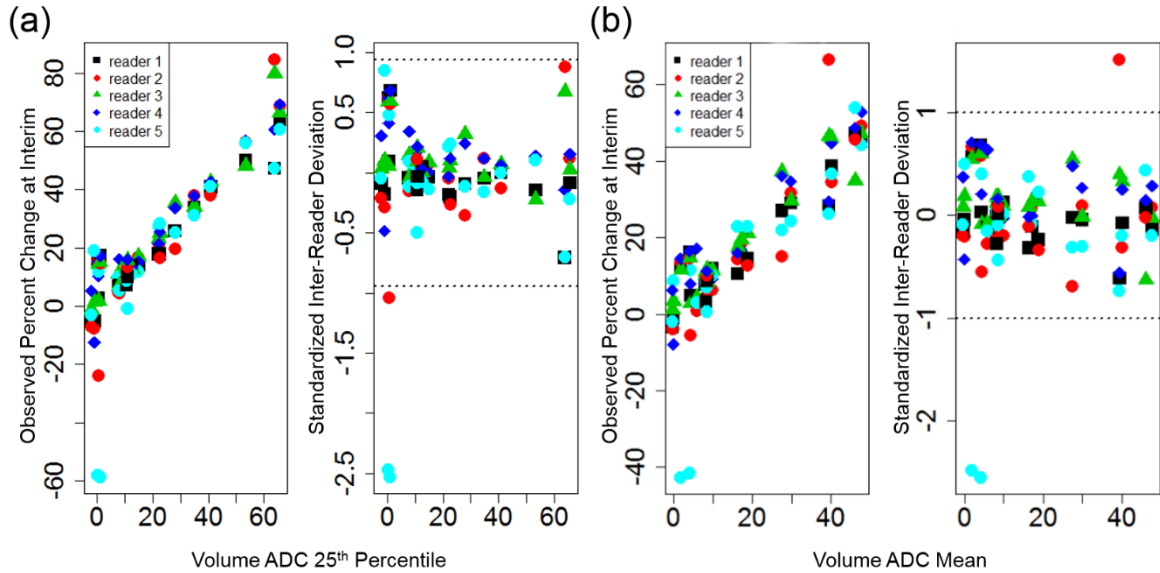


Figure 3-8: Bland-Altman plots comparing inter-reader variability for ADC 25th percentile (a) and ADC mean (b) from MONO scans.

Comparison of contoured volumes across all patients and all possible reader pairings produced average DSC's of 0.73, 0.66, and 0.56 for BL, IM, and FU, respectively. The maximum average DSC for all possible pairings of a single reader across all patients and time-points was 0.69. The minimum average DSC for all possible pairings of a single reader across all patients and time-points was 0.61.

4 Discussion

4.1 Response Prediction by DWI

The results of this study have shown diffusion parameters to be highly useful for classification of surgical patients as pCR or non-pCR. Although the measured diffusion parameters at baseline were significantly different between groups, the relative changes at

interim in these diffusion parameters provided substantially better classifications of the different patient TRG groups. Summary measures of ADC by MONO scans and D by IVIM scans were able to perfectly classify patients by the binary response grouping. The minimum value of D by IVIM scan at BL was an excellent classifier of the binary response grouping, but is based on the measurement in a single voxel, making it unstable.

Results for the MONO exam closely parallel those from a previous study by Van Rossum et al (29). Van Rossum's study was organized nearly identically to this one in terms of size, scan protocol, and patient histology. They found that an optimized cutoff of 29% relative change in Δ ADC median at interim best separated their pCR and non-pCR patients. This is strikingly similar to the 27.7% result found in this study. Agreement between these two papers was achieved in spite of geographically separated patient populations (North America vs Western Europe) and different magnet strengths/manufacturers (GE 3.0 T vs Phillips 1.5 T). This study serves as a good validation of their results.

Thus far, only a single paper has been published on the use of IVIM in esophageal cancer (67). This paper is not comparable to the data in this thesis because it does not concern response to therapy and was performed with a patient population dominated by ESCC.

4.2 Reproducibility of Contour Measurements between Readers

Inter-reader reproducibility by Bland-Altman and ROC analysis showed excellent reproducibility for measurements of several Δ ADC parameter summary measures. The top performing summary measure of Δ ADC was the 25th percentile because it had the highest minimum AUC between all readers (0.987). Slice contouring showed poor reproducibility with a high 95% limit of agreement (1.517) and low AUC across all readers. It is evident that slice contour is insufficient for measuring ADC for the prediction of response.

It is interesting that the lower percentile measures of ΔADC (10th and 25th percentile), show better LOA and AUC in comparison to other percentile measures (Table 3-7). This may have two plausible explanations. The first is that lower values of ADC correspond to higher signal on the diffusion images. Because high signal is the primary marker of tumor voxels in diffusion images, it is more likely that areas of low ADC will be contoured by different readers. A second explanation is the direct interpretation: change in diffusion behavior in the most hindered voxels of the tumor is more representative of response to nCRT. This has been shown previously for prediction of response to CRT in other anatomic sites (68-70). It is likely that the reason for this result is a combination of these two factors.

DSC analysis for spatial agreement of contours may be interpreted using the following ranges: values are considered by where they fall in the following ranges: 0.00–0.20 poor, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good and 0.81–1.00 excellent agreement (71). At BL and IM, average DSC fell into the “good” range, meaning that reproducibility of spatial volume contours was good between readers. Average DSC for individual readers across all patients and time-points was always above 0.61, meaning that all readers within this study were able to perform contouring with spatial location that was in good agreement with their peers. Results similar to this have been obtained by another group for contouring on diffusion images of rectal cancer (72, 73).

Because of the excellent performance of MONO scans with respect to classifying pCR and reproducibility of contour measurements, we have not examined the inter-reader reproducibility for IVIM scans.

4.3 Limitations

A major limitation of this work is the small dataset of surgical patients ($n = 20$). It is unlikely that the perfect separation of pCR and non-pCR patients will continue in a larger dataset. This is evident in the small difference between the 27.7% cutoff and the highest % $\Delta\text{ADC}_{\text{mean}}$ from the non-pCR group (27.04%), as well as the difference between the 27.7% cutoff and the lowest $\Delta\text{ADC}_{\text{mean}}$ from the pCR group (28.39%).

Another limitation is the use of histopathological response as a proxy for survival measures. Although histopathologic measures are highly predictive of survival, direct knowledge of survival is far more powerful for the purpose of retrospective classification.

A third limitation is the location of the tumors in this study. Most of the tumors were located in the distal third of the esophagus ($n = 17$, 85%). Only a few of the tumors were located in the GE junction ($n=2$, 10%). The patient population of this study is not representative of the general population, where incidence in the GE junction is more likely. Accurate contouring is more difficult at the GE junction due to the T₂ shine-through artefact of the stomach. Because of this, diffusion parameter measurement in the general population may be less reliable, which may impact the predictive value of DWI for response to nCRT.

4.4 Future Work

First, these results must be validated in a larger dataset. This will be accomplished through the continued accrual of surgical patients, combination of these results with those reported by Van Rossum et al., and through eventual KM analysis of non-surgical patients. Surgical patients should also undergo KM analysis, as histopathology measures do not hold the same value as survival measures.

After this biomarker has been validated in a larger dataset, it must be tested prospectively in a randomized Phase I clinical trial. In the experimental group of this trial, a patient whose intra-tumoral volume ADC increases sufficiently to predict pCR from chemoradiation alone will be withheld from surgery and placed under a wait-and-see approach. Other patients in this experimental group whose change in volume ADC predicts partial response will be sent for surgery (if able). In the control group, all patients will undergo surgery if able, regardless of change in volume ADC. By comparing survival of these two groups, it can be fully established if using DWI to predict response and avoid surgery in pCR patients holds a clinical benefit.

For routine clinical use, a consistent and fast method of volume contouring of diffusion images must be made available. The volume methods in this paper are fairly time consuming, requiring approximately 5 minutes of time for an experienced reader to complete a contour. In clinical workflow, this is much longer than a radiologist would typically spend reading an image dataset. Thus, it would be advantageous if a robust, fully-automatic method was clinically available for contouring. Recently, a paper has been published which uses convolutional neural networks to contour stroke lesions on DWI of the brain (74). This and other methods of automatic contouring may be investigated for their capabilities towards contouring of diffusion images with reduced reader intervention and oversight.

It is possible that the lack of success of FDG-PET as a predictor of response is caused by the non-specific uptake of FDG at IM such as from radiation induced inflammation. Non-specific uptake makes delineation of the tumor on FDG-PET/CT images difficult, and changes the values of summary measures such as standardized uptake value (SUV). It is possible that combined DWI with FDG-PET would provide better localization for tumor-

specific measurement of FDG-PET summary measures at interim. Co-registration of these images would be improved if they could be acquired in a conserved frame of reference within a PET/MR system. If FDG-PET measures were also found to be useful for prediction of response, both FDG-PET summary measures and ADC summary measures could be measured using a single set of contours acquired in an efficient, single machine setting. This would be time-saving, and may improve delineation of the tumor for both modalities in comparison to contemporaneous and separate acquisitions.

Once enough DKI patients have been accumulated, DKI will be investigated for potential value in prediction of treatment response in esophageal cancer. The kurtosis parameter has been shown as a useful biomarker of cancer in several anatomic sites (75-77). It may hold useful information distinct from ADC that can be used towards predicting the response of esophageal cancer to nCRT.

5 Appendix

5.1 Validation of ADC Map Calculation by Comparison to GE Functool

For validation of ADC calculation by ImageI, ADC maps were generated using GE Functool within a GE AW workstation and imported into ImageI. ADC extraction was performed using each map on volume contours of five esophageal scans. These scans including separate patients, variability in tumor size, and at least one scan from each time-point (BL, IM, and FU). Extracted volume ADC mean were compared between the two ADC mapping systems using an unpaired, two-sided Student's t-test. Results of this comparison are shown below in Table 5-1. Student's t-test did not reveal any significant difference in extracted volume ADC mean between the two methods (all $p > 0.46$).

Table 5-1. Comparison between extracted volume ADC mean from ADC maps generated by ImageI and GE Functool

Scan #	Volume (cm^3)	ImageI Mean ADC ($\times 10^{-3} mm^2/s$)	GE Mean ADC ($\times 10^{-3} mm^2/s$)	% Difference	Student's p-value
Scan 1	19.05	1.667 +/- 0.305	1.667 +/- 0.305	-0.01%	0.99
Scan 2	5.49	1.792 +/- 0.358	1.795 +/- 0.353	-0.17%	0.90
Scan 3	25.61	2.086 +/- 0.414	2.086 +/- 0.414	0.00%	1.00
Scan 4	10.29	2.074 +/- 0.506	2.079 +/- 0.613	-0.25%	0.83
Scan 5	43.90	2.038 +/- 0.422	2.031 +/- 0.399	0.31%	0.46

5.2 ITK-SNAP Settings and Guide

This section is intended to act as a demonstration of our contouring technique in ITK-SNAP. After diffusion images are imported into ITK-SNAP for contouring, the image must be prepared for semi-automated contouring through a series of pre-segmentation steps. In this example, the ITK-SNAP technique is applied to b200 images from a MONO scan. First, the

reader should adjust the contrast by selecting *Tools>Image Contrast>Contrast Adjustment*. This should be adjusted to the reader's preferences. Then, the reader must select the *Active Contour Segmentation* button, which is located on the upper left of the screen.

This will bring up a screen where the reader is asked to shrink the volume of the image set to a small box which barely encompasses the tumor. This is done by dragged dashed red lines to the edges of the tumor. The reader must be careful to include the entirety of all the tumors dimensions by searching the axial, sagittal, and coronal views of the display. An example of this volume minimization can be seen in Figure 5-1.

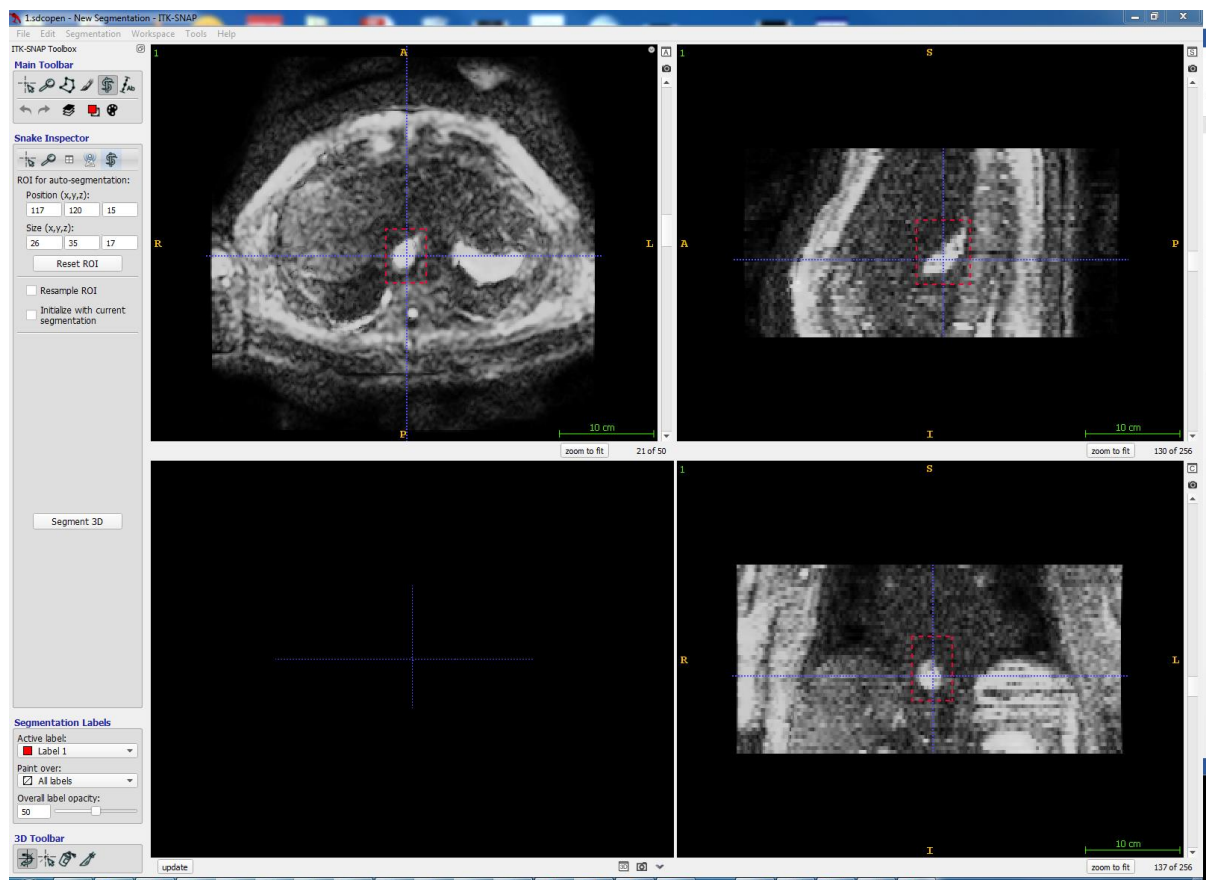


Figure 5-1: Minimization of volume for semi-automated contouring. The new smaller volume is outlined in red lines. The display includes axial (Upper Left), sagittal (Upper Right) and coronal (Bottom Right) views.

Once the contouring volume has been shrunk, the user must select the *Segment 3D* button in the left toolbar. The displayed images will change to include only the minimized volume. The user must click on one of the small blue and white images displayed next to each of the anatomic views. This will change all of the displays into blue and white “speed images”, where a prediction of the semi-automatic contouring is displayed. White pixels represent pixels which are predicted to be in the contour and blue pixels represent pixels which are predicted to be excluded from the contour. For the protocol used in this paper, the *Thresholding* method should be selected in the pull-down tab on the right toolbar. The user should then select the *More...* button on the right toolbar to bring up a window showing application of threshold limits to an image intensity histogram. The user should adjust the upper threshold to its maximum limit. The lower threshold should be adjusted so that the predicted speed images include the entirety of the tumor while minimizing predicted contouring of outside tissue. After this step is complete, advancement to the next step can be accomplished by clicking the *Next* button on the right toolbar. An example of this completed step can be seen in Figure 5-2.

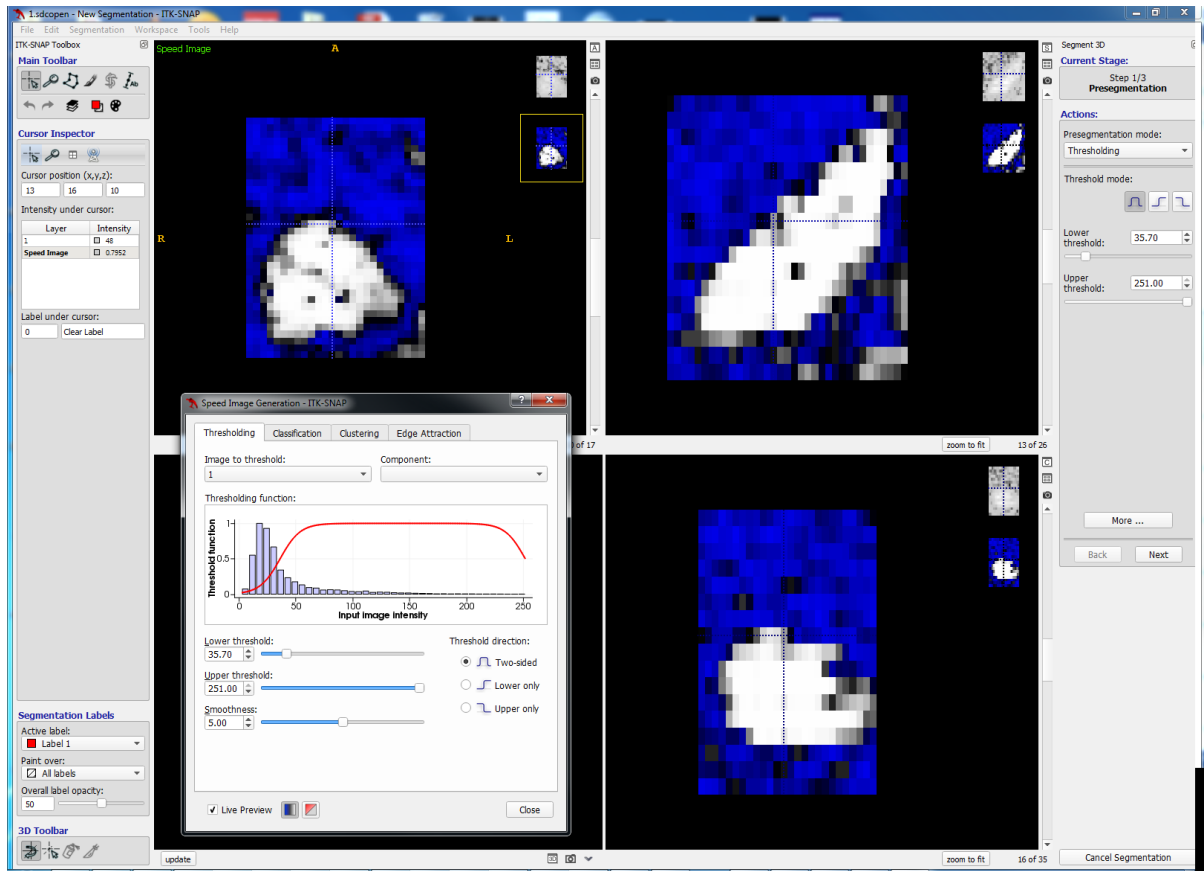


Figure 5-2: Final adjustments of thresholding for speed images of a minimized scan volume. The display includes axial (Upper Left), sagittal (Upper Right) and coronal (Bottom Right) views.

The last pre-contouring step is the placement of bubbles seeds, from which the contour will propagate. For efficiency, a minimum number of bubble seeds should be placed. The bubble size should be increased to maximize their dimension while still containing most of their volume within the tumor boundaries. Multiple bubbles should be placed if the sections of the tumor are disconnected, especially in the low resolution slice dimension. To place a bubble, the reticle should first be placed in the middle of the tumor with a mouse click. Then the *Add Bubble at Cursor* button on the right toolbar should be clicked. The radius of the placed bubble can then be adjusted using the slider on the right toolbar. Incorrectly placed bubbles should be

deleted by selecting them in the list on the right toolbar and clicking the *Delete Active Bubble* button. Once bubble seeds have been placed, the *Next* button on the right toolbar should be clicked for advancement of the contouring step. An example of good bubble placement can be seen in Figure 5-3.

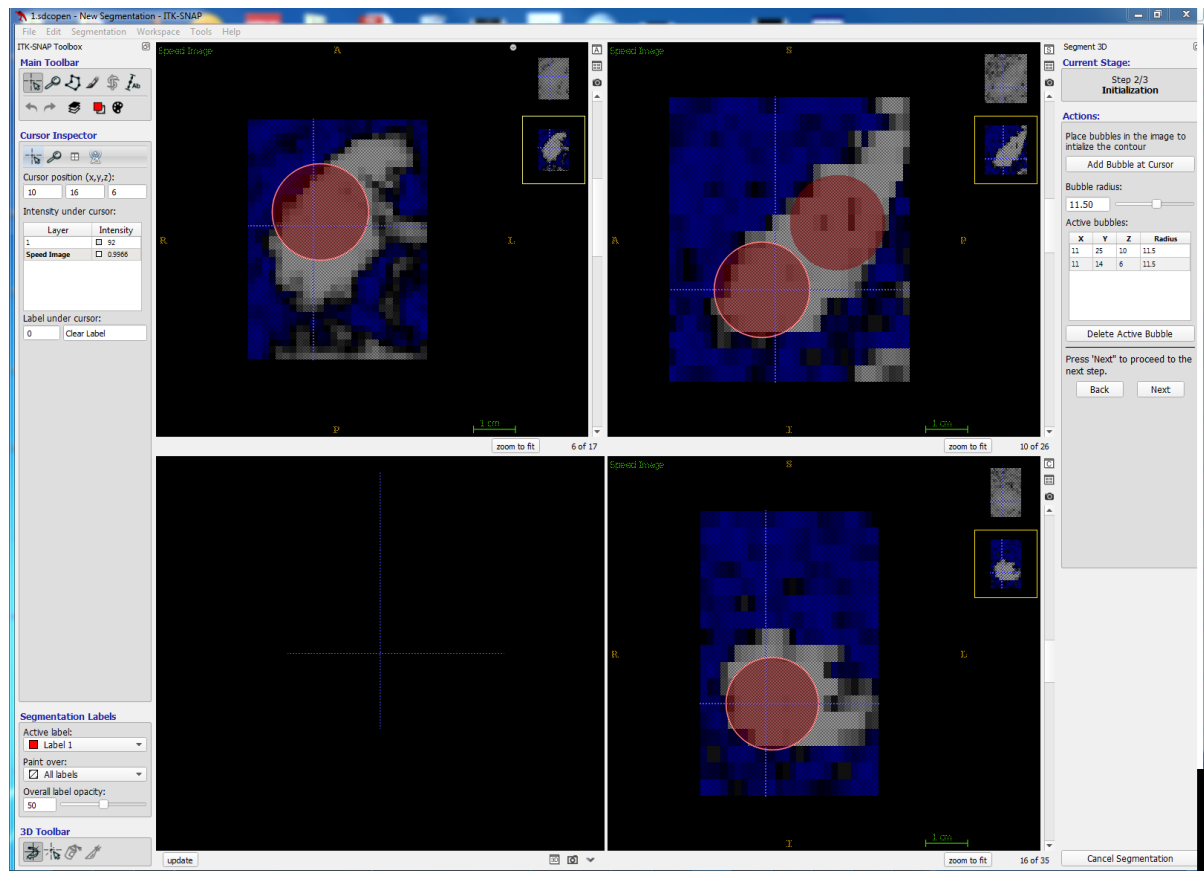


Figure 5-3: An example of good bubble seed placement within the tumor volume.

Next, the active contour must be activated. This can be done by clicking the “play button” on the right toolbar. Once the contour has propagated to the preferences of the reader, the “pause button” (which takes the place of the play button while active contours are being iteratively updated) should be clicked to stop the contour. If the contour is acceptable, it can

then be save and exported as binary masks in DICOM format. Screenshots of before and after active contouring can be seen in Figure 5-4 and Figure 5-5.

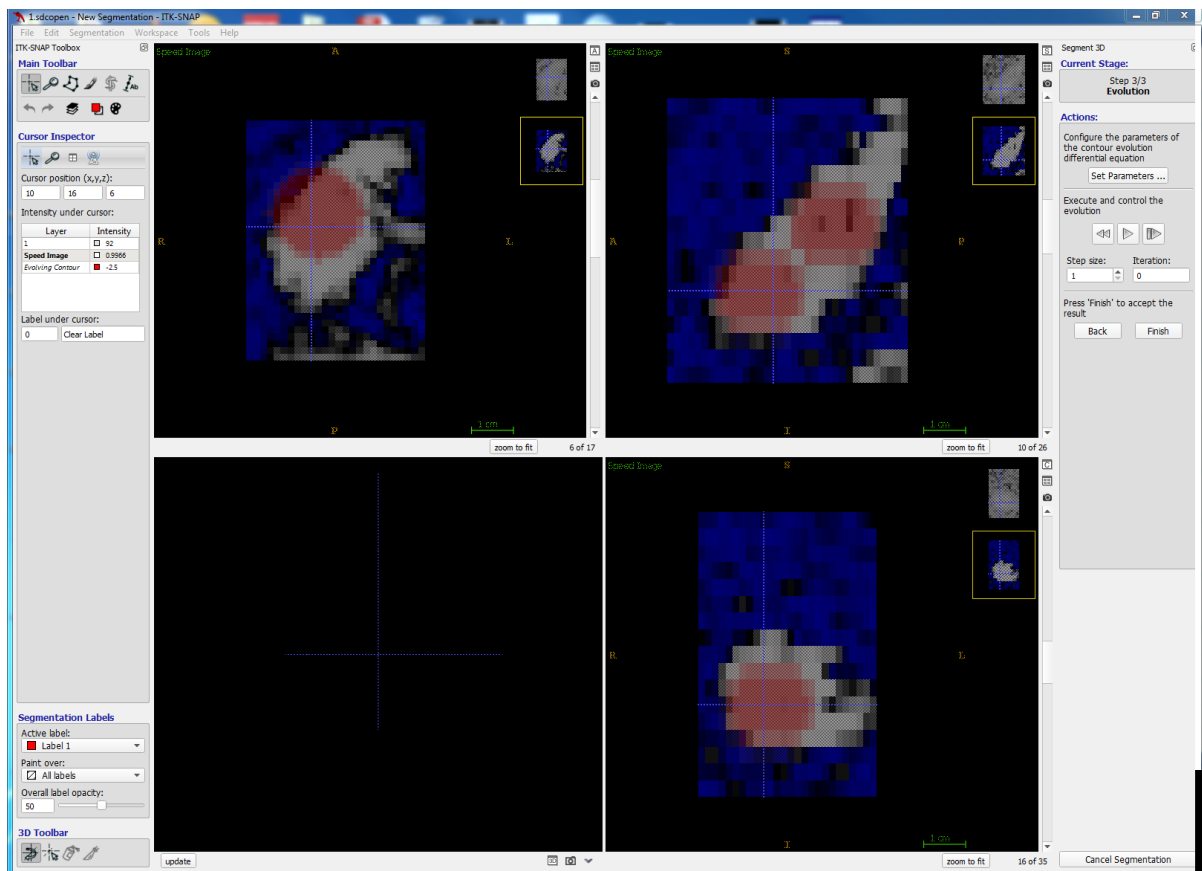


Figure 5-4: Screenshot of active contouring step before active contouring is activated.

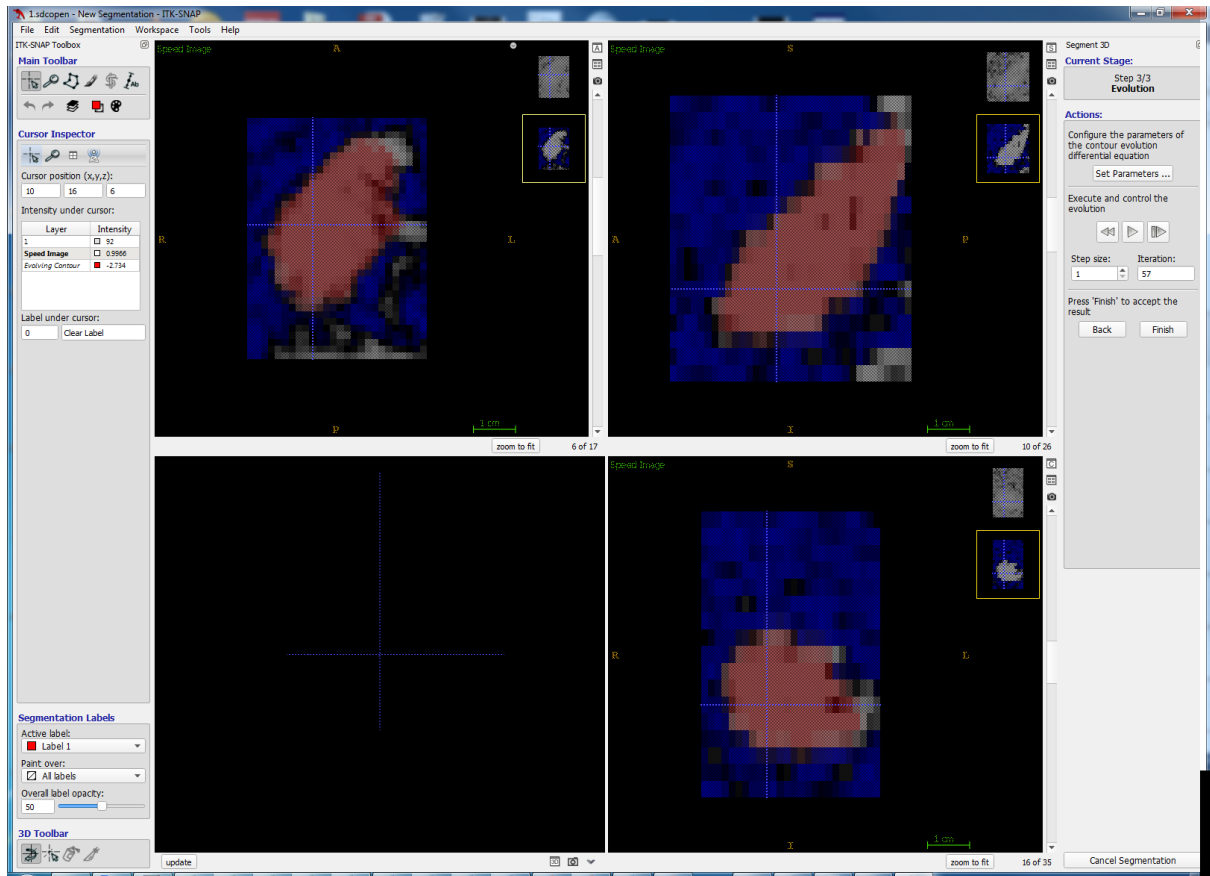


Figure 5-5: Screenshot of completed contour after iterative update of contour boundaries during active contouring has been stopped with appropriate timing.

If the contour fails to propagate throughout the entire tumor during the execution step, Active Contour Segmentation may be repeated with additional bubbles or a lower signal threshold. If the contour propagates too far, but still encompasses the entire tumor, the contour can instead be corrected later.

ITK-SNAP is an open-source software offering free downloads and tutorials. It is applicable to a wide range of biological imaging modalities. Excellent documentation, guides, and more can be found at <www.itksnap.org>.

5.3 Alternative Manual ImageI Contouring Method

This technique is fairly simple to implement, but requires significant manual input. Despite this, good manual contours can be achieved with this method in under 5 minutes per scan. This is comparable to the semi-automated method for BL scans. For IM and FU scans where diffusion is less restricted and signal is lower on diffusion images, this method may be faster than the semi-automated method. It can be accomplished in a series of three steps.

First, the reader must completely contour the tumor on all slices. This contour should be done roughly, the only goal being to completely encompass the tumor without including a significant amount of adjacent structures (cardium, aorta, etc). In ImageI, this is done with the paintbrush tool. An example of this rough outlining on a single slice can be found in Figure 5-6.

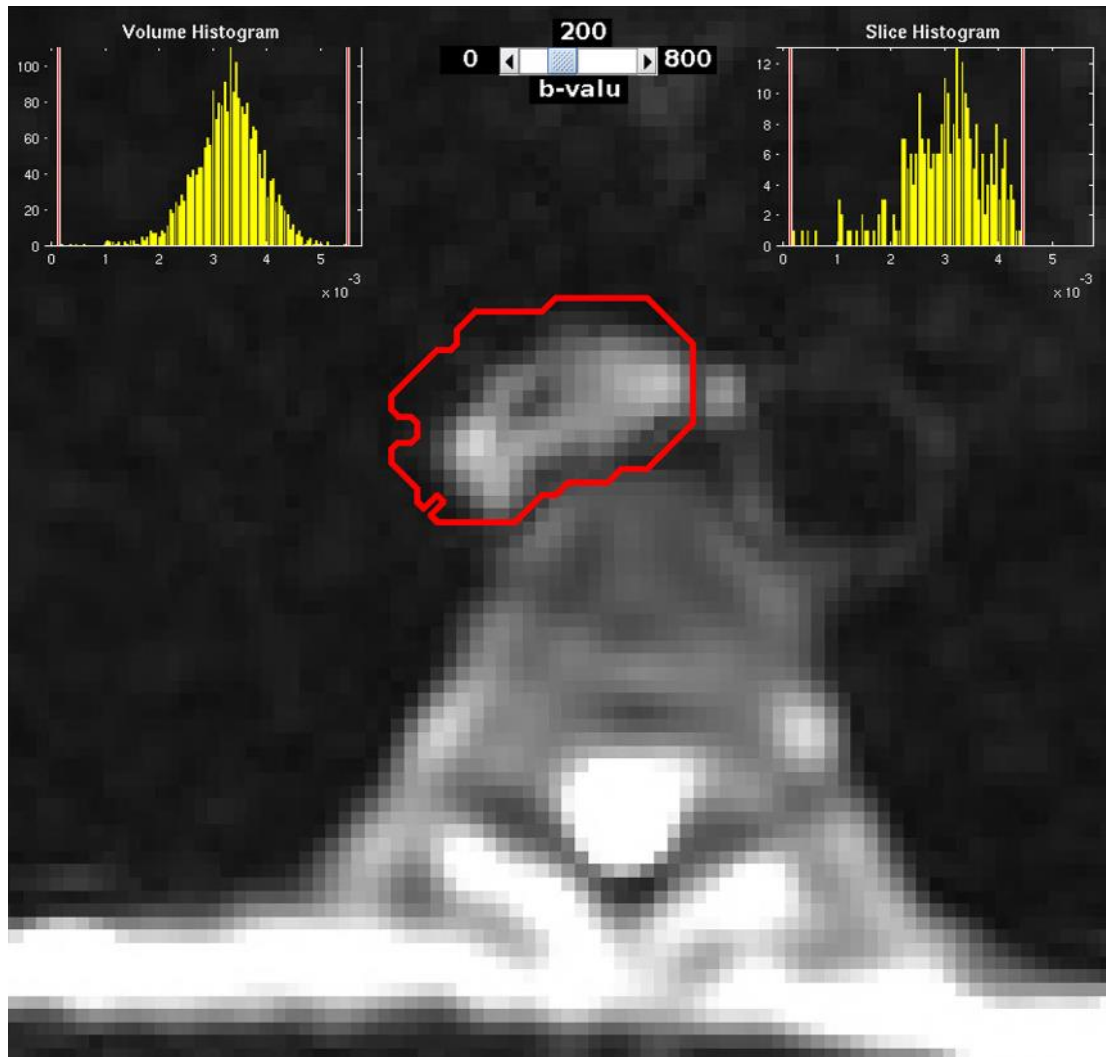


Figure 5-6: Rough outlining of tumor using several sweeps of the paintbrush tool. This should take on the order of a few seconds for each slice.

In the second step of this method, the rough contour will be drawn into the tumor edges by excluding lower signal values. This can be done across using the volume histogram slider in ImageI. Changes in the edges of the tumor should be tracked on a representative slice of the tumor. The readers should slowly raise the minimum included signal value bar until the edges of the tumor (including lumen) are contoured well. If the lumen is visible on any slices of the

scan, one of these slices should be chosen for observation during histogram signal exclusion. An example of this signal exclusion technique is shown for a single slice in Figure 5-7.

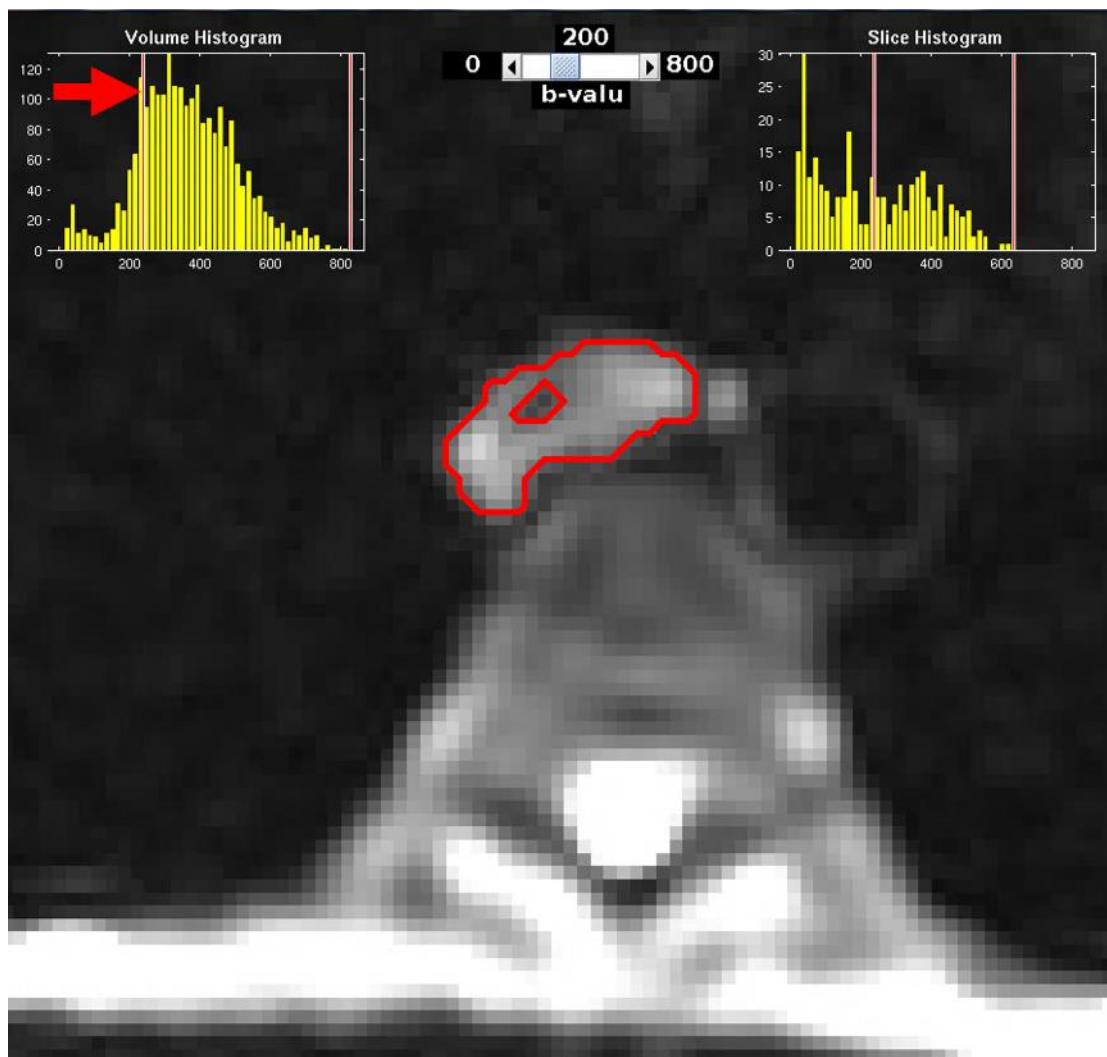


Figure 5-7: Example of edge contouring and lumen exclusion through raising minimum included signal level using the volume histogram tool in ImageJ. A red arrow shows the movement of the slider, which was manually dragged by the user.

After the second step is applied, the contour should be nearly perfect. To complete the contour, the reader must examine every slice and manually extend or contract edges of the contour. The reader should also fill in any missing holes within the tumor contour that do not

correspond to lumen. An example of this effort is shown in Figure 5-8, where the edges of the lumen were expanded and the edges of the tumor were drawn in on the posterior side.

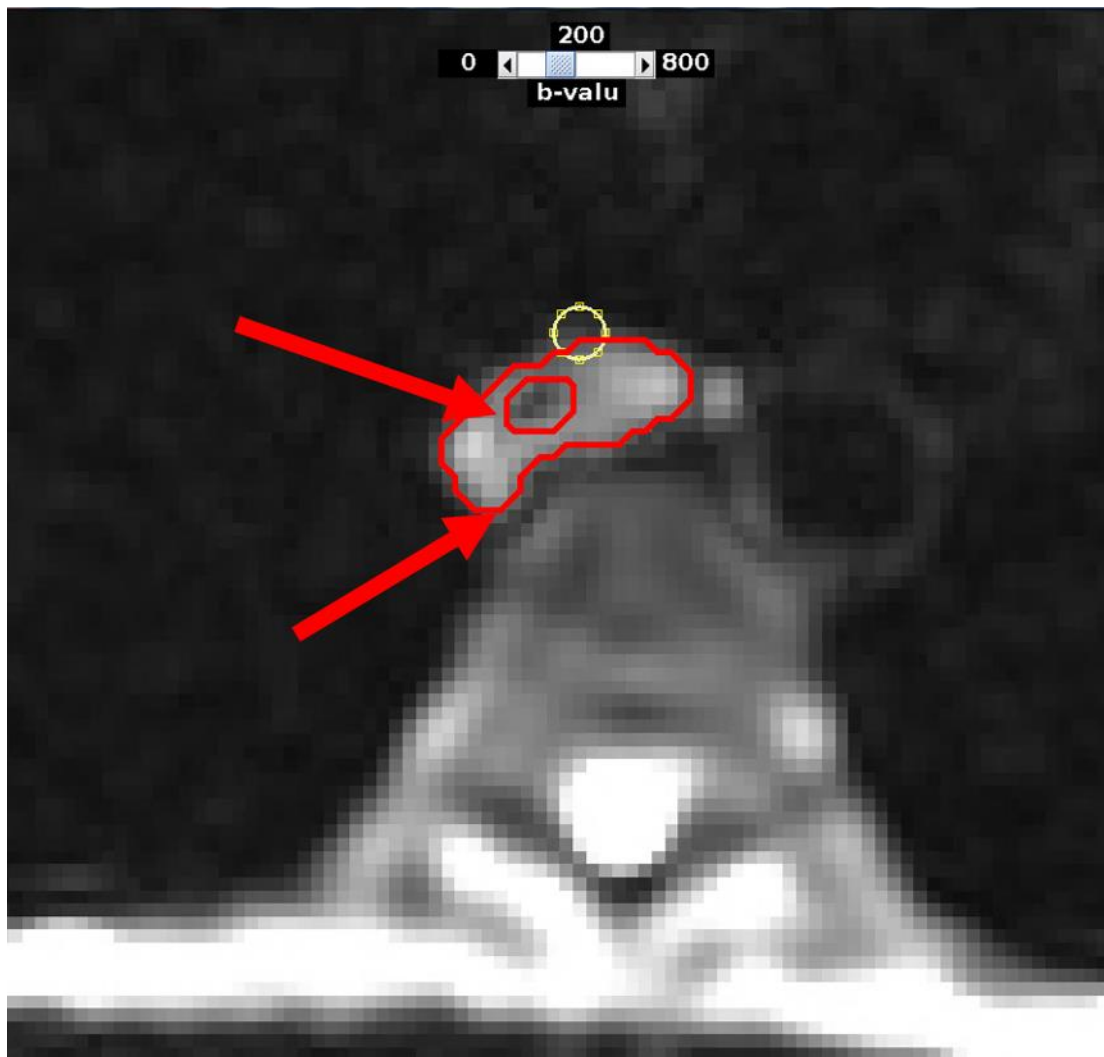


Figure 5-8: Final completion of the contour by manual adjustment of edges with eraser tool. The red arrows show points where the contour from Figure 5-7 was adjusted.

For further information about ImageI, please contact Dr. Jingfei Ma at the following email address: jma@mdanderson.org.

5.4 Diffusion Image Artefacts

5.4.1 Flashing Artefacts / Single Slice Signal Drop Out

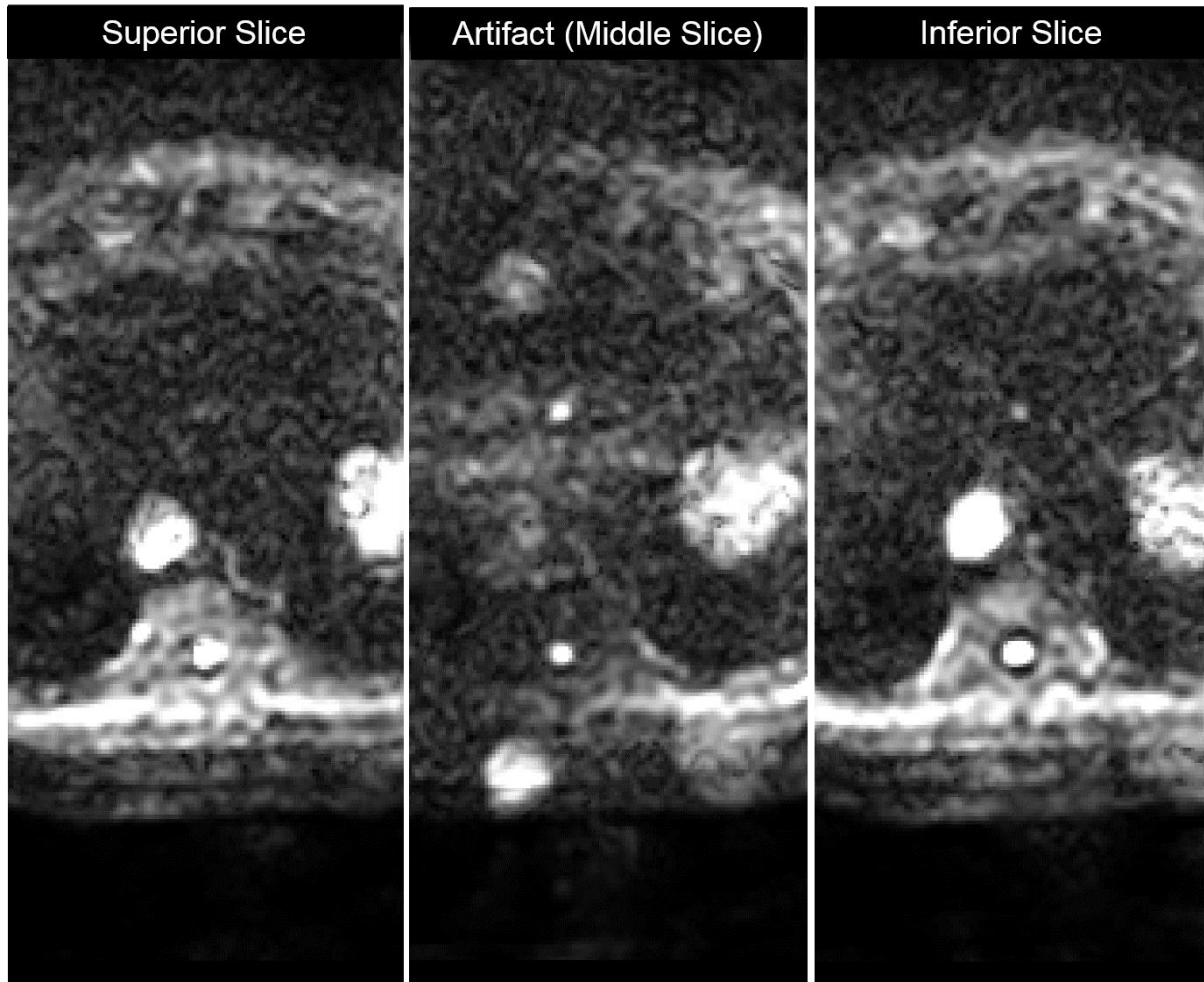


Figure 5-9: Massive shift of tumor signal on a single slice due to motion, causing a “flashing” effect when slices are scrolled through quickly. This is caused large scale motion or by parallel imaging correction failure, which leads to wrap-around artefacts in the phase direction (middle image).

5.4.2 T₂ Shine-through

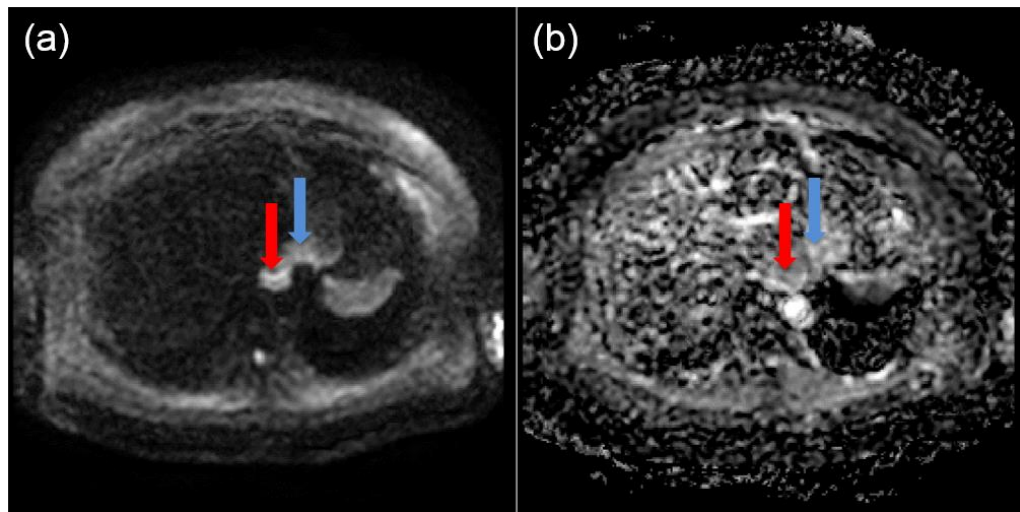


Figure 5-10: This is an example of T₂ shine-through in the b200 image (a) and ADC map (b) of a MONO scan. In this scan, the cancerous area of the esophagus is labeled with a red arrow and healthy stomach is labeled with a blue arrow. Although the stomach is not cancerous and is not experiencing abnormal restriction of diffusion, its T₂ is relatively long, causing it to be bright on the diffusion image. The ADC map accounts for T₂ shine-through, which is why the enhanced stomach on the diffusion image does not match with an area of low ADC (darker) on the ADC map.

5.5 Comparison of Diffusion Parameter Measurements between Alternate Groupings of <1% Viable Cell Patients

Comparison between Δ ADC mean at interim from MONO scans is shown in Figure 5-11. It is seen that patients with <1% viable cell pathology have distinct diffusion behavior from that of 0% viable cells.

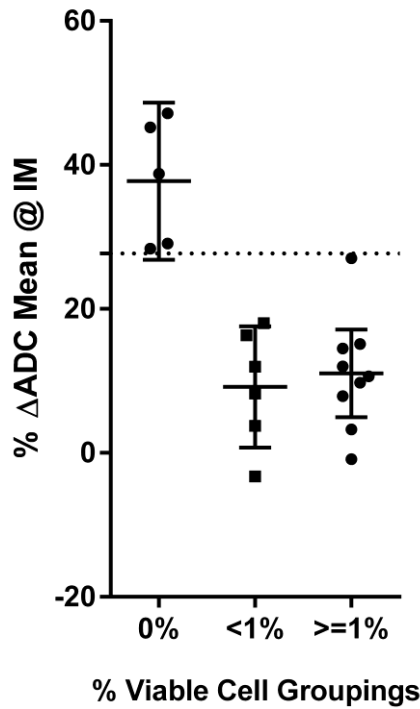


Figure 5-11: Comparison of Δ ADC mean at interim between an alternate set of viable cell groupings. The 27.7% cutoff between pCR and non-pCR is shown as a dotted line.

5.6 Padhani et al. ADC Method Comparison

In this section of the appendix, two methods of calculating mono-exponential ADC are compared to the non-linear LMS (b0, b200, and b800) used in this paper: direct solution by Padhani et al method (b200 and b800), and classic single b-value solution method (b0 and b800). These comparisons are made for mean volume ADC, both as individual measurements and relative change at IM. On average, mean volume ADC by Padhani et al. method was 42% lower than the estimate by LMS used in this thesis. This is an intuitive finding, as exclusion of b0 images would avoid perfusion effects at low b-value from propagating into the model. Adding b0 images forces the curve fitting to account for loss in signal due to perfusion, increasing ADC to account it. On average, mean volume ADC by the classic single b-value

method (b0, b800) was 5.2% lower than the estimation method of this paper. This is driven by the same mechanism, where inclusion of b200 images in this paper's estimation method forces ADC to account for fast drop in signal at b200 due to perfusion effects in lower b-value ranges. Plots depicting correlation of the alternate methods with the method of this paper by linear regression can be found in Figure 5-12.

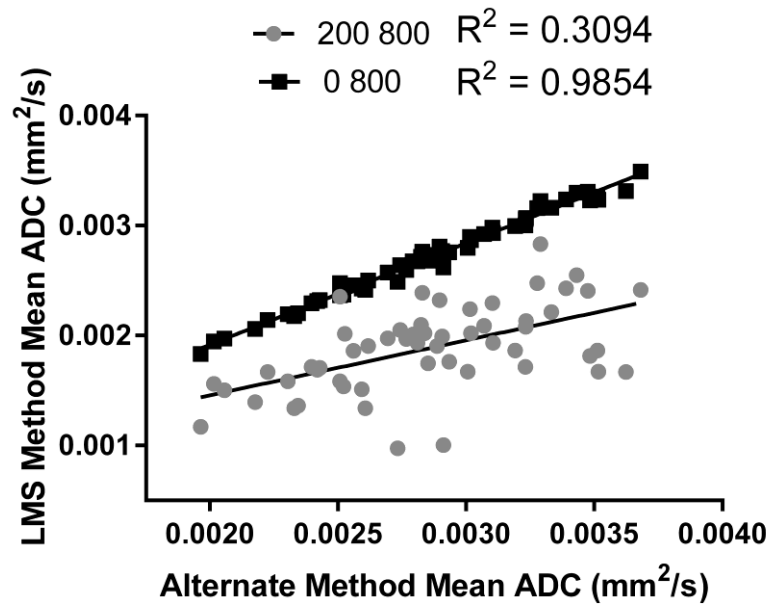


Figure 5-12: Correlation of LMS Curve Fitting Method Mean ADC across all MONO Scans with Mean ADC by Alternate Methods of ADC Estimation. R^2 values from linear regression are shown at the top.

Differences in estimated mean volume ADC between any two methods were not different for any scan time-point (BL vs IM vs FU: KW test p-value >0.3 for either comparison) and was also not different between pCR and non-pCR groupings (MW p-value >0.12 for both comparisons). This implies that the differences were systematic. For separation of pCR and non-pCR groupings, AUC=1 was maintained for only the curve fitting and 0,800 methods. The method by Padhani et al. decreased AUC to 0.893. It is possible that a larger patient dataset

may favor one method of ADC estimation over the others for the purpose of predicting response in esophageal cancer. For this study, there is no sufficient evidence indicating favorability of one estimation technique over another.

5.7 Comparison of Largest Slice and Single Slice Contour Measurements

If isolating the largest slice from a volume contour was a good approximation of measurements by single slice contouring, mean ADC extracted from largest slice should correlate well with mean ADC from single slice contouring. This was compared individually for each of the readers in the inter-reader study. This correlation can be seen in Figure 5-13. Measurements of ADC mean by single slice contouring did not correlate well with measurement of ADC mean by largest slice according to linear regression. This analysis is highly simplified for purpose of comparing measurement techniques (Bland-Altman analysis would be the proper method), but is enough to show that these techniques do not obtain comparable measurements in this application (across all readers: max $R^2 = 0.6162$, min $R^2 = 0.2887$)

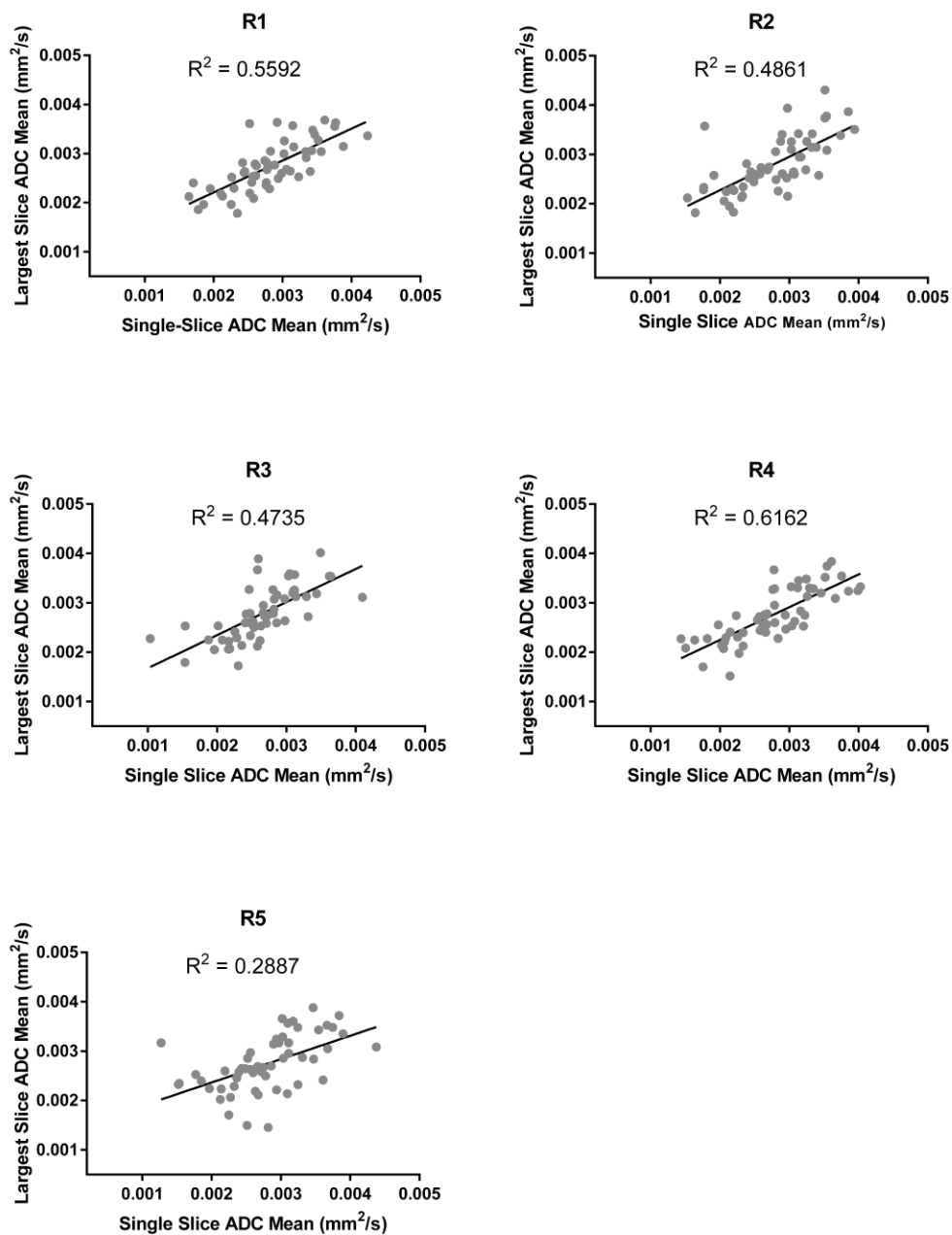


Figure 5-13: Correlation of ADC mean measurements by Largest Slice (y-axis) and Single Slice (x-axis) methods. Graphs are titled with their respective readers, and R² values are shown at the top of the graphs.

5.8 Correlation of IVIM D and MONO ADC

Because both ADC from MONO scans and D from IVIM scans are parameters to estimate the true diffusion coefficient, these parameters were compared. D from IVIM was compared with LMS MONO ADC and Padhani et al. MONO ADC. Volume mean D by IVIM was, on average, 41.5% lower than volume mean ADC by MONO with LMS fitting and 21.1% lower than volume mean ADC by MONO with the Padhani et al. method. A correlation of diffusion coefficient estimations is shown below in Figure 5-14.

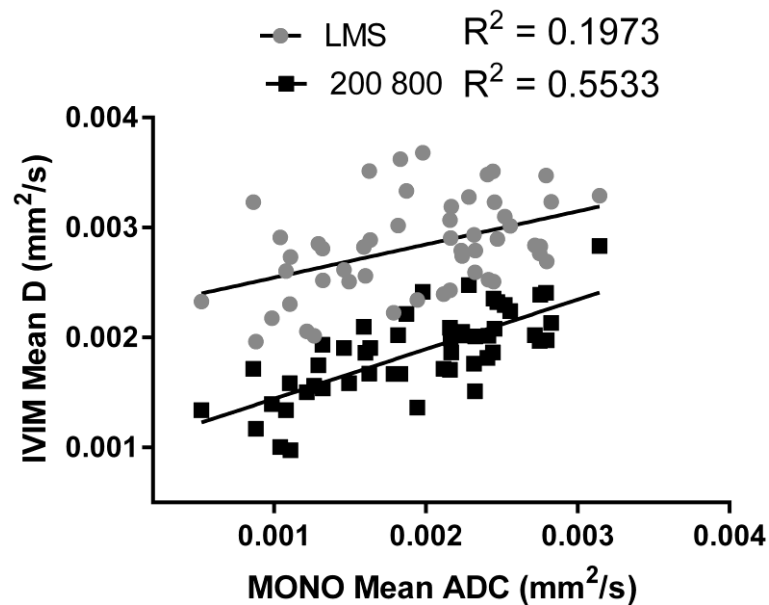


Figure 5-14: Correlation of Volume Mean D by IVIM with Volume ADC Mean by MONO through two different estimation methods. “LMS” corresponds to the MONO ADC method used in this paper, while “200 800” corresponds with the method put forth by Padhani et al. R^2 values from linear regression are shown at the top.

MONO ADC by the Padhani et al. method correlates better with D by IVIM. Neither of the ADC by MONO estimation methods were highly correlated with D by IVIM.

6 References

1. Lin, S. H., and Z. Liao. 2011. Decision making in radiation oncology. In *Medical radiology Radiation oncology*,. J. J. Lu, and L. W. Brady, eds. Springer, Heidelberg. 329-358.
2. Urba, S. G., M. B. Orringer, A. Turrisi, M. Iannettoni, A. Forastiere, and M. Strawderman. 2001. Randomized trial of preoperative chemoradiation versus surgery alone in patients with locoregional esophageal carcinoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 19: 305-313.
3. Berger, A. C., J. Farma, W. J. Scott, G. Freedman, L. Weiner, J. D. Cheng, H. Wang, and M. Goldberg. 2005. Complete Response to Neoadjuvant Chemoradiotherapy in Esophageal Carcinoma Is Associated With Significantly Improved Survival. *Journal of Clinical Oncology* 23: 4330-4337.
4. Donahue, J. M., F. C. Nichols, Z. Li, D. A. Schomas, M. S. Allen, S. D. Cassivi, A. Jatoi, R. C. Miller, D. A. Wigle, K. R. Shen, and C. Deschamps. 2009. Complete pathologic response after neoadjuvant chemoradiotherapy for esophageal cancer is associated with enhanced survival. *The Annals of thoracic surgery* 87: 392-398; discussion 398-399.
5. Atkins, B. Z., A. S. Shah, K. A. Hutcheson, J. H. Mangum, T. N. Pappas, D. H. Harpole, Jr., and T. A. D'Amico. 2004. Reducing hospital morbidity and mortality following esophagectomy. *The Annals of thoracic surgery* 78: 1170-1176; discussion 1170-1176.
6. van Rossum, P. S., L. Goense, J. Meziani, J. B. Reitsma, P. D. Siersema, F. P. Vleggaar, M. van Vulpen, G. J. Meijer, J. P. Ruurda, and R. van Hillegersberg. 2016. Endoscopic biopsy and EUS for the detection of pathologic complete response after neoadjuvant

- chemoradiotherapy in esophageal cancer: a systematic review and meta-analysis. *Gastrointestinal endoscopy* 83: 866-879.
7. Heneghan, H. M., C. Donohoe, J. Elliot, Z. Ahmed, V. Malik, N. Ravi, and J. V. Reynolds. 2016. Can CT-PET and Endoscopic Assessment Post-Neoadjuvant Chemoradiotherapy Predict Residual Disease in Esophageal Cancer? *Annals of surgery* 264: 831-838.
 8. Lee, J. M., S. Y. Yang, P. W. Yang, C. T. Shun, M. T. Wu, C. H. Hsu, C. C. Lin, J. C. Cheng, Y. H. Wang, T. H. Chuang, J. S. Chen, H. H. Hsu, P. M. Huang, S. W. Kuo, and Y. C. Lee. 2011. Polymorphism in epidermal growth factor receptor intron 1 predicts prognosis of patients with esophageal cancer after chemoradiation and surgery. *Annals of surgical oncology* 18: 2066-2073.
 9. Tomimaru, Y., M. Yano, K. Takachi, K. Kishi, I. Miyashiro, M. Ohue, H. Ohigashi, Y. Sasaki, O. Ishikawa, and S. Imaoka. 2008. Correlation between pretherapeutic d-dimer levels and response to neoadjuvant chemotherapy in patients with advanced esophageal cancer. *Diseases of the esophagus : official journal of the International Society for Diseases of the Esophagus* 21: 281-287.
 10. Makino, T., M. Yamasaki, H. Miyata, S. Yoshioka, S. Takiguchi, Y. Fujiwara, K. Nakajima, T. Nishida, M. Mori, and Y. Doki. 2010. p53 Mutation status predicts pathological response to chemoradiotherapy in locally advanced esophageal cancer. *Annals of surgical oncology* 17: 804-811.
 11. Pennathur, A., M. K. Gibson, B. A. Jobe, and J. D. Luketich. 2013. Oesophageal carcinoma. *Lancet (London, England)* 381: 400-412.

12. Zhang, Y. 2013. Epidemiology of esophageal cancer. *World Journal of Gastroenterology : WJG* 19: 5598-5606.
13. Edge, S. B., and American Joint Committee on Cancer. 2010. *AJCC cancer staging manual*. Springer, New York.
14. Ell, C., A. May, O. Pech, L. Gossner, E. Guenter, A. Behrens, L. Nachbar, J. Huijsmans, M. Vieth, and M. Stolte. 2007. Curative endoscopic resection of early esophageal adenocarcinomas (Barrett's cancer). *Gastrointestinal endoscopy* 65: 3-10.
15. Wang, D. B., Z. Y. Sun, L. M. Deng, D. Q. Zhu, H. G. Xia, and P. Z. Zhu. 2016. Neoadjuvant Chemoradiotherapy Improving Survival Outcomes for Esophageal Carcinoma: An Updated Meta-analysis. *Chinese Medical Journal* 129: 2974-2982.
16. Stahl, M., M. Stuschke, N. Lehmann, H. J. Meyer, M. K. Walz, S. Seeber, B. Klump, W. Budach, R. Teichmann, M. Schmitt, G. Schmitt, C. Franke, and H. Wilke. 2005. Chemoradiation with and without surgery in patients with locally advanced squamous cell carcinoma of the esophagus. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 23: 2310-2317.
17. Sgourakis, G., I. Gockel, A. Radtke, T. J. Musholt, S. Timm, A. Rink, A. Tsiamis, C. Karaliotas, and H. Lang. 2010. Minimally invasive versus open esophagectomy: meta-analysis of outcomes. *Digestive diseases and sciences* 55: 3031-3040.
18. Lazopoulos, A., N. Barbetakis, G. Lazaridis, S. Baka, I. Mpoukovinas, V. Karavasilis, I. Kioumis, G. Pitsiou, A. Papaiwannou, N. Katsikogiannis, A. Mpakas, K. Tsakiridis, S. Lampaki, A. Karavergou, M. Kipourou, M. Lada, K. Zarogoulidis, and P. Zarogoulidis. 2015. Open thoracotomy for pneumothorax. *Journal of Thoracic Disease* 7: S50-S55.

19. Yibulayin, W., S. Abulizi, H. Lv, and W. Sun. 2016. Minimally invasive oesophagectomy versus open esophagectomy for resectable esophageal cancer: a meta-analysis. *World journal of surgical oncology* 14: 304.
20. Kelsen, D. P., R. Ginsberg, T. F. Pajak, D. G. Sheahan, L. Gunderson, J. Mortimer, N. Estes, D. G. Haller, J. Ajani, W. Kocha, B. D. Minsky, and J. A. Roth. 1998. Chemotherapy followed by surgery compared with surgery alone for localized esophageal cancer. *The New England journal of medicine* 339: 1979-1984.
21. Xi, M., Z. Liao, W. L. Hofstetter, R. Komaki, L. Ho, and S. H. Lin. 2017. FDG-PET Response After Induction Chemotherapy Can Predict Who Will Benefit From Subsequent Esophagectomy After Chemoradiotherapy for Esophageal Adenocarcinoma. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine*.
22. Fischer, A. H., K. A. Jacobson, J. Rose, and R. Zeller. 2008. Hematoxylin and eosin staining of tissue and cell sections. *CSH protocols* 2008: pdb.prot4986.
23. Thies, S., and R. Langer. 2013. Tumor Regression Grading of Gastrointestinal Carcinomas after Neoadjuvant Treatment. *Frontiers in Oncology* 3: 262.
24. Mandard, A. M., F. Dalibard, J. C. Mandard, J. Marnay, M. Henry-Amar, J. F. Petiot, A. Roussel, J. H. Jacob, P. Segol, G. Samama, and et al. 1994. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. *Cancer* 73: 2680-2686.
25. Chang, F., H. Deere, U. Mahadeva, and S. George. 2008. Histopathologic examination and reporting of esophageal carcinomas following preoperative neoadjuvant therapy:

- practical guidelines and current issues. *American journal of clinical pathology* 129: 252-262.
26. Padhani, A. R., G. Liu, D. M. Koh, T. L. Chenevert, H. C. Thoeny, T. Takahara, A. Dzik-Jurasz, B. D. Ross, M. Van Cauteren, D. Collins, D. A. Hammoud, G. J. Rustin, B. Taouli, and P. L. Choyke. 2009. Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations. *Neoplasia (New York, N.Y.)* 11: 102-125.
 27. van Rossum, P. S., D. V. Fried, L. Zhang, W. L. Hofstetter, M. van Vulpen, G. J. Meijer, L. E. Court, and S. H. Lin. 2016. The Incremental Value of Subjective and Quantitative Assessment of 18F-FDG PET for the Prediction of Pathologic Complete Response to Preoperative Chemoradiotherapy in Esophageal Cancer. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* 57: 691-700.
 28. Xu, Q. G., and J. F. Xian. 2015. Role of quantitative magnetic resonance imaging parameters in the evaluation of treatment response in malignant tumors. *Chinese Medical Journal* 128: 1128-1133.
 29. van Rossum, P. S., A. L. van Lier, M. van Vulpen, O. Reerink, J. J. Lagendijk, S. H. Lin, R. van Hillegersberg, J. P. Ruurda, G. J. Meijer, and I. M. Lips. 2015. Diffusion-weighted magnetic resonance imaging for the prediction of pathologic response to neoadjuvant chemoradiotherapy in esophageal cancer. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 115: 163-170.
 30. Heethuis, S. E., P. S. N. van Rossum, I. M. Lips, L. Goense, F. E. Voncken, O. Reerink, R. van Hillegersberg, J. P. Ruurda, M. E. Philipens, M. van Vulpen, G. J. Meijer, J. J.

- W. Lagendijk, and A. L. H. M. W. van Lier. 2016. Dynamic contrast-enhanced MRI for treatment response assessment in patients with oesophageal cancer receiving neoadjuvant chemoradiotherapy. *Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology* 120: 128-135.
31. Bloch, F. 1946. Nuclear Induction. *Physical Review* 70: 460-474.
 32. Brown, R. W., Y.-C. N. Cheng, E. M. Haacke, M. R. Thompson, and R. Venkatesan. 2014. *Magnetic resonance imaging : physical principles and sequence design*. John Wiley & Sons, Inc., Hoboken, New Jersey.
 33. Mansfield, P., and A. A. Maudsley. 1977. Planar spin imaging by NMR. *Journal of Magnetic Resonance (1969)* 27: 101-119.
 34. Fick, A. 1903. *Gesammelte schriften*. O. Stahel, Würzburg,.
 35. Yankeelov, T., D. R. Pickens, and R. R. Price. 2012. *Quantitative MRI in cancer*. CRC Press, Boca Raton, FL.
 36. Sugahara, T., Y. Korogi, M. Kochi, I. Ikushima, Y. Shigematu, T. Hirai, T. Okuda, L. Liang, Y. Ge, Y. Komohara, Y. Ushio, and M. Takahashi. 1999. Usefulness of diffusion-weighted MRI with echo-planar technique in the evaluation of cellularity in gliomas. *Journal of magnetic resonance imaging : JMRI* 9: 53-60.
 37. Bible, E., F. Dell'Acqua, B. Solanky, A. Balducci, P. Crapo, S. F. Badylak, E. T. Ahrens, and M. Modo. 2012. Non-invasive imaging of transplanted human neural stem cells and ECM scaffold remodeling in the stroke-damaged rat brain by (19)F- and diffusion-MRI. *Biomaterials* 33: 2858-2871.

38. Anderson, A. W., J. Xie, J. Pizzonia, R. A. Bronen, D. D. Spencer, and J. C. Gore. 2000. Effects of cell volume fraction changes on apparent diffusion in human cells. *Magnetic resonance imaging* 18: 689-695.
39. Latour, L. L., K. Svoboda, P. P. Mitra, and C. H. Sotak. 1994. Time-dependent diffusion of water in a biological model system. *Proceedings of the National Academy of Sciences of the United States of America* 91: 1229-1233.
40. Allen, L. M., A. N. Hasso, J. Handwerker, and H. Farid. 2012. Sequence-specific MR imaging findings that are useful in dating ischemic stroke. *Radiographics : a review publication of the Radiological Society of North America, Inc* 32: 1285-1297; discussion 1297-1289.
41. Lim, H. K., J. K. Kim, K. A. Kim, and K. S. Cho. 2009. Prostate cancer: apparent diffusion coefficient map with T2-weighted images for detection--a multireader study. *Radiology* 250: 145-151.
42. Komori, T., I. Narabayashi, K. Matsumura, M. Matsuki, H. Akagi, Y. Ogura, F. Aga, and I. Adachi. 2007. 2-[Fluorine-18]-fluoro-2-deoxy-D-glucose positron emission tomography/computed tomography versus whole-body diffusion-weighted MRI for detection of malignant lesions: initial experience. *Annals of nuclear medicine* 21: 209-215.
43. Valerio, M., C. Zini, D. Fierro, F. Giura, A. Colarieti, A. Giuliani, A. Laghi, C. Catalano, and V. Panebianco. 2016. 3T multiparametric MRI of the prostate: Does intravoxel incoherent motion diffusion imaging have a role in the detection and stratification of prostate cancer in the peripheral zone? *European journal of radiology* 85: 790-794.

44. Sun, Y., T. Tong, S. Cai, R. Bi, C. Xin, and Y. Gu. 2014. Apparent Diffusion Coefficient (ADC) value: a potential imaging biomarker that reflects the biological features of rectal cancer. *PloS one* 9: e109371.
45. Goel, V., P. S. Parihar, A. Parihar, A. K. Goel, K. Waghvani, R. Gupta, and U. Bhutekar. 2016. Accuracy of MRI in Prediction of Tumour Thickness and Nodal Stage in Oral Tongue and Gingivobuccal Cancer With Clinical Correlation and Staging. *Journal of clinical and diagnostic research : JCDR* 10: Tc01-05.
46. Li, X., H. Kang, L. R. Arlinghaus, R. G. Abramson, A. B. Chakravarthy, V. G. Abramson, J. Farley, M. Sanders, and T. E. Yankeelov. 2014. Analyzing Spatial Heterogeneity in DCE- and DW-MRI Parametric Maps to Optimize Prediction of Pathologic Response to Neoadjuvant Chemotherapy in Breast Cancer. *Translational Oncology* 7: 14-22.
47. Sala, E., M. Y. Kataoka, A. N. Priest, A. B. Gill, M. A. McLean, I. Joubert, M. J. Graves, R. A. Crawford, M. Jimenez-Linan, H. M. Earl, C. Hodgkin, J. R. Griffiths, D. J. Lomas, and J. D. Brenton. 2012. Advanced ovarian cancer: multiparametric MR imaging demonstrates response- and metastasis-specific effects. *Radiology* 263: 149-159.
48. Nowosielski, M., W. Recheis, G. Goebel, O. Guler, G. Tinkhauser, H. Kostron, M. Schocke, T. Gotwald, G. Stockhammer, and M. Hutterer. 2011. ADC histograms predict response to anti-angiogenic therapy in patients with recurrent high-grade glioma. *Neuroradiology* 53: 291-302.
49. Torrey, H. C. 1956. Bloch Equations with Diffusion Terms. *Physical Review* 104: 563-565.

50. Stejskal, E. O., and J. E. Tanner. 1965. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *The Journal of Chemical Physics* 42: 288.
51. Le Bihan, D. 2013. Apparent diffusion coefficient and beyond: what diffusion MR imaging can tell us about tissue structure. *Radiology* 268: 318-322.
52. Le Bihan, D., E. Breton, D.ALLEMAND, M. L. Aubin, J. Vignaud, and M. Laval-Jeantet. 1988. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* 168: 497-505.
53. Wang, Y. C., D. Y. Hu, X. M. Hu, Y. Q. Shen, X. Y. Meng, H. Tang, and Z. Li. 2016. Assessing the Early Response of Advanced Cervical Cancer to Neoadjuvant Chemotherapy Using Intravoxel Incoherent Motion Diffusion-weighted Magnetic Resonance Imaging: A Pilot Study. *Chinese Medical Journal* 129: 665-671.
54. Noij, D. P., R. M. Martens, J. T. Marcus, R. de Bree, C. R. Leemans, J. A. Castelijns, M. C. de Jong, and P. de Graaf. 2017. Intravoxel incoherent motion magnetic resonance imaging in head and neck cancer: A systematic review of the diagnostic and prognostic value. *Oral Oncology* 68: 81-91.
55. Cohen, A. D., M. C. Schieke, M. D. Hohenwarter, and K. M. Schmainda. 2015. The effect of low b-values on the intravoxel incoherent motion derived pseudodiffusion parameter in liver. *Magnetic resonance in medicine* 73: 306-311.
56. Jensen, J. H., J. A. Helpert, A. Ramani, H. Lu, and K. Kaczynski. 2005. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic resonance in medicine* 53: 1432-1440.

57. Christou, A., A. Ghiatas, D. Priovolos, K. Veliou, and H. Bougias. 2017. Accuracy of diffusion kurtosis imaging in characterization of breast lesions. *The British journal of radiology* 90: 20160873.
58. Tonoyan, A. S., I. N. Pronin, D. I. Pitsheauri, L. V. Shishkina, L. M. Fadeeva, E. L. Pogosbekyan, N. E. Zakharova, E. I. Shults, N. V. Khachanova, V. N. Kornienko, and A. A. Potapov. 2015. [A correlation between diffusion kurtosis imaging and the proliferative activity of brain glioma]. *Zhurnal voprosy neirokhirurgii imeni N. N. Burdenko* 79: 5-14.
59. Li, H. M., S. H. Zhao, J. W. Qiang, G. F. Zhang, F. Feng, F. H. Ma, Y. A. Li, and W. Y. Gu. 2017. Diffusion kurtosis imaging for differentiating borderline from malignant epithelial ovarian tumors: A correlation with Ki-67 expression. *Journal of magnetic resonance imaging : JMRI*.
60. Mansfield, P. 1977. Multi-planar image formation using NMR spin echoes. *Journal of Physics C: Solid State Physics* 10: L55.
61. Pruessmann, K. P., M. Weiger, M. B. Scheidegger, and P. Boesiger. 1999. SENSE: sensitivity encoding for fast MRI. *Magnetic resonance in medicine* 42: 952-962.
62. Griswold, M. A., P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase. 2002. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic resonance in medicine* 47: 1202-1210.
63. Saritas, E. U., C. H. Cunningham, J. H. Lee, E. T. Han, and D. G. Nishimura. 2008. DWI of the spinal cord with reduced FOV single-shot EPI. *Magnetic resonance in medicine* 60: 468-473.

64. Suo, S., N. Lin, H. Wang, L. Zhang, R. Wang, S. Zhang, J. Hua, and J. Xu. 2015. Intravoxel incoherent motion diffusion-weighted MR imaging of breast cancer at 3.0 tesla: Comparison of different curve-fitting methods. *Journal of magnetic resonance imaging : JMRI* 42: 362-370.
65. Sigmund, E. E., G. Y. Cho, S. Kim, M. Finn, M. Moccaldi, J. H. Jensen, D. K. Sodickson, J. D. Goldberg, S. Formenti, and L. Moy. 2011. Intravoxel incoherent motion imaging of tumor microenvironment in locally advanced breast cancer. *Magnetic resonance in medicine* 65: 1437-1447.
66. Yushkevich, P. A., J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig. 2006. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage* 31: 1116-1128.
67. Lei, J., Y. Tian, S. C. Zhu, Q. Han, Y. Wei, S. Yang, and D. P. Shi. 2015. Preliminary study of IVIM-DWI and DCE-MRI in early diagnosis of esophageal cancer. *European review for medical and pharmacological sciences* 19: 3345-3350.
68. Cho, S. H., G. C. Kim, Y. J. Jang, H. Ryeom, H. J. Kim, K. M. Shin, J. S. Park, G. S. Choi, and S. H. Kim. 2015. Locally advanced rectal cancer: post-chemoradiotherapy ADC histogram analysis for predicting a complete response. *Acta radiologica (Stockholm, Sweden : 1987)* 56: 1042-1050.
69. Moriya, T., K. Saito, Y. Tajima, T. L. Harada, Y. Araki, K. Sugimoto, and K. Tokuuye. 2017. 3D analysis of apparent diffusion coefficient histograms in hepatocellular carcinoma: correlation with histological grade. *Cancer Imaging* 17.
70. Ueno, Y., R. Lisbona, T. Tamada, A. Alaref, K. Sugimura, and C. Reinhold. 2017. Comparison of FDG PET metabolic tumour volume versus ADC histogram: prognostic

- value of tumour treatment response and survival in patients with locally advanced uterine cervical cancer. *The British journal of radiology* 90: 20170035.
71. Cohen, J. 1968. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological bulletin* 70: 213-220.
 72. Lambregts, D. M., G. L. Beets, M. Maas, L. Curvo-Semedo, A. G. Kessels, T. Thywissen, and R. G. Beets-Tan. 2011. Tumour ADC measurements in rectal cancer: effect of ROI methods on ADC values and interobserver variability. *European radiology* 21: 2567-2574.
 73. van Heeswijk, M. M., D. M. Lambregts, J. J. van Griethuysen, S. Oei, S. X. Rao, C. A. de Graaff, R. F. Vliegen, G. L. Beets, N. Papanikolaou, and R. G. Beets-Tan. 2016. Automated and Semiautomated Segmentation of Rectal Tumor Volumes on Diffusion-Weighted MRI: Can It Replace Manual Volumetry? *International journal of radiation oncology, biology, physics* 94: 824-831.
 74. Chen, L., P. Bentley, and D. Rueckert. 2017. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks. *NeuroImage : Clinical* 15: 633-643.
 75. Wu, R., S.-T. Suo, L.-M. Wu, Q.-Y. Yao, H.-X. Gong, and J.-R. Xu. 2017. Assessment of chemotherapy response in non-Hodgkin lymphoma involving the neck utilizing diffusion kurtosis imaging: a preliminary study. *Diagnostic and Interventional Radiology* 23: 245-249.
 76. Wu, D., G. Li, J. Zhang, S. Chang, J. Hu, and Y. Dai. 2014. Characterization of Breast Tumors Using Diffusion Kurtosis Imaging (DKI). *PloS one* 9: e113240.

77. Pang, H., Y. Ren, X. Dang, X. Feng, Z. Yao, J. Wu, C. Yao, N. Di, D. C. Ghinda, and Y. Zhang. 2016. Diffusional kurtosis imaging for differentiating between high-grade glioma and primary central nervous system lymphoma. *Journal of magnetic resonance imaging : JMRI* 44: 30-40.

7 Vita

Benjamin Charles Musall was born in Augusta, Georgia on May 3, 1993, the son of Jacqueline B. Musall and John C. Musall. After completing his high school education at Lakeside High School, Evans, GA, he entered the Georgia Institute of Technology in Atlanta, GA. He received the degree of Bachelor of Science with a major in Nuclear and Radiological Engineering from Georgia Institute of Technology in May, 2015. In August of 2015 he entered The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences.